Refinement of quantitative MRI as an outcome measure in inherited neuropathies

Dr Matthew Rahim Broumand Evans

UCL, Institute of Neurology

Primary Supervisor: Professor Mary M Reilly
Subsidiary Supervisor: Dr John S Thornton

Thesis submitted for the degree of Doctor of Philosophy

University College London

2020

Declaration

I, Matthew Rahim Broumand Evans confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Recent acceleration in discovery of potential drug treatments for inherited neuromuscular diseases (NMD) heralds the urgent need for scientifically sound clinical trials. Given the rarity and slow progression of most of these conditions, simply increasing sample size, or extending trial duration to increase study power, are not viable options.

Trials in NMD are in desperate need of highly responsive outcome measures. Lower limb muscle quantitative MRI (qMRI) allows non-invasive assessment of sequelae of nerve and muscle disease. It has recently been shown to be reliable, valid and responsive in a number of NMD, but further refinement is vital in order to ensure the ability to undertake rigorous and meaningful clinical trials in small cohorts of patients with rare and slowly progressive diseases, over short durations.

This thesis aims to examine and improve qMRI responsiveness for application in trials for NMD. Two separate inherited neuropathies have been studied. In Charcot-Marie-Tooth disease type 1A (CMT1A), extended follow up has revealed that qMRI has large internal responsiveness over five years, measuring significant fat fraction (FF) change of 0.7 ± 0.6%/year with standardised response mean (SRM) of 1.07 over 5 years. Excellent validity of qMRI-determined FF is confirmed by strong correlation with clinical measures at baseline, and longitudinal validity is demonstrated for the first time in CMT1A with strong correlation with changes in CMT examination score and remaining muscle area. In Hereditary Sensory Neuropathy type 1, qMRI measures significant FF change at all anatomical levels examined, with large responsiveness at calf levels (distal calf FF change 2.2 ± 2.7%, SRM=0.83; proximal calf FF change 2.6 ± 3.0%, SRM=0.84 over 12 months). In both diseases, significant muscle fat gradients are shown to exist with the potential to devastate or enhance longitudinal analysis. FF change is predicted by baseline FF and other MRI measures in both diseases. Improvement in qMRI responsiveness is demonstrated through a host of evidence-based manipulations aimed at maximising and homogenising mean FF change.

Quantitative MRI determined FF is shown to be highly responsive as an outcome measure in two different inherited neuropathies, and is ready to be used as a primary outcome measure in drug trials for rare neuromuscular diseases.

Impact statement

Patients with neuromuscular disease are a large cohort suffering significant morbidity and mortality. For the majority, disease modifying treatments are still lacking. However, novel treatments are rapidly emerging, and to be robustly assessed in randomised clinical trials, outcome measures are critical. If an outcome measure is inadequately responsive, clinical trials may be futile: almost certainly resulting in type 2 error, wasting precious clinical trial resources. Quantitative muscle MRI (qMRI) is an emerging outcome measure which has relatively recently been shown to be reliable, valid and highly responsive in patients with neuromuscular disease.

This thesis builds on previous neuroimaging research, providing a detailed evidence-based approach to the implementation of qMRI as an outcome measure in future clinical research in two peripheral nerve diseases: both of which demonstrate slow clinical progression, one of which has extremely low patient numbers meaning high outcome measure responsiveness is crucial.

The central finding in this thesis details the practical application of the 3-point Dixon MRI method of lower limb muscle fat quantification as an outcome measure for these neuromuscular diseases, but also provides more generally applicable lessons. We have demonstrated that with refinements, which are described in detail in this thesis, future national and international clinical trials can achieve equivalent study power whilst enrolling fewer patients over shorter trial duration, compared with previously used primary outcome measures. This provides immediate beneficial impact not only to the quality of the clinical research within academia, but equally importantly to trial participants and potentially to their future wellbeing and that of their families.

The impact of these findings are not limited to the two diseases which have been studied in this thesis. Although individual neuromuscular diseases vary in behaviour with respect to rate of progression, muscles involved and other factors, there is a common histopathological endpoint of muscle fat replacement by intramuscular fat as disease progresses, which is harnessed by 3-point Dixon MRI, allowing these two diseases to be used as representative diseases for the wider field of inherited and acquired neuromuscular disease.

Acknowledgements

The work detailed in this thesis would not be possible without the involvement and support of many people. I would like to thank them here.

I thank my primary supervisor Professor Mary Reilly, and secondary supervisor Dr John Thornton, for their incomparable expertise and guidance, and for their kindness, compassion, encouragement and patience. Professor Reilly, amongst countless things, for instilling in me an excitement for clinical neurology and research. Dr Thornton for his humility, calming influence, and the time spent discussing all aspects of the studies whenever needed.

A special thankyou to my good friend Dr Jasper Morrow, for invaluable guidance, discussions, proofreading and cycling trips.

Professor Yousry for his guidance when needed. Dr Chris Sinclair and Dr Stephen Wastling for MRI physics expertise. Dr Arne Fischmann for grading the baseline STIR and Mercuri images for the CMT1A longitudinal natural history study.

My wife, Hoda, for her love and extraordinary patience. Our beautiful daughters, Juliet and Elizabeth, made an entrance in November '18 and have kept us happily busy. I estimate that we have changed ~5000 nappies in 21 months, a big number but barely a whisker on the number of regions of interest defined for this thesis.

My parents, Richard and Faeghe. My Dad passed away in December '15, part way into the data collection, which was devastating for our family. His was an illustrious life of service to humanity, which I try to honour in my daily life. I see his gentle soul reflected in my daughters. My Mum for limitless and unwavering love and support. My brothers, James, Sia and Rikky, and sisters-in-law Maryam and Layli, for their humour and encouragement. James in particular for meticulous proofreading. My Aunty Libby for her kindness, inspiring stories and friendship. Mitra joon and Almas joon, for their kindness and sacrifice, affording me time to write when Hoda was pregnant, and once the babies had arrived.

My neurology colleagues and friends at Guy's and St Thomas', a truly joyful place to work.

And of course, all the participants without whom there would be no data, and for whom I hope this thesis effects a material change.

Table of contents

List of tables

List of figures

List of equations

List of abbreviations

AM	Adductor Magnus
BF	Biceps Femoris
CI	Confidence interval
CIDP	Chronic inflammatory demyelinating polyradiculoneuropathy
CMAP	Compound muscle action potential
CMT	Charcot-Marie-Tooth disease
CMT1A	CMT type 1A
CMTES	CMT Examination Score
CMTES-LL	CTMES-lower limb components
CMTNS	CMT Neuropathy Score
CMTSS	CMT Symptom Score
CMTX	CMT type X
CMT-FOM	CMT Functional Outcome Measure
CPMG	Carr-Purcell-Meiboom-Gill
CSA	Cross-sectional area
CT	Computerised tomography
DMD	Duchenne muscular dystrophy
DNM2	Dynamin 2
DTI	Diffusion Tensor Imaging
E	Expected efficacy of treatment
EHL	Extensor Hallucis Longus muscle
EMA	European Medicines Agency
EMCL	Extramyocellular lipid
ERT	Enzyme replacement therapy
FA	Fractional anisotropy
FDA	Federal Drug Agency
FF	Fat fraction
FSHD	Facioscapulohumeral dystrophy
HSN	Hereditary sensory neuropathy
HSN1	HSN type 1
IBM	Inclusion body myositis
ICC	Intra-class correlation coefficient

IDEAL	Iterative decomposition of water and fat with echo asymmetry and least-squares estimation
LG	Lateral head of gastrocnemius
LGMD	Limb girdle muscular dystrophy
MG	Medial head of gastrocnemius
MND	Motor neurone disease
MRC	Medical Research Council
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
MTR	Magnetisation Transfer Ratio
NMD	Neuromuscular Disease
OPMD	Oculopharyngeal muscular dystrophy
Pearson ICC	Pearson Inter-class correlation coefficient
PMP22	Peripheral myelin protein 22
qMRI	Quantitative MRI
RF	Rectus femoris
RMA	Remaining muscle area
ROI	Region of interest
SBMA	Spinal and bulbar muscular atrophy
SF-36	Short Form 36
sIBM	Sporadic IBM
SM	Semimembranosus
SMA	Spinal muscular atrophy
ST	Semitendinosus
Sol	Soleus
SPTLC1	Serine palmitoyltransferase, long chain base subunit 1
SPTLC2	Serine palmitoyltransferase, long chain base subunit 2
SRM	Standardised Response Mean
STIR	Short Tau Inversion Recovery sequence
TMPRSS5	Transmembrane protease serine 5
T1	Longitudinal T1 recovery time

T1w	T1-weighted sequence
T2	Transverse T2 relaxation time
TA	Tibialis anterior
TE	Echo time
TR	Repetition time
VI	Vastus intermedius
VL	Vastus Lateralis
VM	Vastus Medialis
3T	3 tesla
6MWT	six minute walk test

1 Background

1.1 Introduction

Neuromuscular disorders have historically been seen as the cinderella subspecialty of neurology due to the availability of few treatments. However, with increasing knowledge of the pathogenesis of neuromuscular diseases (NMD), the advent of next-generation sequencing (Walsh et al., 2017) and advances in molecular genetics, precise genetic diagnoses are now available for many NMD, with subsequent discovery of potential disease-specific therapies. The era of novel drug treatment trials has arrived, with ground-breaking discoveries in several inherited NMD such as Duchenne muscular dystrophy (Blat and Blat, 2015; Bushby et al., 2014; Buyse et al., 2015; Cirak et al., 2011; Syed, 2016) and Spinal muscular atrophy (Finkel et al., 2017; Mercuri et al., 2018), promising results in several preclinical and clinical studies in other NMD (Arechavala-Gomeza et al., 2012; Attarian et al., 2014; D'Ydewalle et al., 2011; Patzkó et al., 2012; Sereda et al., 2003) and strong scientific rationale for further treatments close at hand (Jerath and Shy, 2015; Rossor et al., 2013). All novel compounds must be rigorously assessed in clinical trials.

The relative rarity and slow progression of NMD combined with the fact that new medications may be expected, initially at least, to slow or halt disease progression rather than reverse it, means that it is often challenging to adequately power a trial to assess drug effect. The usual methods of boosting study power, including by increasing sample size, are not viable options.

A potent alternate method by which study power can be increased is by using a clinically meaningful and highly responsive outcome measure. This is where quantitative MRI (qMRI) has come to the fore, with recently proven reliability, validity and large responsiveness, three critical factors which are lacking in various combinations in currently used NMD outcome measures. However, further improvements are urgently needed.

This thesis provides a solid evidence base for refinement of qMRI as an outcome measure for NMD. Chapter 3 details results from extended follow up of our Centre's comprehensive natural history study in Charcot-Marie-Tooth disease type 1A (CMT1A), the first year of which has been previously published (Morrow et al., 2016). A battery of outcome measures including qMRI are analysed longitudinally, showing that MRI-determined calf muscle fat fraction (FF) is highly internally responsive over a 5-year period (FF change 0.7 ± 0.6%/year, p=0.002, standardised response mean (SRM)=1.07) and is for the first time shown to be longitudinally valid by correlation with changes in functional measures (r_s=0.71, p=0.005 with change in CMT

examination score (CMTES)). Chapter 4 examines calf muscle fat gradients in a subset of patients with CMT1A, for the first time quantifying their degree and distribution, drawing attention to the potentially devastating effects of even small errors in longitudinal MRI axial slice placement on qMRI responsiveness and study results. Chapter 5 examines lower limb qMRI measures longitudinally at three separate anatomical levels in a second inherited neuropathy, hereditary sensory neuropathy type 1 (HSN1) examining the distribution of fatty atrophy and detailing methods by which qMRI responsiveness can be further increased. Finally, based on these chapters, in Chapter 6 a detailed and evidence-based suggestions on the place and use of qMRI in upcoming clinical trials in inherited NMD are presented.

Literature pertinent to each study is presented in the relevant chapter. Chapter 1 gives a clinical background of the two inherited NMD being studied: CMT1A and HSN1, reviews principles of outcome measures as applied to NMD, and gives an introduction to qualitative and qMRI techniques used in this thesis, whilst Chapter 2 details methods for all studies.

1.2 Determinants of a good outcome measure

Hastily embarking on clinical trials without appropriate thought to outcome measure characteristics risks leaving trials underpowered to detect small real changes over short trial duration – a waste of time, money and also perhaps unethical (Halpern, Karlawish and Berlin, 2002). Here I examine the three central pillars of a good outcome measure.

1.2.1 Reliability

Reliability of an outcome measure ensures that if conditions remain identical and no real change has occurred, the outcome measure returns the same result each time (Keszei, Novak and Streiner, 2010). Reliability is not a stable characteristic in all settings, being dependent not only on the outcome measure itself, but also on the attributes of the population it is testing, and the trial design. As an example of the effect of the population being studied, scales which are accurate to within 0.5kg would be adequately reliable for assessing weight in adults, but the same scales would not be adequately reliable in new-borns. Reliability can be expressed by the intra-class correlation coefficient which represents the true intersubject variation divided by the intersubject variation plus the error (Equation 1-1). A reliability coefficient of 1.0 indicates that all differences between scores represent real differences between individuals, whereas a reliability coefficient of 0.25 indicates that 25% of the variance

is due to true difference, with the rest due to measurement error. In general, reliability coefficients above 0.75 are considered good. Improving outcome measure reliability increases study power by minimising measurement error which might obscure genuine change due to treatment. Reliability is a foundation block for outcome measure responsiveness.

Equation 1-1 – Reliability Formula

$$\text{Reliability} = \frac{\text{intersubject variation (true difference)}}{\text{intersubject variation} + \text{random variation}}$$

1.2.2 Validity

Reliability is essential for an outcome measure but clearly inadequate on its own. Validity of an outcome measure is defined as the degree to which it measures what it is intended to measure (Streiner and Norman, 2008). Validity is not a characteristic of the instrument itself, but rather relates to a certain question in a certain population at a certain time. Validity of an outcome measure is multifaceted, and can be examined in terms of criterion validity – comparing its results to that of a gold standard, and construct validity which refers to the measure's ability to tightly reflect that which is meaningful to the patient, and of interest to the researcher – the true endpoint.

Direct measures of morbidity and mortality are the preferred outcome measures, however in some cases a surrogate outcome measure is acceptable (e.g. reduction in blood pressure in place of mortality) but this must correlate with the true end-point (Prentice, 1989). Muscle MRI is interesting as a surrogate outcome measure as it does not predict mortality or morbidity, but instead measures current morbidity, perhaps more reliably and sensitively than other measures.

1.2.3 Responsiveness

Responsiveness is a crucial component of a good outcome measure, referring to the sensitivity of an outcome measure to detect or capture change over time if it occurs (Guyatt, Walter and Norman, 1987; Husted et al., 2000; Kirshner and Guyatt, 1985). The responsiveness of an outcome measure is reliant on its reliability, as if the measure is not reliable, the error in measurement may obscure any real change that has occurred. Measures which aim to place results into a limited number of categories (e.g. ordinal data) are usually not responsive, as a large change is often needed to move categories, as is seen for example with the MRC scoring system for assessing muscle strength (Table 2-1). Similarly, an outcome measure will not be responsive if

it cannot discriminate degrees of difference at either end of the measurement scale (i.e. affected by ceiling and/or floor effect).

There are many ways in which to express sensitivity to change (Lehman and Velozo, 2010), one of which is by the SRM (Guyatt, Walter and Norman, 1987, Liang, Fossel and Larson, 1990) – Equation 1-2, which is also known as Cohen's d. It is the currently favoured measure of responsiveness used in the medical literature, and is the measure that is used in this thesis.

The SRM, is one of the key determinants of study power, and is calculated as the mean change in measurement over the duration of a study, divided by the standard deviation of that change. It can be clearly seen from Equation 1-2, that a measure which has a high level of variability in change scores in relation to the mean change, will have a small SRM. The SRM is particularly useful as it removes the dependence on sample size. The magnitude of the SRM is the key factor: by Cohen's rule of thumb, an outcome measure with SRM <0.2 is considered to have minimal responsiveness, between 0.2 and 0.5 small responsiveness, 0.5-0.8 moderate responsiveness and >0.8 large responsiveness (Beaton, Hogg-Johnson and Bombardier, 1997; Liang, Fossel and Larson, 1990; Piscosquito et al., 2015). The SRM has an inverse square relation to required sample size for a stated statistical power which is expressed by Lehr's formula (Equation 1-3), and from this equation, sample size in each arm of a trial can be calculated.

Equation 1-2 – Standardised response mean

$$\mathrm{SRM} = \frac{(\text{mean change})}{(\text{standard deviation of change})}$$

SRM=standardised response mean

For example, if a medication aims to reduce disease progression by 50%, and an outcome measure with an SRM of 0.50 is used, to detect change in a trial powered at 0.8, $p<0.05$, the number of participants in each arm of the trial would be $16/(0.5 \times 0.5)^2 = 256$, whereas if the SRM of the outcome measure was 1.0, only 64 participants would be needed in each arm. One can appreciate the importance of augmenting outcome measure responsiveness as one method by which N can be reduced, for a certain fixed medication effect. The other equally powerful way of reducing N of course is to develop more effective treatments.

Equation 1-3 – Lehr's formula

$$N = \frac{16}{(SRM \times E)^2}$$

SRM=standardised response mean, N=number required in each arm for a study of 80% power with significance level p<0.05, E=expected efficacy of treatment compared with natural history

A poorly responsive outcome measure leaves a study at risk of accepting a false null hypothesis. This is particularly important in slowly progressive diseases such as CMT1A, in which expected clinical change over the duration of a typical clinical trial may be relatively small.

1.2.4 Minimally clinically significant or important difference

Even if an outcome measure is reliable, valid and highly responsive, it is important to know the minimum clinically significant or important difference of what is being measured (Copay et al., 2007; Jaeschke, Singer and Guyatt 1989), which can be defined as 'the smallest difference in score in the domain of interest, that a patient perceives as beneficial and which would mandate a change in the patient's management' (Jaeschke, Singer and Guyatt 1989). It is well recognised for example that medications can improve physiological measurement without any perceivable improvement in how a patient feels.

Changes in instrument score must be translated into a clinically meaningful effect. This is often done intuitively if clinicians are familiar with the instrument being used. However as this does not apply to qMRI, it is important to include overall disease severity scores and other measures in clinical trials, to confirm the longitudinal validity of change in newly developed outcome measures for which the minimum clinically significant difference is not yet known.[1]

[1] It is noted that some authors differentiate 'sensitivity to change' as meaning change regardless of clinical meaning, from 'responsiveness' indicating that change is clinically meaningful (exceeding variation attributable to chance). In this thesis, responsiveness refers simply to the ability to detect change over time, whereas longitudinal validity is used to refer to added meaningful change over time.

1.3 Outcome measures in inherited neuromuscular diseases

The difficulty identified with many of the outcome measures used in NMD to date is that although they may be reliable and valid, adequate responsiveness is lacking. Here, I look critically at a list of outcome measures commonly used in NMD research, with particular attention to their responsiveness.

1.3.1 CMT neuropathy score and sub scores

The CMT neuropathy score (CMTNS) is the only CMT specific scale in adults, and is the benchmark against which other outcome measures are compared in clinical trials of patients with CMT. It is a composite scale taking into account patient signs and symptoms as well as results of limited electrophysiology. The CMTNS has been shown to be reliable and a valid marker of disease severity (Shy et al., 2005). Early studies showed the CMTNS to be responsive with worsening by 0.68 points/year over a 24-month period in a cohort of 72 patients (Shy et al., 2008), but there have been variable results in more recent studies. Annual change in CMTNS in CMT1A studies has varied between −0.92 to +1.0 point/year and its responsiveness has been small to moderate at best.

1.3.1.1 Vitamin C trials in CMT1A

On the recommendation of the 2005 ENMC International Workshop (Reilly, de Jonghe and Pareyson, 2006), the CMTNS was used as the primary outcome measure in the negative ascorbic acid drug trials (Burns et al., 2009; J. et al., 2009; Lewis et al., 2013; Pareyson et al., 2011; Verhamme et al., 2009). A Cochrane review of the five randomised controlled clinical trials (Gess et al., 2015), which together covered a total of 622 adult patients with CMT1A treated with vitamin C (1-4g daily) and compared with placebo, determined that there was no evidence that ascorbic acid improved the course of CMT1A in adults or children. Importantly, the review demonstrated that the outcome measures used in these five trials showed only small change over the 1-2 year study periods, and recommended that longer study duration should be considered, and outcome parameters more sensitive to change over time be designed and validated for future studies in CMT. The outcome measures used in these five trials included change in CMTNS as the primary outcome measure (either at 12 or 24 months) in all studies, alongside a battery of secondary outcome measures including change in disability by various validated scales, change in ulnar, median or peroneal compound muscle action potential (CMAP), change in sensory impairment, change in muscle strength measures by dynamometer or vigorimeter, and change in quality of life by a number of validated scores including Short Form 36

(SF-36). The mean overall change in CMTNS was 0.37 points/year; 95% CI -0.83 to 0.09. In their trial, Lewis et al. found that the CMTNS paradoxically improved over two years in both treatment (-0.21 points/year) and placebo (-0.92 points/year) arms (Lewis et al., 2013). The CMTES showed a significant benefit for ascorbic acid in post hoc analysis in one trial (Micallef et al., 2009), but the meta-analysis of all three trials that reported this outcome measure found no significant change.

Further, Piscosquito et al. assessed responsiveness of all outcome measures used in the Italian-UK ascorbic acid CMT1A randomised control trial (Pareyson et al., 2011), and found little change in outcome measure scores over two years. The primary outcome measure: CMT Neuropathy Score version 1 (CMTNSv1), showed mean worsening of 0.17 points/year with minimal responsiveness (SRM 0.13). The CMTES declined by 0.19 points/year with SRM of 0.17 and the CMTNS-signs an SRM of 0.19. CMTNS-neurophysiology showed no significant change (Piscosquito et al., 2015).

Since the disappointments in the ascorbic acid trials, the CMTNS has been modified: CMT Neuropathy Score version 2 (CMTNSv2) is a Rasch weighted version, aimed at improving its sensitivity to change by addressing floor and ceiling effect (Murphy et al., 2011; Sadjadi et al., 2014) but has had limited effect on responsiveness (Fridman et al., 2015, 2020). The CMT Functional Outcome Measure (CMT-FOM) has also been developed: a performance based assessment measuring construct functional ability, developed for use in clinical trials in CMT1A. Items have been chosen by expert opinion and on the basis of reliability/validity and responsiveness. It has shown moderate correlation with the CMTES, and is reliable and shows content validity, though further assessment is still warranted (Eichinger et al., 2018; Eichinger et al., 2017).

More recently, a recent randomised double-blinded placebo-controlled trial in patients with HSN1 using oral L-serine supplementation (Fridman et al., 2019) did not show a significant difference in the primary outcome (change in CMTNS at 1 year) between treatment and placebo groups. CMTNS declined by 1.5 units relative to placebo, 95% CI −2.8 to 0.1, p=0.03 in the treatment group whereas the placebo group experienced a mean increase in CMTNS of 1.1 point (±0.53, p=0.04). The CMTES showed similar results to CMTNS (relative change of −1.2 points, 95% CI −2.4 to 0.0, p = 0.05 at 1 year, and 0.83 points, 95% CI −2.7 to 1.0, p = 0.37 at 2 years). The authors comment that the discriminatory power of the primary endpoint was limited by small sample size. Our Centre's study in HSN1 demonstrated a change in CMTNS of 0.23

points/year over 12 months, 95% CI −0.24 to 0.68, with minimal responsiveness (Kugathasan et al., 2019).

1.3.2 Handheld and fixed myometry

The principal difficulty with manual muscle testing as an outcome measure is that much of the assessment relies on effort, not only on the part of the patient, but also the examiner, resulting in poor reliability. This applied both to handheld and fixed dynamometry. Despite these disadvantages, in the Italian-UK ascorbic acid trial, handheld myometric assessments of handgrip and foot dorsiflexion strength were the most responsive outcome measures, though still only showing small to moderate responsiveness (SRM -0.31 and -0.38 respectively). Kugathasan et al. found ankle plantarflexion, inversion and eversion by fixed myometry to change significantly over 12 months in HSN1, though these measures showed only small and moderate responsive (SRM of -0.62, -0.33 and -0.40 respectively).

1.3.3 Neurophysiology

Although neurophysiology is crucial as a diagnostic tool, reliability is poor (Salerno et al., 1999; Schuhfried et al., 2017), with large ranges in normative data. This high variability and poor sensitivity to change (Verhamme et al., 2009) renders neurophysiology a poor candidate for use as an outcome measure. As an example, neurophysiology was used as the primary outcome measure in the paediatric vitamin C CMT study (Burns et al., 2009) with minimal responsiveness in the placebo arm. Verhamme et al. showed no deterioration in neurophysiology in their 5-year natural history study in CMT1A (Verhamme et al., 2009)

1.3.4 Gait analysis

Lencioni et al. examined gait in detail (kinematic and kinetic parameters) in 71 CMT patients at baseline and after 28.9 ± 9.5 months. Overall, several parameters showed moderate responsiveness and when patients were divided into groups according to disease severity, some measures reached high responsiveness (SRM >0.80). Kinematic parameters were more responsive in the minimally affected group, whereas kinetic parameters were more responsive in the most severely affected one. The authors concluded that biomechanical parameters showed moderate-to-high responsiveness and may represent suitable outcome measures for CMT. Of interest, CMTES showed moderate responsiveness (SRM 0.53) in the minimally affected group (Lencioni et al., 2017).

1.3.5 Other outcome measures

There are of course myriad other outcome measures that have been used in patients with CMT, including 10-metre timed walk, nine-hole peg test (SRM 0.28 in Italian-UK ascorbic acid study), Overall Neuropathy Limitation Scale, the SF-36 questionnaire and the Visual Analogue Scale, all of which have been proven minimally or not at all responsive due to no significant change over time. Patient reported outcome measures are unlikely to have the sensitivity to change necessary for a primary outcome measure. Investigations that are used in diagnosis, such as blood tests and muscle and skin biopsy, lack reliability and validity, and do not currently have a role as outcome measures. MRC score is assessed on a descriptive non-linear ordinal scale – with no intrinsic value – is subjective, poorly reproducible and with different meaning in each muscle for the same score.

1.3.6 Composite outcome measures

Composite outcome measures are often used as primary outcome measures in large randomised controlled trials in order to increase statistical efficiency, enabling trials to be smaller and shorter in duration. This is of particular interest for NMD. Cross-sectionally in CMT1A, Mannil et al. (Mannil et al., 2014) concluded that a combination of certain aspects of the CMTNS (ulnar CMAP, leg and arm motor symptoms, leg strength and sensory symptoms), alongside 10-metre timed walk, nine-hole-peg test and ankle dorsiflexion by myometry displayed the strongest power in discriminating between disease severities in patients with CMT1A. This study did not include assessment by qMRI.

1.3.7 New progression biomarkers in CMT1A

In a study of 311 patients with CMT1A, Fledrich et al. found that cutaneous transcripts differentiated disease severity with high sensitivity and specificity, and that in a cohort who had repeat skin biopsy after 2-3 years, the change in cutaneous transcripts correlated with disease progression (Fledrich et al., 2017). On the basis that neurofilament is a byproduct of axonal breakdown, Rossor and colleagues examined plasma neurofilament heavy-chain concentration as a potential biomarker in CMT, finding no significant difference in plasma neurofilament heavy-chain concentrations between CMT patients and healthy controls at baseline, and no significant difference between group over 12 months (Rossor et al., 2016). Subsequently, Sandelius et al. performed a cross-sectional study in the same cohort of 75 patients with CMT and 67 matched controls, comparing disease severity measured by Rasch modified CMTES and CMTNS with plasma neurofilament light-chain concentration. The authors found

that neurofilament light-chain concentration was significantly higher in CMT patients (including CMT subsets of CMT1A, CMTX and HSN1) compared with healthy controls ($p<0.0001$) and correlated with disease severity as measured by Rasch modified CMTES ($r=0.43$, $p<0.0001$) and CMTNS ($r=0.37$, $p=0.044$). This result suggests that plasma neurofilament light-chain concentration is promising as a biomarker of disease activity (Sandelius et al., 2018). The difference between these two studies in the same population was thought to be attributable to formation of neurofilament heavy-chain aggregates in the former study with falsely low results (Lu et al., 2011). Most recently, Wang et al. (Wang et al., 2020) reported on the first Schwann cell specific protein: Transmembrane protease serine 5 (TMPRSS5) which was found to be elevated 2.07 fold in two independent cohorts of CMT1A patients compared with healthy controls ($p<0.0001$), therefore providing a potential disease biomarker for clinical trials. TMPRSS5 however was not significantly correlated with disease score, nerve conduction velocities or with age.

1.4 Lower limb MRI as an outcome measure for neuromuscular disease

1.4.1 Qualitative MRI

Use of MRI in clinical neuromuscular medicine has a long and well-established role for the purposes of diagnosis, based on thigh and calf pattern recognition and fat/water distribution. MRI sequences typically used for this purpose are conventional non-contrast axial T1-weighted imaging – for assessment of muscle fat content, and T2 weighted and/or short tau inversion recovery (STIR) imaging – for assessment of water content (Liu, Chino and Ishihara 1993).

In the overall neuromuscular diagnostic algorithm, whereas MRI has previously often been undertaken to guide genetic testing (Chan and Liu, 2002; Degardin et al., 2010; Stramare et al., 2010; Wattjes, Kley and Fischer, 2010), its current role is in a state of flux, given the recent advances in genetics, allowing rapid and relatively cheap screening of multiple genes.

Over the past two decades there has been rapid expansion in the number of publications reporting qualitative changes and patterns of muscle fat involvement in a host of inherited, mostly muscle diseases (Bönnemann et al., 2014; Bugiardini et al., 2018; Dahlqvist, Oestergaard et al., 2019; Del Grande et al., 2011; Diaz-Manera et al., 2018; Fanin et al., 2007; Faridian-Aragh et al., 2014; Feng et al., 2018; Jungbluth, 2017; Lodi et al., 1997; Mercuri et al., 2005, 2007, 2002; Paradas et al.,

2010; Regula et al., 2016; Sookhoo et al., 2007; Straub, Carlier and Mercuri, 2012; Tasca et al., 2012; ten Dam et al., 2016; Tomasová Studýnková et al., 2007; Wattjes, Kley and Fischer, 2010). Some patterns of muscle fat infiltration are considered almost pathognomic of a certain NMD as in Bethlem myopathy (Morrow et al., 2013) and Limb Girdle Muscular Dystrophy (LGMD) D1 (Sandell et al., 2013), whereas others are less specific, as is seen for many of the other limb girdle muscular dystrophies. With respect to CMT1A, there are papers reporting various patterns of individual involvement, however there is currently no role for MRI in diagnosis of CMT.

For the purpose of more precise assessment, several semi-quantitative scales, usually with four, five or six grades, which correlate well with functional measures (Liu, Chino and Ishihara 1993) have been suggested including Mercuri (Mercuri et al., 2002) – Table 1-1, Fischer (Fischer et al., 2008) – Table 1-2, and Goutallier (Slabaugh et al., 2012) scales. Similarly, scales for assessment of STIR hyperintensity have been suggested (Poliachik et al., 2012) though these have not been validated. Although very useful for descriptive and diagnostic assessment, given that scoring is subjective and highly operator dependent, these methods show great inter- and intra-observer variability and results are poorly reproducible (Alizai et al., 2012; Shahabpour et al., 2008). As an example, in their natural history study of 38 patients with LGMD2I, Willis et al. found that for the qualitative MRI scoring (by Mercuri grading), two observers were in agreement for only 65% of muscle groups scored, and that agreement was highest at the extremes of the grades and least in the middle, where there was only agreement for 39% of grade 2b muscles (Willis et al., 2014).

In addition, the limited number of grades (and the fact that grading is ordinal rather than continuous) translates to reduced sensitivity to even moderate change in FF over time (poor responsiveness). For example, Grade 2b of the Mercuri scale ranges from 30-60% muscle fat involvement, and between stage 3 and 4 of the Fischer scale lies a 50% progression in fat infiltration (Table 1-2). Other semi-quantitative methods based on T1 (Sproule et al., 2011) and T2 weighted (Jurkat-Rott et al., 2009) images have also been suggested, though these suffer from similar problems with lack of sensitivity to change.

Table 1-1 – Mercuri grading scale

Grade	Description
0	Normal appearance
1	Scattered small areas of increased density giving an early 'moth-eaten' appearance
2a	Numerous discrete areas of increased signal comprising less than 30% of the volume of the individual muscle
2b	Numerous areas of increased density comprising 30–60% of the volume of the individual muscle
3	Washed-out appearance due to confluent areas of increased density
4	End-stage appearance - complete replacement of muscle

Qualitative grading scale of muscle fat infiltration on T1-weighted MRI (Mercuri et al., 2002)

Table 1-2 – Fischer grading scale

Stage	Description
0	Normal appearance
1	Traces of decreased signal intensity on CT, or increased signal intensity on T1-weighted MRI
2	Decreased signal intensity on CT, increased signal intensity on MRI with beginning of confluence in <50% of muscle
3	Decreased signal intensity on CT, increased signal intensity on MRI in more than 50% of the muscle examined
4	Entire muscle replaced by lower density on CT or increased signal on MRI

Qualitative grading scale of muscle fat infiltration on T1-weighted MRI (Fischer et al., 2008)

1.4.2 Quantitative MRI

Recent studies have shown that lower limb qMRI is valid (Hiba et al., 2012; Wokke et al., 2013; Wren et al., 2008), reliable (Andersen et al., 2017; Fischmann et al., 2014; Morrow et al., 2014) and responsive (Morrow et al., 2016; Willcocks et al., 2014; Willis et al., 2014) as an outcome measure in certain NMD.

1.4.2.1 Reliability

Reliability of qMRI in NMD is well recognised both within and between centres, the latter important in preparation for multicentre trials. In a study of healthy adult volunteers (Morrow et al., 2014), quantitative 3T MRI (3-point Dixon determined FF, MTR and T2 mapping) was performed at baseline in 47 participants, and repeated at two weeks in 15 of these subjects, revealing excellent scan-rescan and inter-observer intraclass correlation coefficients showing it to be consistent, highly reliable and reproducible. Excellent repeatability of qMRI determined FF has been shown amongst others, in a study in seven healthy children (Ponrartana et al., 2014), a group of patients with LGMD2I (Willis et al., 2013) and Facioscapulohumeral muscular dystrophy (FSHD) (Andersen et al., 2017): Pearson ICC R=0.99, P<0.0001. Between centres, Morrow et al demonstrated excellent test-retest reliability of MRI determined FF within healthy calf muscles by the MRC Centre MRI protocol at a second remote site (Morrow et al., 2018). Good test-retest and inter-observer reliability has also been

demonstrated for other qMRI measures: T1 mapping, T2 mapping and MTR techniques in a homogenous group of healthy volunteers (Fischmann et al., 2014).

1.4.2.2 Validity

The validity of qMRI determined FF has been shown by correlation with clinical (Fischmann et al., 2013) and myometric measures in different NMD. In the first year of this Centre's IBM-CMT longitudinal observational study (Morrow et al., 2016), there was strong correlation between clinical parameters (strength, function and overall severity) and MRI measures providing support for MRI as a means to quantify chronic muscle pathology. In a study of nine boys with DMD (Wren et al., 2008), qMRI measurements of fat infiltration had stronger correlation ($p< 0.05$) with functional grade than did measurements obtained with manual muscle testing ($p=0.07$) or quantitative strength measured with dynamometry.

1.4.2.3 Responsiveness

In this Centre's CMT/IBM natural history study (Morrow et al., 2016), MRI-measured FF provided a highly responsive outcome measure. In the CMT1A group, significant 12-month change was measured in whole muscle calf FF: mean absolute FF change ± s.d. of 1·2 ± 1·5%, $p=0·002$, with a standardised response mean (SRM) of 0·83. This far exceeds the SRM of all previously reported clinical outcome measures in NMD (Verhamme, de Haan et al., 2009; Pareyson et al., 2011). The excellent sensitivity to change of MRI fat quantification has also been demonstrated in other NMD. Other qMRI measures including T2 relaxation time (Kim et al., 2010; Willcocks et al., 2014) and MTR (Sinclair et al., 2012) have also been shown to be highly responsive. All relevant literature relating to qMRI responsiveness is detailed in Chapter 3.

1.4.2.4 Other advantages of qMRI

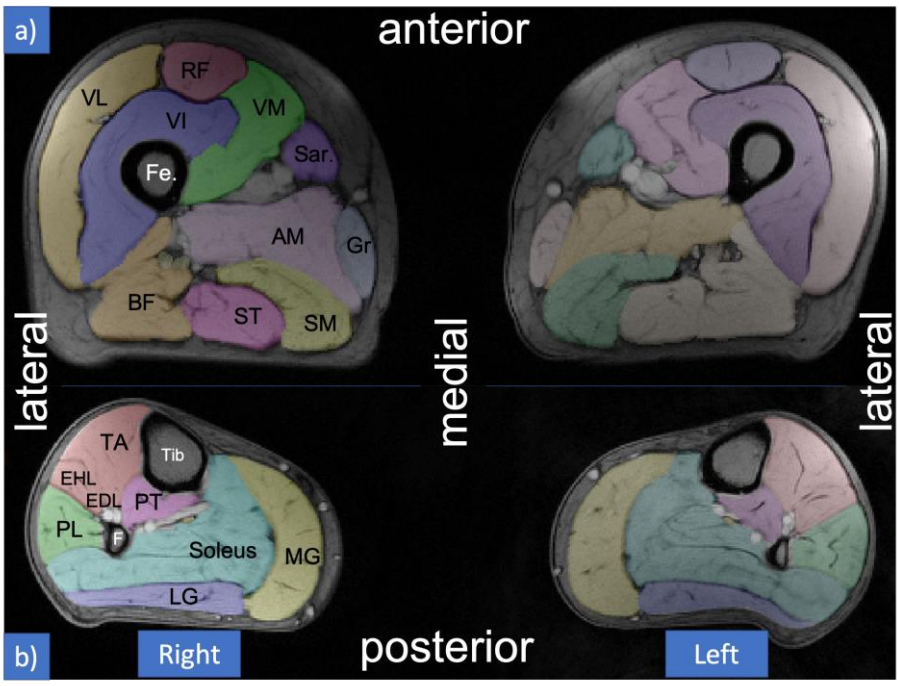
Additionally, MRI has further clear advantages over other outcome measures for NMD trials discussed above. It is safe and does not use ionising radiation. It is non-invasive and can assess the whole breadth of muscles individually and in detail, examining some muscles, either deep or small, which simply cannot be assessed by other methods. It allows multiplanar scanning, which is of particular relevance in NMD, in which contractures can occur impeding patients from lying in anatomical position during imaging. It is non-operator and non-participant dependent, being independent of variation in symptoms and of learning effects. Importantly, data can be stored and analysed at a later date, by one or several examiners, as demonstrated in this thesis (Chapter 4).

1.5 Lower limb MRI

1.5.1 Anatomy

Figure 1-1 demonstrates all muscles which will be analysed in this thesis. Manually defined regions of interest are superimposed, delineating the muscles. These are axial MR images acquired through the mid-thigh (panel a) and mid-calf (panel b). Both legs are shown at both levels.

Figure 1-1 – Lower limb anatomy with superimposed regions of interest



Panel a)=thigh, panel b)=calf. R=right, L=left. Fe=femur, Tib=tibia, F=fibula. See text below for remaining abbreviations

In the thigh, muscles analysed are: VL=vastus lateralis, VI=vastus intermedius, VM=vastus medialis, RF=rectus femoris, Sar=sartorius, Gr=gracilis, AM=adductor magnus, BF=biceps femoris, ST=semitendinosus, SM=semimembranosus.

In the calf, muscles analysed are: TA=tibialis anterior, EHL=-extensor hallucis longus, PT=posterior tibial group, MG=medial head of gastrocnemius, LG=lateral head of gastrocnemius, Soleus and PL=peroneus longus.

1.5.2 MRI sequences

Summarised below are the qualitative and quantitative sequences that are used in this thesis.

1.5.2.1 Qualitative MRI sequences

1.5.2.1.1 Short tau inversion recovery (STIR)

There are problems differentiating between water and fat on T2 weighted images, both appearing bright. To avoid this confounding problem, a STIR sequence is used. This is a fat-suppressed T1 weighted image. Longitudinal magnetisation of fat is first flipped 180 degrees by an inversion radiofrequency pulse and then allowed to relax back to its equilibrium along the magnetic field direction (B0). Water magnetisation, which is also flipped 180° by the same inversion pulse, is subsequently (after the appropriate inversion time) excited for imaging by a 90-degree spin echo pulse as the longitudinal magnetisation of fat crosses the nul point, at which point there is no longitudinal magnetisation to be flipped into the x-y plane from fat. Because fat has a characteristically short T1 relaxation (approximately 220 ms at 1.5 Tesla), if the T1 is set at 150ms, the signal from fat can be suppressed. At this T1, all other tissues (including water), have non-zero magnetization before the 90-degree pulse, and therefore produce large signals. STIR is insensitive to B_0 magnetic field inhomogeneity.

1.5.2.2 Quantitative MRI sequences

In contrast to descriptive semi-quantitative scales of T1 and T2 weighted images, assessment of the MRI signal itself allows analysis of a continuous variable with potential to measure smaller changes over time. In general, three different MRI measures are derived: measures of fat accumulation, water accumulation and muscle area or volume.

1.5.2.2.1 T1 and T2 mapping

MRI contrast is created by capitalising on differences in proton density, T1 recovery time, and T2 relaxation times between adjacent tissues. Of importance to neuromuscular MRI, fluid based tissues have long T1s and T2s, whereas fat based tissues have shorter T1s and T2s (Table 1-3). In T1-weighted images, tissues with short T1s (fat) give the highest signal intensity. In T2-weighted imaged, tissues with long T2s (water) give the highest/brightest signal. With respect to neuromuscular MRI therefore, normal muscle has low signal intensity and appears dark, whereas fat has high signal intensity and appears bright on T1-weighted sequences. Muscle oedema can't be reliably distinguished from normal muscle on T1-weighted sequences, therefore T2-weighted sequences are used in which normal muscle appears dark and both fat and muscle oedema appear bright (Table 1-3).

T1 and T2 mapping allow T1 recovery and T2 relaxation times of tissue to be quantified. An image is generated in which the brightness of each pixel is a direct

reflection of the T1 recovery (T1 mapping) or T2 relaxation (T2 mapping) time. The map can then be analysed by calculating mean values by drawing regions of interest (ROI) within muscles. T2-relaxation time is a sensitive measure of muscle pathology and has been shown to be raised in a variety of muscular dystrophies (Phoenix et al., 1996) including Duchenne muscular dystrophy (Arpan et al., 2013; Huang et al., 1994; H. Kim et al., 2010), juvenile dermatomyositis (Maillard et al., 2004), myotonic dystrophy (Wokke et al., 2013) and in motor neuron disease (Bryan et al., 1998). Given the dependence of T2 time on FF, newer sequences such as IDEAL-CPMG pulse sequence have been developed to separate T2 signal of water from that of fat (Mankodi et al., 2016).

Of particular importance to assessment of NMD, a raised T2 points towards inflammatory, changes when fat is absent, however when fat infiltration is present, because of the much longer T2 time of fat compared to that of muscle oedema, the muscle global T2 becomes largely a measure of tissue lipid content.

In our Centre and others, T1 mapping has been shown in both volunteer and patient analysis to be prone to B1 and B_0 inhomogeneity (Hollingsworth et al., 2013), to be affected by artefact, and have poor reliability and poor sensitive to change compared with other fat quantification techniques (Morrow et al., 2014), and thus will not be further addressed in this thesis.

Table 1-3 – Quantitative MRI characteristics of different tissues

	Muscle	Fat	Fluid based tissues
T2 time (ms)	~50	60-80	700-1200
T1 time (ms)	860-900	240-250	1500-2000

Sequence	Muscle	Denervated	Fat
T1 weighted	intermediate	dark	bright
T2 weighted	intermediate	bright	bright
STIR	intermediate	bright	dark

ms=milliseconds. STIR=short tau inversion recovery

1.5.2.2.2 Magnetisation transfer ratio

Earlier, and potentially reversible changes than the end-stage fat replacement of muscle can be examined by imaging water distribution in muscle. Though this can be done qualitatively with STIR imaging, magnetisation transfer imaging allows quantification of these changes.

In brief, magnetisation transfer creates contrast between tissues which is dependent on magnetisation exchange between two pools of water molecules present in any tissue: free, mobile protons which exist in free water and in some lipid tissues, and bound, restricted protons which are bound in proteins and macromolecules. The latter have a very short T2 time and are therefore not normally visualised on standard MRI. If an off resonance radiofrequency pulse is applied to selectively saturate the bound pool of protons, the saturated protons may enter the free pool of protons, or transfer their magnetisation to free water protons, resulting in a decrease in the MR visible signal in areas of macromolecules affected by magnetisation transfer. The magnetisation transfer ratio is then calculated – the magnitude of which is determined by the proportion of water molecules in the bound pool.

Sinclair et al. (Sinclair et al., 2012) assessed MRI magnetisation transfer ratio (MTR) in ten patients with CMT1A, nine with chronic inflammatory demyelinating neuropathy (CIDP) and ten healthy controls, finding that the MTR in the two patient groups was significantly lower than controls: median MTR 50.5 (±1.6) p.u in controls, 41.5 (±10.6) p.u in CMT1A and 39.3 (±8.7) p.u. in CIDP. MTR was significantly reduced in certain muscles which appeared normal by semi-quantitative assessment, indicating that MTR may be more sensitive to early muscle changes than conventional MRI sequences. The MTR of muscle has been shown to be decreased in a variety of other NMD including limb girdle muscular dystrophy (McDaniel et al., 1999).

1.5.2.2.3 Muscle fat quantification

Fat can be quantified by various methods including T1 relaxometry (see above), proton MR spectroscopy and chemical-shift sensitive methods.

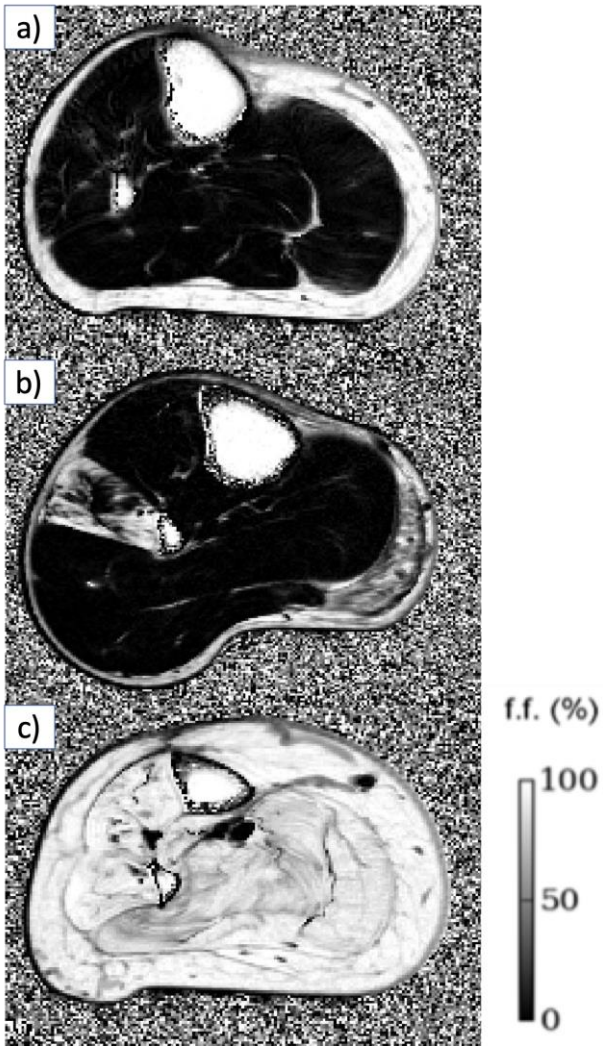
The Dixon method (Dixon, 1984; Glover and Schneider, 1991; Karampinos et al., 2011) is a simple spectroscopic imaging technique for water and fat separation which relies on the difference in chemical shift between water and fat. It requires a minimum of two separate gradient echo images with a modified spin echo pulse sequence – though a third is also acquired to correct for field inhomogeneities (Glover and Schneider, 1991). The first is a conventional spin echo image with in-phase water and fat signals. The second is acquired so that the water and fat signals are out of phase. From these two separate acquisitions, a water only, and a fat only image can be generated by a mathematical algorithm, allowing direct fat and water quantification in each voxel.

This technique was initially unsuccessful as it was marred by B_o inhomogeneity, however since the original description, several modifications have been made to

further improve the imaging (Ma, 2008). Example Dixon fat maps are seen in Figure 1-2.

Most studies in NMD have used either two or three gradient echoes (2-point or 3-point Dixon), though investigation is being done into using more echoes for improved signal to noise ratio (Grimm et al., 2019). A number of studies have quantified fat replacement of muscle by MRI in both cross-sectional (Sproule et al., 2011; Willis et al., 2013; Dahlqvist et al., 2014; Wokke et al., 2014) and longitudinal studies, including in OPMD (Fischmann et al., 2012), LGMD (Willis et al., 2013), DMD (Arpan et al., 2014; Wren et al., 2008) and FSHD (Janssen et al., 2014). Literature is reviewed in detail in subsequent chapters.

Figure 1-2 – Dixon fat map of the mid right calf in three patients with CMT1A



Axial Dixon fat map at mid-right calf level for three patients with CMT1A: a) mild, b) moderate and c) advanced, demonstrating the appearance of progressive muscle fat infiltration

1.6 Aims of this thesis

Neuromuscular disease research is in the midst of discovery of potentially life-changing and lifesaving treatments, each one of which must be rigorously assessed in randomised controlled trials. A highly responsive outcome measure is needed.

Quantitative MRI is reliable, valid and responsive to change over time, but ongoing improvements, particularly in qMRI responsiveness are critical in order to ensure clinical trials in these rare and clinically heterogenous diseases can be adequately powered to detect drug effect over short duration.

The overarching aim of this PhD thesis was to refine quantitative MRI as an outcome measure for NMD, in preparation for future clinical trials. This was achieved by longitudinal assessment of qMRI measures in two separate inherited neuropathies: CMT1A and HSN1. Chapter 3 reports on the extension of our Centre's CMT1A natural history study in which qMRI, clinical and functional measures were assessed over prolonged follow up. Results revealed large internal responsiveness and longitudinal validity of combined bilateral calf qMRI determined FF and of functional measures over up to five years. In Chapters 4 and 5, logical stepwise application of differing methods of qMRI data analysis were aimed at improving qMRI responsiveness. In Chapter 4, CMT1A muscle calf fat gradients were examined at baseline, with different methods of longitudinal analysis informing significant improvements in qMRI responsiveness over 12 months. In Chapter 5, qMRI was applied longitudinally at three anatomically distinct lower limb levels in a cohort of patients with HSN1, with various adjustments in data analysis, including a novel baseline severity specific approach identifying further methods by which to improve responsiveness.

This thesis concludes with comprehensive evidence-based suggestions for use of qMRI as an outcome for future clinical trials in NMD.

2 Methods

This thesis describes three separate studies each using similar methods, all of which are described in this chapter. Any methodological details specific to one or another study are clearly indicated as such.

2.1 Background

2.1.1 Refinement of MRI as an outcome measure

The challenge facing upcoming neuromuscular clinical trials is to measure small therapeutic effects over relatively short durations. This requires highly responsive outcome measures to achieve adequate trial statistical power: underpowered clinical trials increase the chance of both false positive and negative results. There is also the risk of overestimation of the effect size of a true treatment effect.

In diseases with higher prevalence, trials designed to detect small therapeutic effects mainly achieve adequate power by using large sample sizes – with not insignificant associated cost. As an example, the recent Odyssey Outcomes trial recruited almost 19,000 participants (Szarek et al., 2019) – which is not unusual in cardiovascular medicine multicentre, randomised placebo-controlled trials. Neuromuscular diseases (NMD) however are rare, without the possibility of being able to supercharge a trial by recruiting such an extraordinary number of participants. The most common of the NMD, Charcot-Marie-Tooth disease type 1A (CMT1A) has prevalence of ~1 in 2500, and many diseases are rarer, such as Hereditary Sensory Neuropathy (HSN), of which there are only a handful of patients in the UK, and even fewer in other individual countries worldwide. This paucity of numbers, and geographic dispersion means that simply increasing participant numbers to increase clinical trial power is not a viable option for many NMD, particularly when physical impairments make travel difficult, if not impossible. International collaboration is often necessary, but even then available study numbers are small and co-ordination between centres not without its challenges. Related to this, patients with rare diseases are at risk of 'research-fatigue' due to a large number of successive trials using the same cohort of patients, and in this setting, participant enrolment becomes more difficult and attrition rates considerable with consequent reduction in statistical power.

How then do we increase power without increasing sample size? The alpha level cannot be increased beyond 0.05 without unreasonable risk of type I error, and effect

size is difficult to change. As discussed in Chapter 1, both drug efficacy and outcome measure responsiveness have a potent effect on trial power according to Lehr's formula. With respect to the former, given the chronic nature of NMD, new treatments would be expected to slow disease progression, or at best halt it in the first instance, rather than reverse it. Thus, a reliable, valid and highly responsive outcome measure is needed. Fat quantification by MRI, in particular using the 3-point Dixon technique has been shown on 12-month follow-up data to be highly responsive (Morrow et al., 2016), with standardised response mean far exceeding any previously reported outcome measure used in trials of patients with neuromuscular diseases. Further refinement is both possible and needed however. An increase in outcome measure responsiveness is achieved by increasing the ratio of mean change to standard deviation of change, the latter encompassing the combination of true disease-driven change variation, and variation due to random measurement errors. It is this refinement that lies at the heart of this PhD thesis.

In Chapter 3, I present the extension and culmination of our Centre's CMT1A natural history study, the major aim of which was to assess outcome measure responsiveness in a heterogeneous cohort of CMT1A patients over extended follow up, through longitudinal assessment of various MRI and clinical biomarkers. The thesis then turns to refinement of quantitative MRI (qMRI) responsiveness in two distinct inherited neuropathies, identifying a host of evidence-based methods by which this is imminently achievable. Based on the findings from these three chapters, an outline of the role of qMRI in future clinical trials is then proposed.

2.1.2 Participant groups

The two neuromuscular diseases studied in this thesis are CMT1A, which is the most common inherited neuropathy, and in which an explosion in potential treatments is at hand, and Hereditary Sensory Neuropathy type 1 (HSN1), one of the rarer inherited peripheral neuropathies, but one for which a potential drug treatment is ready for clinical trials (Fridman et al., 2019). Thus perhaps these two neuromuscular diseases are in most urgent need of a highly responsive outcome measure at this critical time.

2.1.2.1 Charcot-Marie-Tooth disease

Charcot-Marie-Tooth disease, also known as hereditary motor and sensory neuropathy, comprises a large group of inherited motor and sensory neuropathies, with as their core feature a neuropathy primarily affecting either the peripheral nerve axon (CMT type 2) or its myelin sheath (CMT type 1). The prevalence of CMT is reported to be between 10 and 80/100,000 worldwide, (Kurihara et al., 2002;

Braathen, 2012) and between 10-20/100,000 in Europe (Lefter, Hardiman and Ryan, 2017). With the help of next generation sequencing techniques, many novel genes have been shown to be associated with CMT in which the neuropathy is only a part of the clinical syndrome. Although usually slowly progressive over decades, CMT can be severe and rapidly progressive in certain genetic subtypes.

CMT1A is the most common of the inherited neuropathies, accounting for ~50% of all cases of CMT. It is a length-dependent sensorimotor neuropathy which is caused by duplication of the region of chromosome 17p11.2 containing the 22 kilodalton peripheral myelin protein (PMP-22) gene (Lupski et al., 1991; Raeymaekers et al., 1991; Timmerman et al., 1992). It is a demyelinating neuropathy in which secondary axonal degeneration leads to weakness and wasting in addition to sensory loss (Krajewski et al., 2000) affecting the patient in a length-dependent manner (Reilly, Murphy and Laura, 2011; Shy et al., 2005). Although much is known about its pathogenesis, no drug treatments for CMT1A are currently available, although rapid progress is being made (D'Ydewalle et al., 2011; Garofalo et al., 2011; Patzkó et al., 2012; Gutmann and Shy, 2015; Saporta et al., 2015). Recent trials have employed outcome measures with no to minimal responsiveness, and there is a need for a highly responsive outcome measure to go forward with meaningful clinical trials.

2.1.2.2 Hereditary Sensory Neuropathy

The hereditary sensory and autonomic neuropathies (HSN) are a genetically diverse group of inherited peripheral nerve diseases. The clinical spectrum is wide, but unified by the common finding of a progressive, sensory neuropathy which, when marked, may lead to severe complications including unintentional self-injury, skin ulcers and osteomyelitis sometimes necessitating amputation (Houlden, Blake, and Reilly, 2004). There is variable autonomic involvement and sometimes quite disabling motor deficits develop in certain genetic subtypes later in the disease course (Skre, 1974; Ouvrier, 2010; D'Ydewalle et al., 2011; Patzkó et al., 2012). The classification of HSN has been strengthened by the discovery of twelve genes since 2001, although these still account for less than 20% of cases.

The most common form of HSN is HSN type 1A, which is an autosomal-dominant predominantly sensory neuropathy in which motor involvement is variable, and neuropathic pain common (Houlden et al., 2006). It is caused by a mutation in the *SPTLC1* gene, which encodes a subunit of the serine palimtoyltransferase (SPT) enzyme which catalyses the rate limiting step in sphingolipid synthesis. Mutations in *SPTLC1* alter the substrate specificity of SPT leading to production and accumulation of neurotoxic deoxysphingolipids (Eichler et al., 2009; Penno et al., 2010). L-serine

supplementation has been identified as a possible therapy in this form of HSN1 (Fridman et al., 2019; Garofalo et al., 2011). A recent randomised controlled trial in HSN1 revealed that L-serine may slow disease progression (Fridman et al., 2019), though further trials with more responsive outcome measures are needed. The fact that this is a predominantly sensory neuropathy poses unique challenges for selection and development of an outcome measure.

2.1.2.3 Healthy controls

The use of matched healthy controls is crucial. At baseline, controls provide a measure of the distribution of normative values, which may differ between implementations on different MRI machines. For longitudinal assessment, they provide a control for healthy aging effects as a potential confound, or for measurement changes due to MRI system errors which may be wrongly attributed to the disease process if a control group is not also examined contiguously with the patient participants.

2.1.3 Study Sample sizes

2.1.3.1 CMT natural history study

20 patients provide 80% power to detect significance of a mean change of 0.66 SD units, assuming 10% dropout using a paired t-test and a 5% significance level (Chow, Shao and Wang, 2008). The patient data reported here is that of five-year follow-up data, with correspondingly greater power, assuming greater disease-related effect sizes over the longer time interval: longitudinal data on 10 patients still provides 80% power to detect significance of a mean change of 0.89 SD Units using a paired t-test and a 5% significance level. We chose to include a healthy control group of equal size.

2.1.3.2 CMT gradient study

All patients from the CMT1A natural history study who had multi time point data of adequate quality were included.

2.1.3.3 HSN natural history study

All HSN patients attending the inherited neuropathy outpatient clinic at the National Hospital for Neurology and Neurosurgery, London who consented, were enrolled, in order to maximise study power.

2.1.4 Ethical considerations

The CMT1A natural history study and HSN study were approved by the local research ethics committee. All participants were provided with written information sheets and

signed informed consent after at least 24 hours consideration. The CMT1A MRI Gradient Study was an analysis of previously acquired data, covered by the same ethical approval and consent process.

2.1.5 Study design

2.1.5.1 CMT1A natural history study

This study was a prospective longitudinal observational study with assessments set nominally 12 months apart. Chapter 3 reports on the continuation of the CMT1A/IBM longitudinal natural history study, data from the first year of which has been previously published (Morrow et al., 2016). Following the first publication analysing 12 month data for both disease groups, analysis of the two disease groups was separated. This thesis is focused upon the CMT1A cohort. At baseline, 20 patients with CMT1A and 20 matched controls were enrolled and had baseline assessment comprising collection of demographic and clinical information on a standardised form, MRI, myometry and clinical evaluation. This cohort was then examined at three further visits (Figure 3-7) at each of which all previous assessments were repeated. At the second visit (mean interval ± s.d. of 12.8 ± 1.1 months), there remained 17 CMT1A patients and 20 controls. At the third visit (28.7 ± 2.5 months) there remained 14 CMT1A patients and 11 controls, and at the fourth visit (47.9 ± 6.3 months) there remained 10 CMT1A patients and 8 controls. Participant dropout for CMT1A patients was attributable to death, surgery, illness, inability to travel and personal reasons to withdraw from the study. For controls, withdrawal of their patient 'buddy' with whom they co-attended based upon a personal relationship, meant that they too withdrew from the study.

2.1.5.2 CMT1A gradient study

This chapter of the thesis describes a retrospective analysis of existing 3-point Dixon imaging data from successive anatomical levels in the lower limbs.

Data from thirteen patients with CMT1A and eight matched controls were included. These participants underwent baseline assessment comprising collection of demographic and clinical information and 3-point Dixon MRI and a repeat assessment at 12 months.

2.1.5.3 Hereditary Sensory Neuropathy type 1

This study was a prospective longitudinal observational study with assessment performed at an interval of 12 months. At both visits, demographic and clinical information was collected on a standardised form. Participants underwent MRI, myometry, quantitative sensory testing, neurophysiology, skin biopsy and clinical assessment (including CMTNS version 2). Data from 34 patients with HSN1 and 10

matched controls were analysed at baseline. Data from 25 patients with HSN1 and 10 controls were reassessed at mean=364 ± 7 days. Three patients were lost to follow up, and six MRI scans were could not be analysed due to technical failures or data transfer errors. This thesis deals with the MRI aspects of this study in detail, the author (MRBE) having performed all imaging assessments and analysis. Our group's already published paper (Kugathasan et al., 2019) details results of other outcome measures, and of a more basic MRI assessment.

2.1.6 Participants

2.1.6.1 CMT1A natural history study

Patients with genetically confirmed CMT1A (chromosome 17p11.2 duplication) attending the inherited neuropathy outpatient clinic at the National Hospital for Neurology and Neurosurgery, and who were aged ≥ 17 years were invited to participate. Participants with concomitant neuromuscular diseases, significant organ comorbidities, who were pregnant, who had had surgery to the feet within the previous 12 months, or who did not satisfy MRI safety criteria were excluded from the study. Healthy controls were matched for age, sex, weight and height, and included research staff, as well as relatives of CMT1A patients. All assessments were done at the National Hospital for Neurology and Neurosurgery, London, UK.

2.1.6.2 CMT1A gradient study

Imaging from all CMT1A patients and controls who had been enrolled in the CMT1A natural history described above, and who had baseline and follow-up imaging was assessed. Data from those patients with adequate quality proximal-distal imaging at both baseline and 12 month follow-up were included in this gradient study.

In total, data from thirteen patients with CMT1A were included, having mean age of 46.2 years (range 19 – 67 years), median age at disease onset of six years (range 1 -27) and median CMT examination score of 7 (range 0-18). Data from eight age and gender matched healthy controls were included (mean age 53.4 years, range 25 – 80y). All assessments were done at the National Hospital for Neurology and Neurosurgery, London, UK.

2.1.6.3 HSN study

Patients aged ≥18 years with genetically confirmed HSN1 due to *SPTLC1* or *SPTLC2* mutations were identified from the Neurogenetics Laboratory, National Hospital for Neurology and Neurosurgery and Bristol Genetics Laboratory. Exclusion criteria were concomitant neuromuscular diseases, significant comorbidities and safety-related MRI contraindications. Healthy controls were matched for age, sex, weight and height, and included research staff, as well as relatives of CMT1A patients. All

assessments were done at the National Hospital for Neurology and Neurosurgery, London, UK.

2.2 Clinical assessments

2.2.1 CMT1A natural history study

Participants underwent a battery of assessments at baseline, and the subsequent three follow-up visits.

2.2.1.1 CMT1A specific rating scale

The CMT examination score (CMTES) was used in this study as the measure of clinical severity (Murphy et al., 2011). As is detailed on page 33, the CMTNS (which includes neurophysiology) has been used as the primary outcome measure in previous trials in patients with CMT, and has proven minimally responsive. Indeed CMTES has fared better in terms of responsiveness in one study (Micallef et al., 2009). In light of these findings, neurophysiology was therefore not performed in this study.

2.2.1.2 Clinical assessment

All participants had a medical history taken, underwent a standard neurological examination, and completed an SF-36 questionnaire at each visit. As there are currently no other validated outcome measures in CMT1A, no further clinical assessments were done.

2.2.1.3 Medical Research Council scale for strength assessment

MRC grading assessment of muscle strength was performed at each visit according to a modification of the MRC scale which was first published in 1943 (Medical Research Council, 1943). Baseline and first follow-up assessments were performed by Dr Jasper Morrow, and the second and third follow up by the author (MRBE).

Movements assessed were:

- Neck: flexion and extension
- Upper limbs: shoulder abduction, elbow extension, elbow flexion, wrist extension, wrist flexion, finger extension, index finger abduction, little finger abduction, thumb abduction, long finger flexors, short finger flexors.
- Lower limbs: hip flexion, hip extension, hip abduction, hip adduction, knee flexion, knee extension, ankle dorsiflexion, ankle plantarflexion, ankle eversion, ankle inversion, great toe extension.

Muscle strength was graded using a modified MRC scale (O'Brien, 2000) (Table 2-1).

Table 2-1 – MRC grading scale for muscle strength

Grade	Description
5	Normal contraction against full examiner resistance
5-	Weakness barely detectable
4+	Movement against gravity and near maximal resistance
4	Movement against gravity and moderate resistance
4-	Movement against gravity and minimal resistance
3	Full range of movement against gravity (but no more)
2	Movement possible but not against gravity
1	Only a flicker of movement seen or felt
0	No movement

Upper limb/neck and lower muscle scores were summed to obtain a total upper limb/neck and total lower limb score for each subject. For this purpose, 5- was scored as 4.75, 4+ as 4.25 and 4- as 3.75. The maximum score obtainable was 120 for upper limb/neck and 110 for lower limb.

As is seen from descriptions of each grade, there is a subjective element, particularly in the scores surrounding '4', which in large part accounts for the low sensitivity of this scale as an outcome measure for longitudinal assessment.

2.2.1.4 Myometry

Patients and controls underwent detailed lower limb myometry on a HUMAC NORM dynamometer (CSMi, Massachusetts, USA) which most often occurred on the same day as the MRI assessment. The myometry assessment was performed following the MRI when possible in CMT1A patients to try and avoid any possible effects of exercise on the imaging. Baseline and first follow-up visits were performed by Dr Jasper Morrow (see acknowledgments) and all second and third follow-up appointments by the author. There was detailed training from one assessor to the other over a three month period to minimise any inter-operator bias. Myometry was performed as detailed in Table 2-2. All movements at knee and ankle were assessed bilaterally using both isometric and isokinetic protocols and the maximum torque in Newton metres recorded. Isometric assessments consisted of four attempts of three seconds separated by ten seconds rest. Given the more complex movement, the isokinetic assessments started with a practice run and ten second rest, then three successive movements through full range were performed. The best attempts from both isometric and isokinetic assessments were selected for analysis. The machine setup was recording at baseline visit, allowing identical machine positioning on follow-up testing. Although movements at the knee were recorded by the author, this data

was not analysed as part of this thesis, given results of the first assessment showing minimal involvement of thigh muscles in this cohort of patients (Morrow et al., 2016).

Table 2-2 – HUMAC NORM myometry as performed for the CMT1A natural history study

Joint	Type	Movement	Angle/Speed
Knee	Isometric	Extension	45°
		Extension	90°
		Flexion	45°
		Flexion	90°
	Isokinetic	Extension and Flexion	60°/s
		Extension and Flexion	120°/s
Ankle	Isometric	Plantarflexion	10°
		Dorsiflexion	10°
	Isokinetic	Plantarflexion/Dorsiflexion	60°/s
	Isometric	Eversion	10°
		Inversion	10°
	Isokinetic	Inversion and Eversion	60°/s

2.2.2 CMT1A gradient study

The analysis relevant to the CMT1A gradient study is of the 2D 3-point Dixon MRI data collected as part of the CMT1A natural history study. Age, Weight, Height, BMI and CMTES data were also included in the assessment, the latter as a marker of overall disease severity.

2.2.3 HSN natural history study

As noted above, participants underwent a battery of assessments at baseline and follow-up. This thesis is dedicated to the MRI branch of this study, of which details are set out below.

2.3 MRI system

In all studies, participants were examined lying feet first and supine. MR imaging was at 3 Tesla (TIM Trio, Siemens, Erlangen, Germany). A multi-channel peripheral angiography coil and 'spine matrix' coil elements were used to image both lower limbs within the same field of view.

2.3.1 CMT1A studies

Axial-slice matrices and fields of view were 256x120 and 400x188mm for calf-level images, except for fat fraction (FF) acquisitions where matrices were 512x240 pixels for calf imaging.

2.3.2 HSN study

As part of a wider protocol, the 3-point Dixon technique was used for fat quantification of the proximal calf muscles with 88×5 mm contiguous slices, 448x210 matrix.

2.4 MRI block positioning

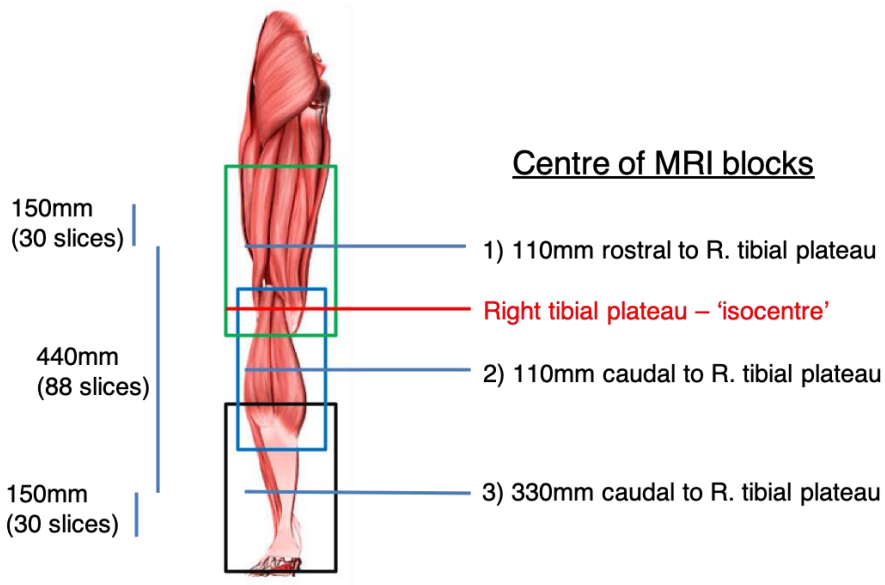
2.4.1 CMT1A studies

For baseline and visit two, calf imaging was centred one quarter of the total distance from the tibial tuberosity to the lateral malleolus (as measured by surface anatomy). This distance at baseline was recorded and used for block positioning at visit two to ensure consistency. For visits three and four however, an improved scout-based imaging method with internal bony landmarks was used, as described in our group's paper (Fischmann et al., 2014).

2.4.2 HSN study

3D 3-point Dixon imaging was performed for measurement of FF. Three overlapping imaging blocks were positioned (prescribed relative to scout images): one set 110mm above, the second 110mm below and the third 330mm below the right tibial plateau. Each imaging block consisted of 88 slices each with 5mm thickness and no gap, thus with a section of each block overlapping, there are ~148 slices available for analysis (Figure 2-1).

Figure 2-1 – 3D 3-point Dixon MRI block placement in HSN study



Axial slices are 5mm thick with no gap

2.5 Choice of anatomical location for imaging

2.5.1 CMT1A natural history study

For the CMT1A natural history study, the lower limbs were chosen for imaging investigation rather than the upper limbs. CMT1A is a length-dependent neuropathy, affecting distal legs in the first instance (Gutmann and Shy, 2015) before thighs and upper limbs. There is measurable clinically meaningful deficit at the ankles, and previous imaging studies in lower limbs, with which to compare results. Finally, the MRI scanning time is potentially shorter in the lower limbs as it is easier to image both sides together for these than the arms. Thigh data were not included in the ongoing longitudinal analysis presented in this thesis, given that previous analysis of this cohort's 12 month baseline and follow-up data revealed minimal thigh involvement - only 3/20 patients having full region of interest (ROI) thigh FF >5% by qMRI (Morrow et al., 2016).

2.5.2 CMT1A gradient study

For the CMT1A gradient study, it was felt that data from a single calf would be adequate for the analysis, in light of the results from the CMT1A natural history study showing little difference between limbs, and little change in responsiveness when comparing both, combined to single-leg measurements.

2.5.3 HSN study

Imaging was also obtained from the lower limbs in this study. Although a predominantly sensory neuropathy, it is known from natural history studies that motor nerve involvement is not uncommon as the disease progresses (Houlden, Blake and Reilly, 2004). As per CMT1A, HSN is a length-dependent neuropathy, and the pathology is expected to start distally, progressing proximally over time.

2.6 MRI sequences

Scanning time for all studies was less than 60 minutes and included some or all of the following sequences: T1-weighted and short tau inversion recovery (STIR) qualitative imaging, FF measurement using the 3-point-Dixon technique (Glover and Schneider, 1991), T1 relaxometry by the DESPOT method with B1 correction (Deoni, Rutt and Peters, 2003), pseudo-T2 relaxometry from dual-contrast turbo-spin-echo images, magnetisation transfer (MT) ratio derived from two 3D-FLASH images with and without an MT pre-pulse with B1 correction (Sinclair et al., 2012).

2.6.1 Dixon fat fraction measurements

2.6.1.1 CMT1A studies

For Dixon FF imaging (Glover and Schneider, 1991), three 2D gradient-echo acquisitions were performed (parameters: TE1/TE2/TE3=3.45/4.60/5.75ms, TR=100ms, flip angle=10^{o}, bandwidth 420Hz/pixel, NEX=4, 10 x 10mm slices with 10mm gap, 512x240 matrix calf, iPat=2). Phase unwrapping was performed using PRELUDE (FSL, FMRIB, Oxford) (Smith et al., 2004) and after fat (F) and water (W) image decomposition, FF calculated as FF = 100% x F/(F+W). The TE=3.45ms image was used for the ROI placement and as a reference for inter-method image interpolation and registration using FLIRT (FSL, FMRIB, Oxford), so that that the same ROIs could be applied to extract data from the T2 and MTR maps.

2.6.1.2 HSN study

As part of a wider protocol, three 3D gradient-echo acquisitions were performed at three anatomical levels. Parameters were TE1/TE2/TE3=3.45/4.60/5.75 ms, TR=23 ms, flip angle= 5^{o}, bandwidth 446 Hz/pixel, NEX=1, 88 contiguous 5mm slices, 448x210 matrix, iPat=2. Phase unwrapping was performed using PRELUDE (FSL, FMRIB, Oxford) (Smith et al., 2004) and after fat (F) and water (W) image decomposition, FF calculated as FF = 100% x F/(F+W). The TE=3.45ms image was used for the ROI placement.

2.6.2 Other sequences – CMT natural history study

2.6.2.1 T1 relaxometry

DESPOT-1 (Deoni, Rutt and Peters, 2003), T1-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^{o}, TR/TE=23/3ms, and BW=440Hz/pixel acquired in a single, non-selective slab with 80 x 5mm longitudinal phase-encoded partitions. Flip-angles were corrected using B_1 maps obtained as below and T_1 calculated according to Deoni (Deoni, Rutt and Peters, 2003). The T1 mapping data were not included in the following analyses since earlier analyses suggested they were not useful, largely for technical reasons. The 3D FLASH images were however also used to help anatomically localise slices from the other sequences selected for analysis.

2.6.2.2 T2 relaxometry

Non-fat-suppressed dual-contrast turbo-spin-echo (TSE) images (6500/13/52ms or 6500/16/56ms; 10x10mm slices with 10mm gap, iPat=2, BW=444Hx/pixel, refocusing flip angle 180^{o}, NEX=1, 6/8 *k*-space sampling, 256x128 matrix (thigh), 256x120 matrix

(calf)) were acquired. Pseudo-T2 was calculated from the respective pixel intensities ITE1 and ITE2 from the TE1 and TE2 images.

During the first year of the study (following 54 baseline and 6 follow-up scans), a routine MRI scanner software upgrade resulted in slightly different echo times (above) due to altered sequence timing constraints. An empirically derived correction equation (corrected T2 = 1.0933 x post-upgrade T2 – 0.0245) was applied to the post upgrade calf T2 values to correct for this (Morrow et al., 2016).

2.6.2.3 Magnetisation transfer ratio

MTRs were calculated from two 3D-FLASH images with (M1) and without (M0), an MT pre-pulse (500° amplitude, 1200Hz offset, 10ms duration) (TR/TE=65/3ms or 68/3ms, α=10º, BW=440Hz/pixel, NEX=1, 6/8 *k*-space sampling, iPat=2, 40x5mm longitudinal phase encoding partitions, 256x128 matrix (thigh), 256x120 matrix (calf)) according to MTR = (M0-M1)/M0 x 100 percentage units (p.u.). Percentage units were used by convention to avoid ambiguity with fractional change expressed as a percentage. MTR maps were RF-inhomogeneity corrected using the B1 maps according using a mean-over-all-subjects B1 inhomogeneity correction factor of k = 0.0085 (Sinclair et al., 2012).

2.6.2.4 T1 weighted imaging

Standard T1-weighted images were acquired with a turbo-spin-echo readout prior to commencing the quantitative protocol (TR/TE=671/16ms, 10 slices, 10mm thickness, 10mm slice gap, 256x192 matrix, iPat acceleration of 2, 444 Hz/pixel bandwidth (BW), TSE factor=3, refocusing flip angle (fa) 130º, NEX=2, acquisition time (TA) = 45s).

2.6.2.5 Short tau inversion recovery imaging

STIR (TR/TE/Inversion Time = 5500/56/220ms, NEX=1, flip angle 180°, parallel imaging factor (iPat)=2) imaging was performed (10x10mm slices, 10mm gap, 256x128 matrix (calf)).

2.7 Qualitative image analysis

2.7.1 CMT1A natural history study

Analysis of baseline fatty infiltration on the T1 weighted MR images was performed using the 6-point Mercuri scale (Table 1-1) (Mercuri et al., 2002) and for STIR hyperintensity using a three-point scale (0=none, 1=mild, 2=marked) by a radiologist with 4 years post-speciality experience (Dr Fischmann – see acknowledgements) who was blinded to subject group. Repeat images were graded without reference to

baseline images. The same muscles were assessed as in the quantitative analysis below, with assessment over the entire imaged muscle volume, rather than just on a single slice.

2.8 Quantitative image analysis

2.8.1 Slice selection

2.8.1.1 CMT1A natural history study

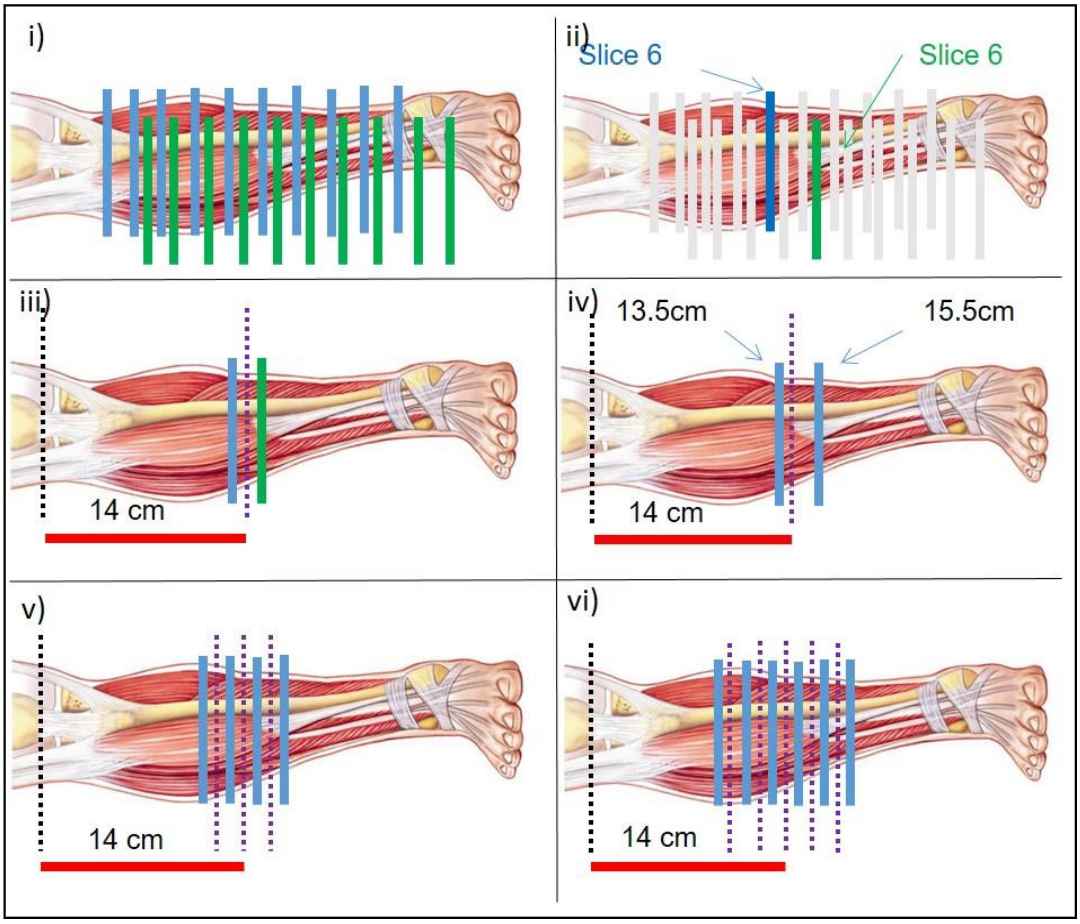
For the baseline scan, the sixth most superior slice in the calf was used for analysis, unless all muscles were not visible on that slice, in which case an adjacent slice was selected. At each follow-up, the slice chosen for analysis was that which was most similar in appearance to that used from the first visit, based on measured distance from bony landmarks (tibial plateau or tip of the fibular head) identified on the 3D-FLASH images. This was done, and visually checked by the author (MRBE) for each follow-up image set.

2.8.1.2 CMT1A gradient study

Using the 3D-FLASH images, at both baseline and follow-up, the distance of each of the ten axial slices distal to the right tibial plateau (on a line drawn between the right tibial plateau and right medial malleolus) was documented, in order to enable precise baseline and longitudinal comparison. A point 14cm distal to the right tibial plateau was chosen as the central point of the data series for all scans. This was a point at which all participants had at least one axial slice both proximally and distally, thus maximizing the number of slices available for analysis.

Five methods to select the axial slice(s) for longitudinal analysis were compared, choosing: (a) slice 6 – one of the numerically central slices; (b) the slice closest to the fixed distance 14cm distal to the right tibial plateau; or calculating the weighted mean FF (FF weighted by cross-sectional area) of (c) two, (d) four and (e) six slices centred at the point 14cm distal to the right tibial plateau. If one of the slices was itself at 14cm distal to the right tibial plateau, then c) used that slice alone, whereas d) and e) used the average of three and five slices about that point respectively. Weighted mean FF assumed linear FF gradients between adjacent slices. The slice selection process is detailed in Figure 2-2.

Figure 2-2 – Methods of axial slice selection for the CMT1A gradient study



Imaging 'block' (i) consisted of 10 axial slices each separated by 2cm (blue=baseline, green=follow up, with hypothetical positioning mis-match). Five methods were used to select the slice(s) for longitudinal analysis, choosing the: (ii) numerically central slice, iii) slice closest to a fixed distance 14cm distal to the right tibial plateau, or calculating the weighted mean of (iv) two, (v) four and (vi) six slices centred 14cm distal to the right tibial plateau

2.8.1.3 HSN natural history study

For the baseline scan, the imaging slices chosen for analysis were at the centre of the imaging blocks (Figure 2-1): for the distal calf a distance of 330mm distal to the tibial plateau, for the proximal calf a distance of 110 mm distal to the tibial plateau. For the distal thigh, a distance of 110mm above the tibial plateau – such that slices were separated by 220mm. Slice position for each lower limb was measured separately to ensure no detrimental effect on slice selection of possible pelvic tilt. At follow-up, the same process was applied and slices visually checked.

2.8.2 Defining regions of interest

The author (MRBE) manually defined all whole and small ROI blinded to clinical details and visit number for the three studies included in this thesis. Blinded to patient

visit, ROIs for each participant were drawn with reference to each other, so as to ensure equivalent ROI placement.

2.8.2.1 CMT1A natural history study

Whole and small regions of interest were defined bilaterally on the selected slice as shown in Figure 2-3 on an unprocessed Dixon acquisition (TE=3.45ms) using ITK-SNAP36 software. Whole ROI included the entire cross-sectional area of the muscle up to but not including the surrounding fascia, and excluding prominent neurovascular structures. Small regions of interest were defined on an area within the muscle belly avoiding artefact and neurovascular structures. Left and right calf ROIs were thus defined for six muscles/muscle groups: tibialis anterior, peroneus longus, lateral gastrocnemius, medial gastrocnemius, soleus and the posterior tibial group of muscles.

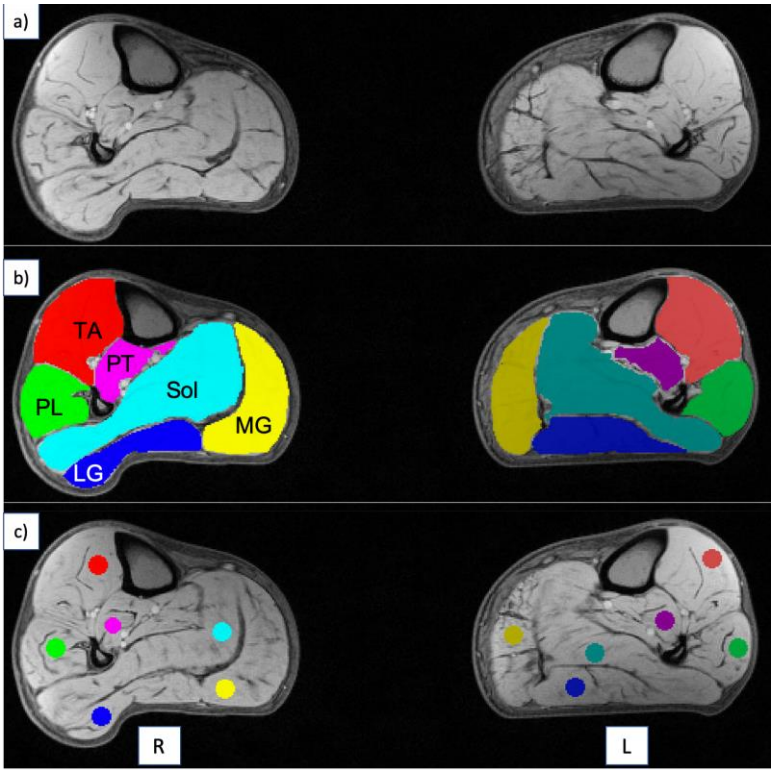
2.8.2.2 CMT1A gradient study

Right calf whole muscle ROI were defined on an unprocessed Dixon acquisition (TE=3.45ms) using ITK-SNAP36 on each of the ten axial slices (Figure 2-4). The entire cross-sectional area of the muscle was included up to but not including the surrounding fascia, and excluding prominent neurovascular structures. As is demonstrated in Figure 2-4, depending on exact block positioning and calf anatomy, some muscles were not included on all ten slices. For example, the most distal slices often did not include the medial or lateral head of the gastrocnemius muscle. ROIs were defined for seven muscles/muscle groups: medial gastrocnemius, lateral gastrocnemius, peroneus group (peroneus longus and brevis), soleus, tibialis anterior, extensor hallucis longus and the posterior tibial group (tibialis posterior, flexor digitorum longus and flexor hallucis longus).

2.8.2.3 HSN study

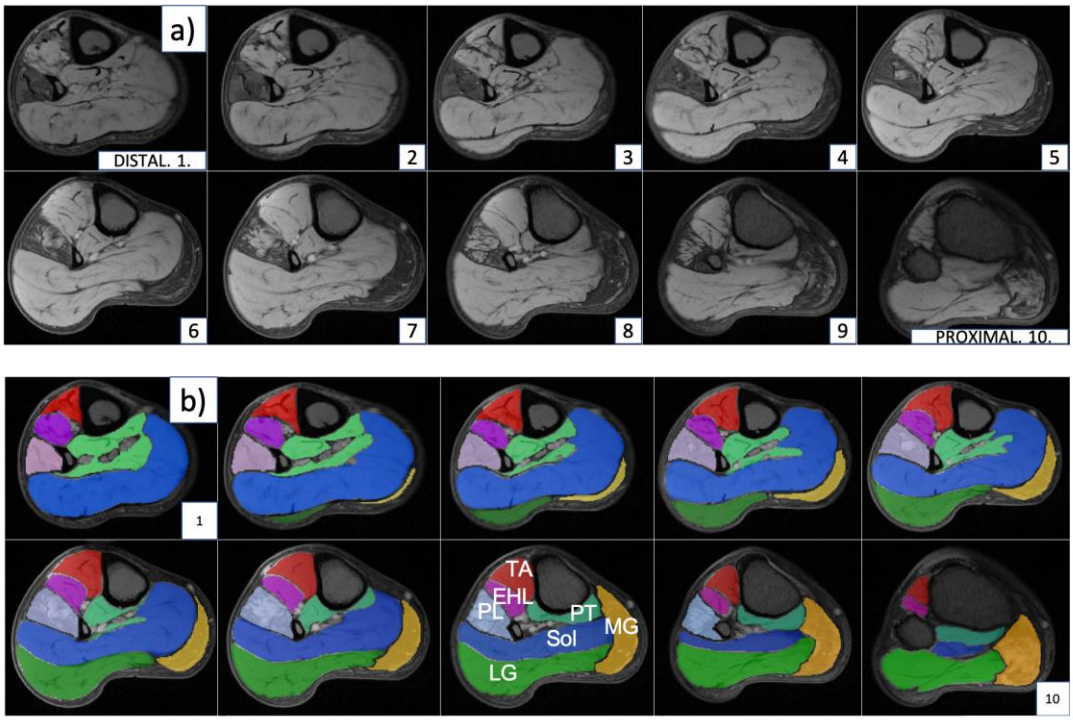
Regions of interest were defined bilaterally for distal thigh, proximal calf and distal calf on selected levels of an unprocessed Dixon acquisition (TE=3.45ms) using ITK-SNAP36. In the distal calf, ROIs were defined for five muscles: tibialis anterior, extensor hallucis longus, peroneus longus, soleus and the deep posterior group. In the proximal calf, ROIs were defined for seven muscles: tibialis anterior, extensor hallucis longus, peroneus longus, medial head of gastrocnemius, lateral head of gastrocnemius, soleus and the deep posterior group. In the thighs, ROIs were defined for ten muscles: vastus medialis, vastus lateralis, vastus intermedius, rectus femoris, gracilis, sartorius, adductor magnus, semimembranosus, semitendinosus and biceps femoris (Figure 2-5).

Figure 2-3 – Muscle segmentation for CMT1A natural history study



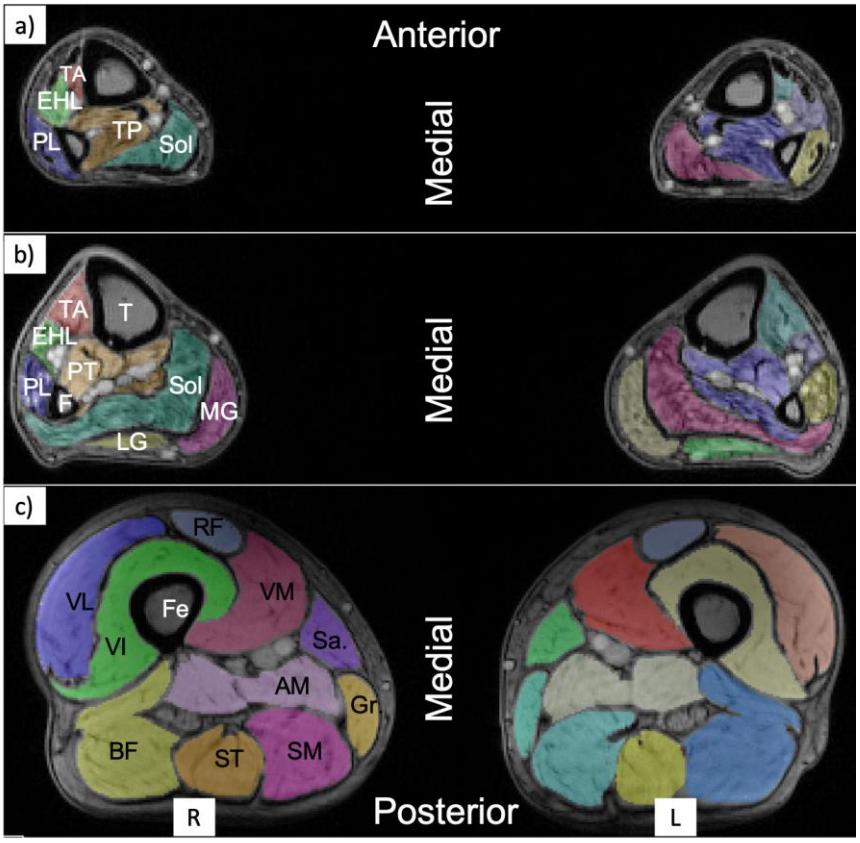
a) Gradient echo (TE=345ms) of the right and left calves as captured in the same field of view. This gradient echo sequence is one of the three sequences used in 3-point Dixon imaging. b) whole muscle ROI. c) small ROI. TA=tibialis anterior, PL=peroneus longus, LG=lateral head of gastrocnemius, MG=medical head of gastrocnemius, Sol=soleus, PT=tibialis posterior, R=right, L=left

Figure 2-4 – Regions of interest segmented for CMT1A gradient study



Panel a): All ten Gradient echo acquisitions (TE=345ms) of the right calf from distal (slice 1 – closest to ankle joint) to proximal (slice 10). Panel b) whole muscle ROIs applied to the GRE sequence. TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral head of gastrocnemius, MG=medical head of gastrocnemius, Sol=soleus, PT=tibialis posterior

Figure 2-5 – Whole muscle segmentation for HSN study



Muscle segmentation at distal calf (a), proximal calf (b) and distal thigh (c) level in a patient with HSN1. TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, MG=medial head of gastrocnemius, LG=lateral head of gastrocnemius, Sol=soleus, PT=deep posterior group, RF=rectus femoris, VM=vastus medialis, VL=vastus lateralis, VI=vastus intermedius, Gr=gracilis, Sa=sartorius, AM=adductor magnus, SM=semimembranosus, ST=semitendinosus and BF=biceps femoris. T=tibia, F=fibula, Fe=femur, R=right, L=left

2.8.3 Transfer of regions of interest to Fat Fraction, T2 and MTR maps

All maps were inspected visually for artefact. Muscles which had areas of gross artefact were excluded from the analysis. Small regions of interest were transferred to the three different co-registered maps (FF, T2 and MTR) for the CMT1A natural history study. Adjustments to the small ROIs were made if necessary to ensure each was within the muscle on all maps. Whole muscle ROIs were transferred to the inherently co-registered FF maps for all three studies. We decided not to use whole ROI for T2 and MTR analysis because of the possibility of capturing non-muscle tissue at the muscle boundaries due to movement between Dixon sequences and the remaining acquisitions.

2.9 Data and statistical analysis

MRC Centre Physicists, Dr Sinclair (baseline and first two follow-ups) and Dr Wastling (final follow-up) extracted using automated software tools mean T2 time, mean FF and mean MTR from the small regions of interest (for CMT1A natural history study), and mean FF and cross-sectional area (CSA) from the whole muscle regions of interest (for all studies). All extracted data was checked for errors and outliers.

In addition to the individual values, summary measures were calculated by the author thus:

- CMT1A natural history study
    - Small ROI (by mean from all relevant ROI)
        - Right and left calf.
        - Combined bilateral calf.
    - Whole ROI (by CSA-weighted average of FF from each relevant ROI)
        - Triceps surae group – soleus, lateral and medial gastrocnemius.
        - Anterior group – tibialis anterior and extensor hallucis longus.
        - Right and left calf.
        - Combined bilateral calf.
- CMT1A gradient study (by CSA-weighted average of FF from each relevant ROI). NB – each summary measure was calculated for each method (a → e) of slice selection.
    - Peroneal group – weighted average of FF from tibialis anterior, extensor hallucis longus and peroneus longus (Figure 2-6).
    - Tibial group – weighted average of FF from tibialis posterior, soleus and the gastrocnemii (Figure 2-6).
    - Whole (right) calf.
- HSN study (by CSA-weighted average of FF from each relevant ROI)
    - Combined bilateral distal calf.
    - Combined bilateral proximal calf.
    - Combined bilateral distal thigh.
    - Combined overall calf (distal and proximal).
    - Right and left calf (both levels) and distal thigh.
    - Severity weighted combination based on baseline fat fraction.

As a measure of the contractile CSA, remaining muscle area (RMA) was defined in all studies, for all measures as:

Equation 2-1 – Calculation of remaining muscle area

$$RMA\,(mm^2) = CSA(mm^2)\,x\left(\frac{(100 - fat\ fraction(\%))}{100}\right)$$

RMA=remaining muscle area, CSA=cross-sectional area. Fat fraction refers to mean or CSA-weighted mean fat fraction as appropriate

Longitudinal changes were quantified for each individual muscle and summary measures, and for each method of slice selection (CMT1A gradient study).

In the CMT1A natural history study, further T2 analysis was done on CMT1A muscles with normal FF, i.e. FF below a threshold as defined by the 95th centile in controls.

Given the non-uniformity of timing of follow-up scans, change in all parameters for each subject was 'annualised' allowing direct comparison longitudinally. In the CMT1A natural history study, the terms 'timepoint two', 'timepoint three' and 'timepoint four' were adopted to refer to

- Timepoint two – annualised results from visit 2.
- Timepoint three – annualised results from visit 3.
- Timepoint four – annualised results from all participants with ≥ three visits.

All data was analysed using IBM-SPSS Statistics Version 26.0. (Armonk, NY; IBM Corp.).

All data were assessed for normality by the Shapiro-Wilk test alongside visual inspection of frequency distribution graphs.

In all studies, responsiveness was evaluated for all outcome measures using the standardised response mean (Equation 1-2). SRM was categorised by magnitude according to Cohen's rule of thumb: < 0.2 minimal responsiveness; 0.2-0.5 small responsiveness; 0.5-0.8 moderate responsiveness; >0.8 large responsiveness.

2.9.1 CMT1A natural history study

Summary statistics for CMT1A and control groups of each quantitative measure on an individual muscle, and summary measure basis were calculated (mean ± standard deviation). Mean baseline measures and mean change in individual muscle and summary measures were compared within groups by paired t-test or Wilcoxon Rank Sum test, and between groups using independent two tailed t-test or Mann-Whitney U test as appropriate. Inter-muscle differences were assessed using ANOVA with post hoc comparisons using Bonferroni's method. Correlations were assessed using Pearson or Spearman correlation coefficients as appropriate. Linear regressions

were performed at baseline to assess separately, the dependence of T2 and MTR on FF. Patients were grouped based on baseline FF (as determined by 95th centile in controls): <6.4%, 6.4-60% and >60% and rate of change in FF compared according to these categories. Similar analysis was done based on STIR hyperintensity. One-way ANOVA with post-hoc Tukey analysis was used to identify if any baseline MRI muscle characteristics were predictive of subsequent change. Statistical significance was set at <5%.

2.9.2 CMT1A gradient study

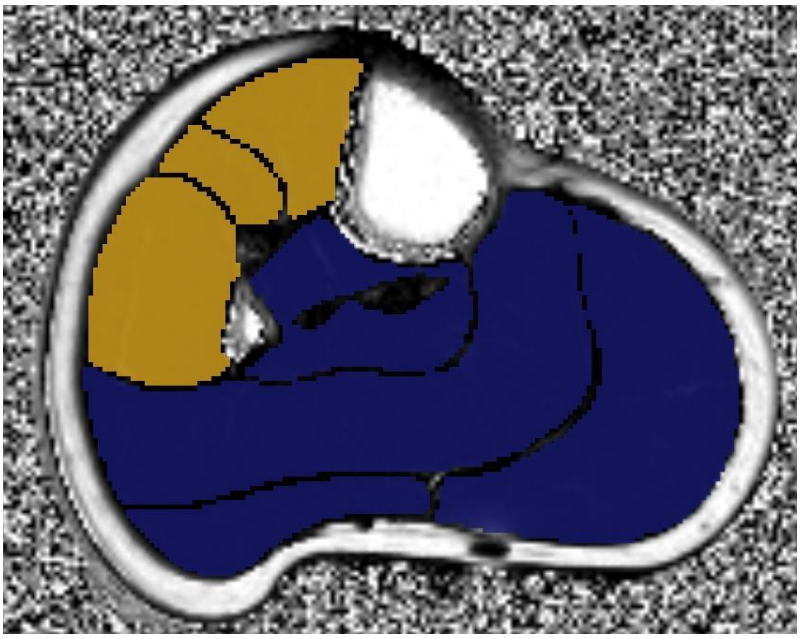
All baseline analyses were performed with FF, CSA and RMA values calculated by method c (Figure 2-2) whereas longitudinal analysis was done for each of the five methods of slice selection. Mean FF of individual muscles within the CMT1A group were compared using ANOVA with post hoc comparisons using Bonferroni's method. Mean baseline measures and mean change in individual muscle and summary measures were compared within groups by paired t-test or Wilcoxon ranked-sum test, and between groups using independent two tailed t-test or Mann-Whitney U test as appropriate. A positive FF gradient was recorded if in the muscle concerned there was an increase in FF over three consecutive axial slices, and the average absolute change over that distance was at least 2%/cm. A negative FF gradient was recorded if the average absolute change in FF was -2%/cm over three consecutive slices. The occurrence of distal-proximal muscle fat-fraction gradients for each participant at baseline and follow-up were represented graphically, and the frequency and distribution of gradients calculated for each muscle and summary measure. Correlations were assessed using Pearson or Spearman correlation coefficients as appropriate. Patients were grouped based on baseline FF (normal FF defined by 95th centile in controls): <4.8%, 4.8-70% and >70%, and rate of change in FF compared according to these categories.

2.9.3 HSN study

For baseline data, mean ± standard deviation or median/interquartile range are given as appropriate. Inter-muscle differences were assessed using ANOVA with post hoc comparisons using Bonferroni's method. Independent sample t-test or Mann-Whitney U test (as appropriate) was used for comparison between controls and patients for MRI FF measurements. Longitudinal change is presented as means/95% CI or medians/IQR as appropriate. Comparison of mean change was evaluated using one-sample two-tailed t-tests or one-sample Wilcoxon signed-rank tests as appropriate. Patients were grouped based on baseline FF: <20%, 20-70% and >70% and rate of

change in FF compared according to these categories. Statistical significance was set at <5%.

Figure 2-6 – Right calf axial MRI slice demonstrating segmentation of muscles into peroneal and posterior tibial innervated groups for the CMT1A gradient study



Yellow muscles are innervated by the peroneal nerve, and blue muscles by the posterior tibial nerve

3 Charcot-Marie-Tooth disease type 1A – natural history study

3.1 Introduction

Charcot-Marie-Tooth disease type 1A (CMT1A) is the most common inherited peripheral neuropathy with prevalence estimated to be between 10 and 80/100,000 worldwide, (Kurihara et al., 2002; Braathen, 2012). It is a demyelinating, length-dependent, motor and sensory neuropathy caused by a 1.5 Mb duplication of the region of chromosome 17p11.2 which contains the 22 kilodalton peripheral myelin protein *(PMP-22)* gene. As the disease slowly progresses, denervated muscle atrophies and is replaced by intramuscular fat, contributing to distal limb weakness. Although much is known about its pathogenesis, there are no current drug treatments for CMT1A, although rapid progress has been made (D'Ydewalle et al., 2011; Garofalo et al., 2011; Gutmann and Shy, 2015; Patzkó et al., 2012; Saporta et al., 2015). Emerging potential treatments require rigorous assessment in scientifically sound clinical trials, for which appropriate outcome measures are needed.

Measuring change in such a slowly progressive neuropathy over a reasonable time period presents great challenges. As discussed in Chapter 1, many different outcome measures have been used over the past 10-15 years, though each has been limited by its poor sensitivity to change. Quantitative MRI (qMRI) is an ideal outcome measure for CMT1A and other neuromuscular diseases (NMD). The pathological consequences of both primary muscle and nerve diseases: intramuscular water accumulation and fat deposition, can be readily quantified by MRI, which is safe, non-invasive, and non-operator or participant dependent.

To date in NMD, qMRI has been proven reliable, valid and responsive in several mostly muscle diseases (Table 3-1). With respect to CMT1A, the first year of our Centre's CMT1A/IBM natural history study of which the continuation is presented in this chapter, showed qMRI determined combined bilateral calf fat fraction (FF) to be more responsive than any previously or currently used outcome measure in CMT1A with significant FF change of 1.2 ± 1.5% and standardised response mean (SRM) of 0.83 over 12 months (Morrow et al., 2016). This was a ground-breaking finding for NMD research, however further work is needed to identify and refine responsive outcome measures in preparation for rapidly approaching clinical trials.

The importance of maximising outcome measure responsiveness cannot be overstated for reasons discussed in Chapter 1, chief amongst which are the slow

progression and rarity of the diseases being studied. Of equal importance is ensuring that the change which is detected is clinically meaningful and relevant to the patient.

There are two types of responsiveness as defined by Husted and colleagues. Internal responsiveness characterises 'the ability of an outcome measure to change over a particular prespecified time period' – this can be assessed over the course of a randomised controlled trial of a medication which has previously been shown to be effective, or in natural history studies as is the case in the current study. In this study, outcome measure internal responsiveness is expressed by the SRM, which is the ratio of mean change to standard deviation of change. The magnitude of the SRM is critical and can be classified according to Cohen's rule of thumb as: <0.2 minimal responsiveness; 0.2-0.5 small responsiveness; 0.5-0.8 moderate responsiveness; >0.8 large responsiveness. Responsiveness is a key determinant of study power through its' inverse square relationship with sample size as detailed in Chapter 1, such that a two-fold increase in SRM results in a four-fold decrease in sample size for the same study power.

On the other hand, external responsiveness (which the author refers to in this thesis as longitudinal validity), reflects the extent to which change in an outcome measure relates to corresponding change in a reference measure (i.e. a gold standard) or metric of health status which is accepted to indicate change in the patient's condition (Husted et al., 2000). For studies in inherited neuropathies, the current gold standard is the CMT neuropathy score (CMTNS), which has been shown to be reliable and a valid marker of disease severity in CMT (Shy et al., 2005). If there is a clearly strong relationship between the new outcome measure (e.g. qMRI in this case) and the gold standard, this then provides support that the new instrument measures clinically meaningful change, and that it can be put forward for consideration to be used as a primary outcome measure in clinical trials. Demonstration of longitudinal validity is particularly important as it is generalisable, providing evidence for the novel outcome measure to be used beyond the disease in which the longitudinal validity has been demonstrated (Husted et al., 2000).

The study presented in this chapter extends qMRI, clinical and functional outcome measurements in our CMT1A cohort across five years. A battery of outcome measures are examined, assessing internal responsiveness, cross-sectional validity and where possible, longitudinal validity over prolonged follow up.

3.2 Background literature

Longitudinal studies in neuromuscular disease (NMD) have most often assessed muscle qualitatively or semi-quantitatively. Studies evaluating qMRI determined outcome measures are not plentiful, with the vast majority done in patients with muscle disorders such as Duchenne muscular dystrophy (DMD), driven by emerging treatments and a recent explosion in drug trials (Haas et al., 2015; McDonald et al., 2017). Beyond the studies in DMD, there are a number of natural history studies and a scattered handful of interventional studies reporting on longitudinal qMRI. Neuropathies have barely been studied by qMRI, with only two such natural history studies in CMT1A – both from our Centre. Table 3-1 summarises details of longitudinal studies which have used qMRI determined FF as an outcome measure.

3.2.1 Quantitative MRI longitudinal studies

Here I summarise pertinent longitudinal studies in NMD which use qMRI measures: FF, T2 time, magnetisation transfer ration (MTR) and cross-sectional area (CSA) as outcome measures, before presenting results of our CMT1A natural history study.

3.2.1.1 Fat fraction and T2 relaxation time

3.2.1.1.1 Duchenne Muscular Dystrophy

3.2.1.1.1.1 Natural history studies

Bonati and colleagues conducted a 12 month natural history study in 20 boys with DMD, comparing 2-point Dixon MRI of thigh muscles with clinical measures. The authors reported a significant FF change in all thigh muscles of 6.3 ± 3.7% (SRM 1.70), and in the adductor muscles of 7.3 ± 4.4%. (SRM 1.66). When the group of boys was subdivided into ambulant and non-ambulant, overall FF change in thigh muscles in the non-ambulant group was 9.1 ± 3.2% (SRM 2.80). The responsiveness of qMRI was far greater than that of clinical measures (Bonati et al., 2015). In Ricotti et al's recent natural history study, qMRI using the 3-point Dixon method was used to assess upper limb fat infiltration in 15 non-ambulant boys with DMD at four time points over 12 months. There was a significant increase in central forearm FF over 12 months, reaching significance from 6 months compared with controls: FF change 3.9% (95% CI 1.9 to 5.7) at 6 months and 5.0% at 12 months – 95% (95% CI 3.2 to 6.9%), which translates to an SRM of 1.91 and 2.50 at 6 and 12 months respectively. This far exceeded responsiveness of other MRI derived, functional and strength measures (Ricotti et al., 2016). Also in the upper limbs in DMD, Hogrel et al. examined 25 boys (of which 15 were non-ambulatory) with 3-point Dixon MRI/MR spectroscopy at baseline and a 12 month interval, alongside clinical assessments including hand

grip and key pinch strength. FF changed by 1.20% (p=0.09) in ambulatory and 3.20% (p=0.014) in non-ambulatory patients, though SRM was low in each group, indeed not much higher than that of hand-held grip strength, probably due to heterogeneity of the group. Other studies in patients with DMD have also shown progressive increase in muscle FF over time (Forbes et al., 2014; Marden et al., 2005; Torriani et al., 2012) with greater responsiveness than functional and clinical measures. Willcocks et al. measured longitudinal calf T2 time in a cohort of 16 boys with DMD and 15 controls. In the soleus muscle, mean T2 time increased significantly over both 12 months and 24 months (4.5 ± 3.6ms, SRM=1.25), with a more rapid increase in older boys. This is likely to be predominantly the effect of increasing FF. The fat suppressed T2 increased by only a small amount over the study period (Willcocks et al., 2014).

3.2.1.1.1.2 Interventional studies

In a cohort of 15 DMD patients taking steroids, and 15 matched DMD patients who had never taken steroids, Arpan and colleagues measured right lower limb T2 relaxation time and FF in two muscles by MRI/MR spectroscopy. They found that over 12 months, there was a greater FF increase in patients not treated with steroids, and that those patients who started steroids demonstrated reduced T2 values consistent with improvement in muscle oedema, and reduced T2 change compared with steroid naïve patients (Arpan et al., 2014). An earlier study by Kim and colleagues assessed T2 time longitudinally in a group of patients with DMD taking steroids. The authors reported variability in T2 response to treatment, though there was no measurement of fat making interpretation of T2 values difficult. Absence of a control group also made longitudinal interpretation of limited benefit. Limited reporting of descriptive statistics did not allow SRM to be calculated (Kim et al., 2010).

3.2.1.1.2 Becker Muscular Dystrophy

In Becker muscular dystrophy, Bonati and colleagues examined thigh muscles of three patients at baseline and 12 months using 2-point Dixon MRI, comparing this with a functional scale. Baseline thigh FF was 61.6 ± 7.6%, and FF increase over 12 months was 3.7 ± 4.7%. Study numbers were small, and p value for FF change was not given (Bonati et al., 2015).

3.2.1.1.3 Pompe disease

3.2.1.1.3.1 Interventional studies

In an open-label retrospective study in 23 patients with Pompe disease, of whom 14 were on enzyme replacement therapy (ERT), patients were scanned at least twice

and some several times during a four-year period. A modified 3-point Dixon sequence for fat quantification detected overall FF increase of 0.9 ± 0.2%/year with fastest progression in hamstrings and adductor muscles. Patients on ERT progressed significantly slower than untreated patients. The authors reported that when muscle $T2_{water}$ was abnormal, fat progression was 0.61 %/year greater than when T2 was normal (p= 0.02). Longitudinal change in $T2_{water}$ time was not reported. (Carlier et al., 2015). Also in Pompe disease, Figueroa-Bonaparte and colleagues examined thigh muscles of 32 patients with 3-point Dixon MRI alongside patient reported outcome measures and clinical assessment. FF change as measured by qMRI was 1.79 ± 3.05% over 12 months in symptomatic patients (SRM 0.59) (Figueroa-Bonaparte et al., 2018). Finally, the EMBASSY study examined 16 patients with Pompe disease, but found no change in 3-point Dixon after six months (van der Ploeg et al., 2016).

3.2.1.1.4 Oculopharyngeal muscular dystrophy

In OPMD, Fischmann et al. compared thigh and calf 2-point Dixon and T2w qMRI at 1.5 Tesla with semiquantitative MRI and functional assessments in a cohort of eight patients and five controls at baseline and after 13 months. Overall thigh and calf FF increased by 19.2% to 20.7% ($p<0.001$) and T2 time increased by 49.4 to 51.6ms ($p<0.001$) over 12 months with no change in controls. There were no changes in all other measures. Descriptive statistics were not reported (Fischmann et al., 2011).

3.2.1.1.5 Spinal muscular atrophy type II and III

Eighteen patients and 19 controls were assessed with qMRI (T2 mapping and 2-point Dixon) of both thighs over 12 months (at 6, 9 and 12 months), alongside clinical evaluation. A single whole muscle FF was calculated for each patient at each timepoint. SRM of MRI measures (0.66 for overall 'multipeak' FF) exceeded that of clinical measures. There was no change in T2 time over the duration of the study (Bonati et al., 2017). Also in SMA, Chabanon and colleagues collected detailed baseline data including by 3-point Dixon MRI in a cohort of 81 SMA II and III patients with plans to publish longitudinal data soon (Chabanon et al., 2018).

3.2.1.1.6 Limb Girdle Muscular Dystrophy

In LGMD2I, 32 patients from four centres were examined at baseline and 12 months with lower limb thigh and calf 2D 3-point Dixon and standard T1 weighted MR at 3 Tesla alongside functional (10 metre walk, timed up and go, six minute walk test), and muscle strength assessment using handheld myometer. Fat fraction measured by the Dixon method increased in 9/14 lower limb muscles, though descriptive details were inadequate in the manuscript to calculate responsiveness. There were no significant

changes over the study period in T1 weighted MR and other measures collected (Willis et al., 2013). The cohort was reassessed after a period of six years, at which point there had been ongoing significant increase in FF at thigh (30.8% to 47.3%, p<0.001. SRM 1.20) and calf level (15.4% to 26.8%. p<0.001. SRM 1.28), in addition to significant but less responsive decline in six minute walk test (6MWT), 10m walk, timed up and go, and chair rise (Murphy et al., 2019).

3.2.1.1.7 Spinobulbar muscular atrophy and motor neurone disease

Dahlqvist and colleagues examined 29 patients with spinobulbar muscular atrophy (SBMA) at baseline and 18 months, using 2-point Dixon MRI (lower limbs and right upper limb) as the primary outcome measure, alongside stationary dynamometry, 6MWT, and functional rating scales. Overall FF increased by 2.0 ± 1.25%, as too did FF in all individual lower limb and upper limb muscles, but CSA did not change. Clinical measures also changed significantly: knee extension and handgrip strength, 6MWT, stair climbing time, but functional rating scale scores did not change (Dahlqvist et al., 2018). Bryan et al. showed progression in the T2 time in the tibialis anterior muscle over a four month period in 11 patients with motor neuron disease (MND). They reported that T2 time correlated with compound motor action potential amplitude cross-sectionally and longitudinally (Bryan et al., 1998).

3.2.1.1.8 Facioscapulohumeral Dystrophy

Andersen et al. conducted a prospective natural history study in 45 patients with facioscapulohumeral muscular dystrophy (FSHD) using 2-point Dixon MRI at 3 Tesla to assess progression in thigh, calf and paraspinal muscles over a mean of 436.3 ± 43 days. Positioning for follow up was based on the initial scan, and corrected by position of lung apex and acetabulum. Two different observers segmented baseline and follow up, though FF calculated from their baseline scan region of interest (ROIs) correlated strongly. ROIs were of muscle groups rather than individual muscles. Other measures collected included FSHD score, muscle strength using a handheld dynamometer, 6MWT, a 14 step stair test and sit to stand. There was significant change in overall FF of 3.6% (CI 2.6-4.6, p<0.001) with significant progression in all individual muscles as well. There were also significant changes in muscle strength measured by MRC, and in FSHD score, which worsened by 10% (p<0.05) (Andersen et al., 2017), though qMRI was more responsive than functional measures. Unfortunately there was no control group in this study which somewhat hinders interpretation of longitudinal results.

3.2.1.1.9 Charcot-Marie-Tooth disease

As mentioned, in CMT1A, Morrow et al. reported on the first year of our Centre's CMT1A/IBM natural history study (Morrow et al., 2016) demonstrating that FF by 3-point Dixon provided a highly responsive outcome measure over 12 months, with muscle FF increase of 1.2% ± 1.5% (p = 0.008) and SRM of 0.83. There was no significant change in muscle T2 time or MTR. Our Centre has since demonstrated the validity of calf muscle MRI FF by the MRC Centre's lower limb MRI protocol. Ten patients with CMT1A were examined at a remote site (Iowa, USA). Excellent test-retest reliability was found at the remote site. qMRI also demonstrating an increase in calf level FF of 1.8 ± 1.7% (SRM=1.04) across the cohort, with SRM increasing to 2.19 when only patients with calf FF >10% at baseline were analysed (Morrow et al., 2018).

Table 3-1 – Fat fraction change in longitudinal studies using quantitative MRI fat fraction as an outcome measure

Details of longitudinal slice	Cohort	MRI	n	FF mean change ± sd	Duration of study	p	SRM
Both thighs	Pompe[10]	3-point Dixon	32	1.8 ± 3.1	12 months	0.001	0.59
Combined calf and thigh	Pompe[1]	modified 3-point Dixon	23	0.9 ± 0.2	12 months	p<0.001	4.50
Both thighs	DMD[2]	2-point Dixon	20	6.3 ± 3.7	12 months	p<0.001	1.70
Both quadriceps	DMD[2]	2-point Dixon	20	6.4 ± 4.7	12 months	p<0.001	1.36
Both hamstrings	DMD[2]	2-point Dixon	20	5.2 ± 2.8	12 months	p<0.001	1.86
Both adductors	DMD[2]	2-point Dixon	20	7.3 ± 4.4	12 months	p<0.001	1.66
Both thighs in ambulant group >7y	DMD[2]	2-point Dixon	20	9.1 ± 3.2	12 months	p<0.001	2.80
Both forearms	DMD[3]	3-point Dixon	25	3.2 ± 8.5	12 months	0.014	0.38
Both thighs	BMD[4]	2-point Dixon	3	3.7 ± 4.7	12 months	-	0.79
Unilateral forearm	DMD[5]	3-point Dixon	7	5.0 ± 2.0	12 months	-	2.50
Unilateral forearm	DMD[5]	3-point Dixon	7	3.9 ± 2.1	6 months	-	1.9
Bilateral - 5 levels combined	FSHD[6]	2-point Dixon	45	3.6 ± 3.3	12 months	<0.001	1.10
Both thighs (3 slices each leg)	SMA[7]	2 and 6-point Dixon	18	-	12 months	<0.0001	0.23
Thigh and calf	OPMD[8]	2-point Dixon	5	19.2 to 20.7%	12 months	significant	-
Both calves	CMT[9]	3-point Dixon	20	1.2 ± 1.5	12 months	0.002	0.83
Both calves	IBM[9]	3-point Dixon	20	2.6 ± 2.9	12 months	0.002	0.90
Both thighs	IBM[9]	3-point Dixon	20	3.3 ± 3.3	12 months	0.0007	1.00
Combined over 5 anatomical areas	SBMA[11]	3-Point Dixon	29	2.0 ± 1.3	18 months	<0.001	1.60
Both calves	CMT[12]	3-point Dixon	10	1.8 ± 1.7	12 months	0.009	1.04
Both calves - baseline FF >10%	CMT[12]	3-point Dixon	5	2.9 ± 1.3	12 months	0.008	2.19
Individual lower limb muscles	LGMD2I[13]	3-point Dixon	32	between 1.3 and 21%	12 months	0.004 to 0.025	-
Both thighs	LGMDR9[14]	3-point Dixon	23	30.8 to 47.3%	6 years	<0.001	1.20
Both calves	LGMDR9[14]	3-point Dixon	23	15.4 to 26.8%	6 years	<0.001	1.28

p=p value for paired t-test or Wilcoxon test as appropriate. N=number of participants in arm, sd=standard deviation, FF=fat fraction (%), IBM=inclusion body myositis, CMT=Charcot-Marie-Tooth disease, DMD=Duchenne Muscular Dystrophy, FSHD=facioscapulohumeral dystrophy, SMA=spinal muscular atrophy, BMD=Becker Muscular Dystrophy, SBMA=spinobulbar muscular atrophy 1) (Carlier et al., 2015), 2) (Bonati, Hafner, et al., 2015), 3) (Hogrel et al., 2016), 4) (Bonati et al., 2015), 5) (Ricotti et al., 2016), 6) (Andersen et al., 2017), 7) (Bonati et al., 2017), 8) (Fischmann et al., 2011), 9) (Morrow et al., 2016), 10) (Figueroa-Bonaparte et al., 2018), 11) (Dahlqvist et al., 2018), 12) (Morrow et al., 2018), 13) (Willis et al., 2013), 14) (Murphy et al., 2019)

3.2.1.2 Cross-sectional area and remaining muscle area

CSA is an alternative to MRI derived tissue measures which can be assessed longitudinally, though its measurement is complicated by the effects of aging, gender, and a large disparity in individual lower limb muscle size. Remaining muscle area (RMA) which is also referred to in the literature as contractile CSA, is derived from a

combination of CSA and FF, and in theory may be expected to be more responsive than CSA.

Several studies, including some of those discussed above have found no significant longitudinal change in CSA despite significant changes in FF: in SMA (Sproule et al., 2011), CMT (Morrow et al., 2016), IBM (Morrow et al., 2016), IBM (Spector et al., 1997), SBMA (Dahlqvist et al., 2018), LGMD R9 (Murphy et al., 2019), but others have reported positive results.

In their study of non-ambulatory patients with DMD, Ricotti and colleagues reported a significant increase ($p<0.01$) in central forearm CSA of 140mm^2 (95% CI 50.9 to 229.1) at 12 months, (but not 3 or 6 months), giving an SRM of 1.45. There was no change in CSA at other forearm levels (Ricotti et al., 2016). The authors did not report results of longitudinal correlation between CSA and FF change which would have been helpful in terms of assessing longitudinal validity.

In a small group of patients with diabetic neuropathy and matched controls, Andreassen et al. showed a loss of muscle volume of 4.5% and 5.0%/ year in ankle dorsiflexors and plantarflexors respectively, compared with loss of 1.7% and 1.8%/year in controls (Andreassen et al., 2009). The paper did not report standard deviations, and the prolonged follow up of 12 years makes interpretation difficult.

Murphy et al. reported a significant reduction in RMA in 8/14 individual lower limb muscles, and all combined muscles at six years in a cohort of 23 patients with LGMD R9. Maximum SRM was in semitendinosus which changed from 315mm^2 to 206mm^2 over six years (SRM -0.71). The whole thigh and whole calf changed from 3288mm^2 to 3087mm^2 and 3637mm^2 to 3365mm^2 respectively with SRM of -0.55 and -0.52 for thigh and calf respectively. In each case however, SRM for the RMA was smaller than for the corresponding FF (Murphy et al., 2019).

Amato and colleagues used right thigh volume by T1w MRI as the primary outcome measure in their phase 2 clinical trial of treatment of 11 sIBM patients with bimagrumab (3 other patients received placebo). This decision was based on the effect of the drug which is to block the TGF-beta pathway leading to muscle atrophy. The authors reported an increase in thigh volume at 8 weeks (6.5% increase in right thigh [p=0.009], 7.6% increase in left thigh, [p=0.02]) in the treatment arm. The change correlated with significant change in 6MWT (Amato et al., 2014). It is not yet clear whether increase in muscle CSA leads to improved function – this needs further investigation. Pichiecchio et al. found a small increase in muscle volume of 7.5% of the less clinically affected anterior thigh compartment in 11 patients with Pompe

disease treated with ERT. There was an associated 45% increase in muscle strength, which the authors interpreted as an increase in muscle quality alongside the small increase in size. There was no CSA change in the more clinically affected posterior thigh (Pichiecchio et al., 2009). There was also ongoing lower limb fat accumulation during treatment, despite increased CSA, perhaps indicating that ERT is ineffective on FF once past a certain threshold

3.2.1.3 Longitudinal correlation with strength and function

Evidence of strong cross-sectional correlation between qMRI measures and patient strength (Kan et al., 2009; Sproule et al., 2011; Hiba et al., 2012; Janssen et al., 2014) and function (Kim, et al., 2010; Torriani et al., 2012; Wren et al., 2008) is well recognised for many NMD including CMT1A. Longitudinal correlation however is less well reported, and when present is often less strong than cross-sectional correlations.

In their six year follow-up study in LGMDR9, Murphy et al. reported that changes in 6MWT and FF across six years correlated moderately in soleus ($r=-0.6$, $p<0.05$), rectus femoris and weakly in some composite muscle groups: thigh ($r=-0.47$, $p<0.05$) and quadriceps ($r=-0.46$, $p<0.05$). There was weak correlation between FF change and 10-metre walk across six years ($r=-0.52$, $p<0.05$) (Murphy et al., 2019). In IBM, Morrow et al. found correlation between change in quadriceps RMA and knee extension: $r=0.66$ on right ($p=0.005$) and 0.81 on left ($p=0.0001$), but not with FF change (Morrow et al., 2016). Andersen et al. found no correlation between progression in FSHD score and FF or between strength and FF change over 12 months (Andersen et al., 2017). Similarly, Dahlqvist found no correlation between progression in FF and progression in strength or functional measures in SBMA (Dahlqvist et al., 2018). Several longitudinal studies mentioned above did not report on longitudinal correlation. In CMT1A, or other inherited neuropathies, longitudinal correlation between qMRI and other measures has not been demonstrated.

3.2.1.4 Method of longitudinal slice selection

The method by which the longitudinal slice(s) was selected was variable across studies, and when mentioned, based on measurement from either surface anatomy or scout imaging. It has been shown that scout image based slice selection results in narrower limits of agreement and higher intraclass correlation coefficients compared with surface-anatomy-based slice selection (Fischmann et al., 2014).

3.2.2 Summary

Over the past 5-10 years, there has been an expanding number of publications examining qMRI outcome measures longitudinally in NMD. Most studies have

compared qMRI measure responsiveness to that of other established outcome measures, resulting in good evidence that FF measured by the Dixon method is a highly responsive outcome measure in a number of NMD, with responsiveness almost universally surpassing that of other outcome measures. There is a much smaller volume of literature relating to inherited neuropathies however. As demonstrated in Table 3-1, qMRI FF responsiveness is variable, and as discussed in Chapter 1, is a function of the particular disease and cohort of patients under study.

There is also good evidence that qMRI determined CSA is not as sensitive to change as is FF, though this has not been examined as rigorously. Longitudinal assessment of CSA suffers from several drawbacks including a general paucity of natural history data in controls and NMD, and confounding effects of ageing associated muscle atrophy, variations in muscle size between individuals, high standard deviations in longitudinal change due to difficulties in measurement, and the unclear effect of medications on muscle size, and benefit of increased muscle size. RMA shows more promise in recent studies, but on current evidence appears to be less responsive than FF.

Cross-sectional validity of qMRI FF is well demonstrated by strong correlation with strength and function, but longitudinal validity is less well shown, and not demonstrated for qMRI in any neuropathies.

In almost all qMRI studies mentioned above, FF derived from a single axial level (single or bilateral whole calf or thigh) has been analysed longitudinally. Three studies (Andersen et al., 2017; Bonati et al., 2017; Dahlqvist et al., 2018), have used multiple slices, and only one of these studies assessed three contiguous slices (Bonati et al., 2017), but without comparing findings to a single slice. Several studies discussed have considered effects of stratifying patients based on clinical (Bonati et al., 2015; Morrow et al., 2018) or qMRI baseline characteristics, but this needs further detailed assessment.

3.3 Study aims

The aims of this longitudinal study are threefold. Firstly to examine internal responsiveness of a battery of outcome measures including qMRI over prolonged follow-up in a cohort of patients with Charcot-Marie-Tooth disease type 1A. Secondly, to correlate any significant longitudinal change in qMRI measured parameters with significant changes in clinical and functional outcome measures thus establishing longitudinal validity, and thirdly to assess patterns of muscle fat accumulation, and baseline predictors of fatty accumulation. The chapter will conclude with an evidence-based discussion of the optimal use of qMRI as an outcome measure for upcoming trials in Charcot-Marie-Tooth and related neuromuscular diseases.

3.4 Cross-sectional results

The cross-sectional analysis in this CMT1A natural history study is reflected in previous work by my colleague Dr Jasper Morrow (see acknowledgements) and will therefore be summarised here briefly and where relevant to that part of the natural history study central to this thesis. It should be noted that for the purposes of this thesis, all muscle regions of interest were redefined by the author (MRBE), blinded to diagnosis and visit number.

3.4.1 Study participants

20 patients with CMT1A, and 20 healthy matched controls were assessed at baseline. Age, gender, height and weight were well matched between CMT1A and control groups (Table 3-13).

3.4.2 MRI data

Summary of baseline MRI, myometric and clinical data is given in Table 3-2.

Table 3-2 – Baseline calf data in CMT1A and controls

	CMT1A	Controls	p
	Clinical measures		
Age at CMT onset (years)	5.9 ± 4.2	NA	NA
Disease duration (years)	36.7 ± 16.5	NA	NA
CMTSS (0-12)	3.1 ± 2.0	NA	NA
CMTES (0-28)	8.0 ± 5.1	NA	NA
MRC-LL (0-110)	95.4 ± 15.4	NA	NA
MRC total (0-230)	207.8 ± 21.8	NA	NA
SF36 (0-100%)	73.9 ± 15.2	NA	NA
SF36-PF (0-100%)	65.0 ± 23.0	NA	NA
	MRI measures		
Fat fraction small (%)	15.5 ± 25.3	1.7 ± 1.0	0.02
Fat fraction whole (%)	16.2 ± 25.2	3.2 ± 1.6	0.03
Fat fraction R. Calf (%)	14.5 ± 24.0	2.7 ± 1.6	0.04
Fat fraction L. Calf (%)	18.1 ± 26.5	3.8 ± 1.7	0.02
T2 (ms)	66.5 ± 32.5	44.8 ± 3.6	0.003
MTR (pu)	26.1 ± 9.0	32.0 ± 0.9	0.001
CSA (cm^2)	92.8 ± 24.0	117.4 ± 27.3	0.006
RMA (cm^2)	79.7 ± 33.7	113.6 ± 26.1	0.002
	Myometry measures		
Ankle plantarflexion (nM)	33.2 ± 19.5	66.8 ± 20.6	<0.0001
Ankle dorsiflexion (nM)	9.8 ± 7.9	35.7 ± 13.3	<0.0001
Ankle inversion (nM)	15.2 ± 10.1	19.3 ± 6.8	0.04
Ankle eversion (nM)	7.4 ± 4.4	20.8 ± 7.7	<0.0001

Values are mean ± standard deviation. MRC=Medical Research Council score, whole=all-muscle region of interest value, NA=not applicable, CMTSS/ES=Charcot-Marie-Tooth symptom/examination score, SF36-PF=Short-Form 36 Quality of Life Score-physical function domain, p value is for Mann-Whitney U test or independent samples two-tailed t-test where appropriate (CMT1A versus control group), MTR=magnetisation transfer ratio, nM=Newton metres, CSA=cross-sectional area, ms=milliseconds

3.4.2.1 Measures of chronic fatty atrophy

Cross-sectional analysis of FF and CSA demonstrate the dual processes of intramuscular fat accumulation and atrophy in CMT1A patients.

3.4.2.1.1 Fat fraction

In the CMT1A group at baseline, combined bilateral whole calf FF was 16.2 ± 25.2% compared with 3.2 ± 1.6% in controls (p=0.03). Combined bilateral calf FF measured by small regions of interest was 15.5 ± 25.3% in CMT1A and 1.7 ± 1.0% in controls (p=0.02). There was a significant difference between CMT1A and controls in all summary calf muscle FF measures except for right triceps (p=0.07). There were also significant intergroup differences in several individual muscles. Baseline MRI determined FF values in CMT1A and controls are summarised in Table 3-3.

Table 3-3 – Baseline MRI measures in CMT1A and controls

	CMT1A	Control		CMT1A	Control		CMT1A	Control	
Muscle	Mean fat fraction ± sd (%)		p	Mean CSA ± sd (cm^2)		p	Mean RMA ± sd (cm^2)		p
R. TA	13.1 ± 21.1	1.8 ± 1.0	0.02	6.2 ± 1.8	9.3 ± 2.6	<0.0001	5.5 ± 2.3	9.4 ± 2.4	<0.0001
R. PL	23.7 ± 28.0	3.4 ± 2.5	0.003	4.7 ± 1.7	4.9 ± 1.6	0.50	3.7 ± 2.1	4.7 ± 1.6	0.09
R. LG	15.4 ± 28.5	2.6 ± 1.5	0.06	5.8 ± 2.5	6.9 ± 2.4	0.11	5.1 ± 3.1	6.7 ± 2.3	0.08
R. MG	18.4 ± 28.5	2.8 ± 1.2	0.03	10.1 ± 4.0	12.8 ± 3.8	0.08	8.7 ± 4.9	12.4 ± 3.7	0.03
R. Sol	12.2 ± 23.8	3.2 ± 2.7	0.10	15.2 ± 6.4	20.8 ± 5.0	0.002	13.8 ± 7.2	20.2 ± 4.8	0.001
R. PT	12.1 ± 21.7	1.7 ± 1.0	0.04	4.7 ± 1.2	3.8 ± 1.3	0.04	4.2 ± 1.5	3.8 ± 1.3	0.31
R Triceps	14.0 ± 25.6	2.9 ± 1.8	0.07	31.2 ± 9.3	40.2 ± 10.0	0.02	27.6 ± 12.2	39.3 ± 9.5	0.002
R. Calf	14.5 ± 24.0	2.7 ± 1.6	0.04	46.8 ± 12.7	58.9 ± 13.8	0.006	40.9 ± 17.1	57.2 ± 13.3	0.003
L. TA	14.6 ± 21.7	2.3 ± 1.1	0.02	6.3 ± 2.0	9.7 ± 2.4	<0.0001	5.5 ± 2.5	9.5 ± 2.4	<0.0001
L. PL	20.5 ± 27.2	4.4 ± 2.7	0.01	5.0 ± 1.7	4.9 ± 1.2	0.90	4.1 ± 2.1	4.7 ± 1.2	0.32
L. LG	17.1 ± 28.7	3.0 ± 1.7	0.04	5.2 ± 1.5	7.2 ± 2.3	0.01	4.4 ± 2.1	7.0 ± 2.3	0.001
L. MG	19.9 ± 29.5	4.2 ± 1.7	0.03	9.8 ± 3.2	12.6 ± 4.4	0.08	8.3 ± 4.2	12.0 ± 4.3	0.04
L. Sol	12.6 ± 22.3	3.4 ± 1.6	0.08	15.3 ± 5.7	20.4 ± 4.9	0.003	13.8 ± 6.7	19.7 ± 4.6	0.002
L. PT	12.9 ± 22.1	2.6 ± 1.5	0.05	4.2 ± 1.4	3.8 ± 1.4	0.34	3.8 ± 1.7	3.7 ± 1.4	0.90
L. Triceps	16.6 ± 26.0	3.6 ± 1.8	0.04	20.1 ± 4.9	24.8 ± 6.3	0.02	17.3 ± 7.4	23.7 ± 6.3	0.01
L. Calf	18.1 ± 26.5	3.8 ± 1.7	0.02	46.0 ± 11.6	58.6 ± 13.7	0.008	38.7 ± 16.8	56.3 ± 13.1	0.002
Both calves	16.2 ± 25.2	3.2 ± 1.6	0.03	92.8 ± 24.0	117.4 ± 27.2	0.006	79.7 ± 33.7	113.6 ± 26.1	0.002

Mean combined bilateral calf fat fraction. CSA=cross-sectional area, RMA=remaining muscle area, L.=left, R.=right. p value is for Mann-Whitney U test or independent samples t-test (as appropriate) between CMT1A and control. TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, TP=tibialis posterior. Rows are highlighted in blue where p value is <0.05

3.4.2.1.2 Cross-sectional area and remaining muscle area

In the CMT1A group at baseline, combined bilateral calf CSA was 92.8 ± 24.0cm^2 compared with 117.4 ± 27.2cm^2 in the control group (p=0.006). In all summary calf measures (left and right triceps surae, left and right anterior compartment and left and right whole calf), there was a statistically smaller baseline CSA in CMT1A compared with controls, though the difference did not reach statistical significance in several individual muscles (right and left peroneus longus, right lateral head of gastrocnemius, right and left medial head of gastrocnemius and left posterior tibial group). Of note, both right and left tibialis posterior muscles had larger baseline CSA

in the CMT1A group, this difference reaching statistical significance in the right posterior tibial group (CMT1A: 4.7 ± 1.2cm^2, control: 3.8 ± 1.3cm^2. p=0.04). Also of note is that in both right and left peroneus longus muscles, although the most fat infiltrated muscle on both sides, CSA was not significantly different to that of controls. Baseline bilateral calf RMA was 79.7 ± 33.7cm^2 in CMT1A and 113.6 ± 26.1 cm^2 in controls (p=0.002). In all summary calf RMA measures and most individual muscles, there were significant differences in RMA between CMT1A and controls, with smaller p values than those seen for matching CSA. Baseline CSA and RMA values in CMT1A and controls are summarised in Table 3-3.

3.4.2.1.3 Mercuri grading score

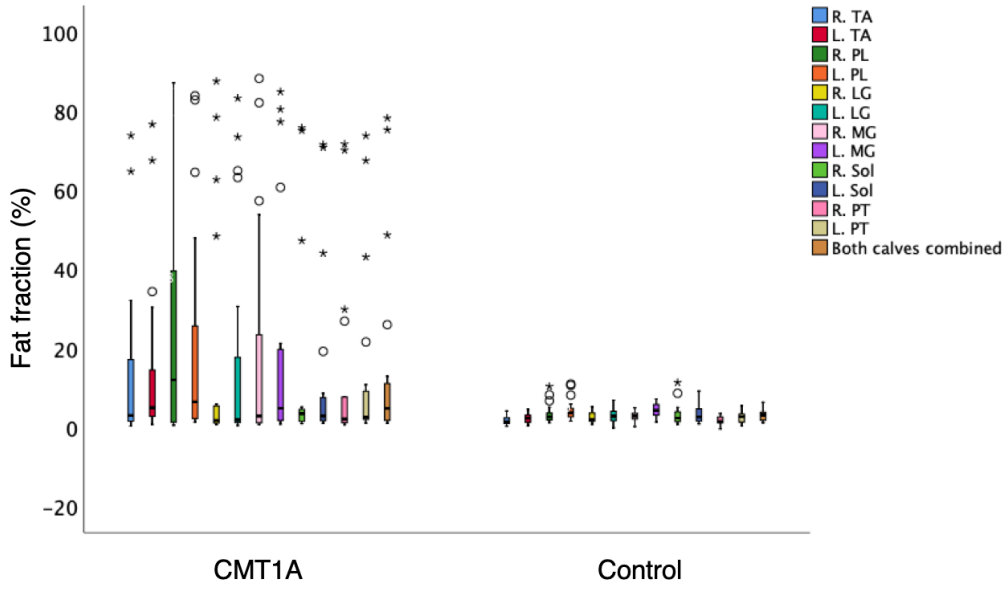
Of the 240 muscles assessed in CMT1A patients, 124 (51.7%) were grades 1-3, 95 (39.6%) were grade 0 and 21 (8.8%) grade 4. Soleus was never grade 4 – consistent with his muscle having the lowest mean FF by qMRI on the left (12.6 ± 22.3%) and second lowest on the right (12.2 ± 23.8%), but other muscles were given the spread of Mercuri grades. When comparing the anterior compartment and triceps surae groups, in the anterior group: 20% of muscles were grade 0 whereas in the posterior groups 48.3% were grade 0. In the anterior group, 62.5% of muscles were grade 1-2b, whereas this accounted for 30.8% in the triceps surae group.

In contrast, of the 240 muscles assessed in controls, 193 muscles (80.4%) were grade 0, 45 muscles (18.8%) grade 1 and two muscles (0.8%) grade 2a – both of the latter soleus muscles of the same participant.

3.4.2.2 Distribution of changes based on measures of fatty accumulation

The distribution of fat accumulation in CMT1A patients and controls at baseline is shown in Figure 3-1. Results for the right and left leg are shown side by side alongside both combined calf measures. Peroneus longus and medial gastrocnemius were most affected by fatty infiltration, and tibialis posterior least affected. The relative sparing of tibialis posterior is again noted. The distribution of changes is in keeping with FF distribution reported in previous publications.

Figure 3-1 – Baseline individual mean muscle fat fraction in CMT1A and controls



Boxes represent median and IQR, whiskers show range, and circles/asterisks are outliers. Difference between baseline mean fat fraction in CMT1A and control groups are significant ($p<0.05$) for all individual muscles except right and left soleus, right lateral gastrocnemius and left posterior tibial group. R.=right, L.=left, TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, TP=tibialis posterior

3.4.2.3 Measures of water accumulation

3.4.2.3.1 T2 relaxation time and magnetisation transfer ratio

In the CMT1A group at baseline, combined bilateral calf T2 relaxation time was significantly longer than in controls (66.5 ± 32.6ms vs 44.8 ± 3.6ms, p = 0.007). Furthermore, T2 time was significantly longer than controls in nearly all individual muscles as well as all summary measures.

At baseline, magnetisation transfer ratio (MTR) was significantly lower in CMT1A patients than controls: combined bilateral calf MTR: 26.1 ± 9.0p.u in CMT1A and 32.0 ± 0.9 p.u. in controls (p=0.001). Of note, there was no significant difference in MTR between CMT1A and control in the tibialis posterior muscles. Baseline T2 relaxation time and MTR for all individual muscles, and muscles groups in CMT1A and controls is summarised in Table 3-4.

Table 3-4 – Baseline T2 relaxation time and magnetisation transfer ratio in CMT1A and controls

	T2 (ms)			MTR (p.u.)		
Muscle	**CMT1A**	**Control**	**p**	**CMT1A**	**Control**	**p**
R. TA	62.3 ± 31.3	41.5 ± 3.0	<0.0001	24.5 ± 7.9	31.7 ± 1.5	<0.0001
R. PL	73.3 ± 40.5	46.4 ± 5.3	0.04	24.1 ± 10.8	31.7 ± 1.4	0.02
R. LG	66.8 ± 35.0	44.7 ± 3.1	0.007	25.9 ± 10.2	31.9 ± 0.9	0.001
R. MG	69.5 ± 35.0	44.8 ± 4.5	0.04	25.6 ± 10.1	32.7 ± 1.3	0.01
R. Sol	62.5 ± 33.4	47.9 ± 7.6	0.04	27.8 ± 9.2	32.0 ± 1.6	0.06
R. PT	60.9 ± 36.8	43.7 ± 3.4	0.04	27.8 ± 9.6	32.7 ± 1.4	0.10
R. Calf	65.7 ± 33.0	44.8 ± 3.8	0.007	26.0 ± 8.9	32.1 ± 0.9	0.001
L. TA	70.0 ± 38.7	43.5 ± 3.1	<0.0001	25.6 ± 9.5	31.8 ± 1.4	<0.0001
L. PL	65.9 ± 33.7	45.6 ± 4.2	0.05	26.2 ± 9.8	32.0 ± 1.4	1.0
L. LG	69.5 ± 45.0	44.3 ± 4.6	0.26	25.7 ± 11.0	31.9 ± 1.7	0.06
L. MG	70.5 ± 38.0	44.5 ± 4.9	0.01	25.8 ± 10.8	32.3 ± 1.4	0.048
L. Sol	61.3 ± 33.4	46.7 ± 5.2	0.03	28.0 ± 8.7	32.0 ± 1.2	0.04
L. PT	66.1 ± 33.4	44.6 ± 2.9	0.001	27.1 ± 8.7	31.5 ± 2.4	0.07
L. Calf	67.2 ± 32.4	44.9 ± 3.5	0.002	26.4 ± 9.1	31.9 ± 1.4	0.003
Both Calves	66.5 ± 32.6	44.8 ± 3.6	0.003	26.1 ± 9.0	32.0 ± 0.9	0.001

T2=T2 relaxation time, ms=milliseconds, MTR=magnetisation transfer ratio, p.u.=percentage Units, R.=right, L.=left, TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior. P value for Mann-Whitney U test. Rows highlighted in blue where p<0.05

3.4.2.3.2 STIR

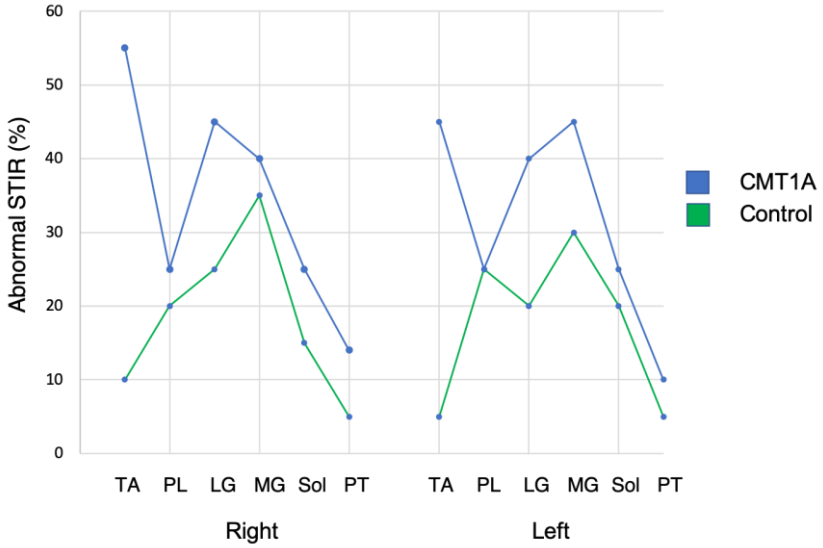
The pattern of STIR changes in CMT1A revealed highest frequency of STIR abnormality in the tibialis anterior muscle in which 55% of patients had STIR abnormality on the right and 45% on the left, and least in tibialis posterior: 15% and 10% for right and left calf respectively. STIR abnormality in peroneus longus was low (30% of muscles on right, 25% on left) despite it being the most fatty-infiltrated muscle in the CMT1A group. There was no significant difference in baseline combined bilateral calf STIR grading between CMT1A and controls: mean 0.4 ± 0.3 versus 0.3 ± 0.3 respectively (p=0.13), however when muscles were analysed individually, there was a statistically significant difference between groups in tibialis anterior: right=0.7 ± 0.7 in CMT1 versus 0.1 ± 0.3 in controls (p=0.01) and left=0.5 ± 0.6 in CMT1A and 0.1 ± 0.2 in controls (p=0.03) though not in other individual muscles. Table 3-5 summarises the frequency of each STIR grading in CMT1A and controls at baseline. Figure 3-2 demonstrates the distribution of STIR abnormality within individual muscles in CMT1A and control.

Table 3-5 – Baseline STIR abnormalities in CMT1A and controls

	STIR grading: n (%)		
	Nil	**Mild**	**Marked**
CMT1A	162 (67.5%)	60 (25%)	18 (7.5%)
Control	197 (82.0%)	33 (13.8%)	10 (4.2%)

Values are count (n) and percent within that group. CMT1A=Charcot-Marie-Tooth disease type 1A

Figure 3-2 – Baseline STIR distribution in CMT1A and controls



Values are the proportion of participants with STIR abnormality in each individual muscle

3.4.3 Myometric data

Myometric values were significantly lower at baseline in all ankle movements of the CMT1A cohort except for isometric right and left ankle inversion at 10 degrees. Ankle dorsiflexion, was relatively speaking weaker when compared with controls than was ankle plantarflexion, and eversion weaker than inversion – as expected from the clinical phenotype (Reilly et al., 2010). Baseline myometry results in CMT1A and controls are summarised in Table 3-6.

Table 3-6 – Baseline myometry in CMT1A and controls

			CMT1A	Control	
Ankle action	**Type**	**Side/angle**	**torque (nM)**		**p**
Plantarflexion	**Isometric**	**R. 10°**	33.7 ± 19.6	68.5 ± 21.6	<0.0001
		L. 10°	32.7 ± 19.8	65.1 ± 17.7	<0.0001
	Isokinetic	**R. 60°/s**	18.7 ± 12.8	58.4 ± 22.5	<0.0001
		L. 60°/s	19.0 ± 12.4	58.4 ± 21.6	<0.0001
Dorsiflexion	**Isometric**	**R. 10°**	9.7 ± 8.1	36.1 ± 13.9	<0.0001
		L. 10°	10.0 ± 7.8	35.3 ± 13.0	<0.0001
	Isokinetic	**R. 60°/s**	11.9 ± 8.4	25.2 ± 8.8	<0.0001
		L. 60°/s	11.9 ± 11.0	25.0 ± 10.0	0.0003
Inversion	**Isometric**	**R. 10°**	15.6 ± 10.0	19.2 ± 6.5	0.20
		L. 10°	14.8 ± 10.4	19.6 ± 7.3	0.11
	Isokinetic	**R. 60°/s**	16.0 ± 9.9	24.5 ± 8.2	0.008
		L. 60°/s	14.3 ± 9.5	24.4 ± 9.2	0.002
Eversion	**Isometric**	**R. 10°**	7.1 ± 4.1	21.4 ± 7.9	<0.0001
		L. 10°	7.7 ± 4.8	20.2 ± 7.7	<0.0001
	Isokinetic	**R. 60°/s**	8.4 ± 4.0	16.3 ± 5.8	<0.0001
		L. 60°/s	7.8 ± 3.1	16.8 ± 6.1	<0.0001

Values are mean torque in newton metres ± standard deviation. R=right, L=left. p value is for independent samples two-tailed t-test. Rows are highlighted in blue where p <0.05

3.4.4 Clinical data

Baseline clinical data is summarised in Table 3-13 and Table 3-16. There was a wide spread of ages (mean 42.8y ± 13.9y, range 21y – 81y) and clinical severity: CMT examination score (CMTES) (range 0-18/28) in the CMT1A cohort.

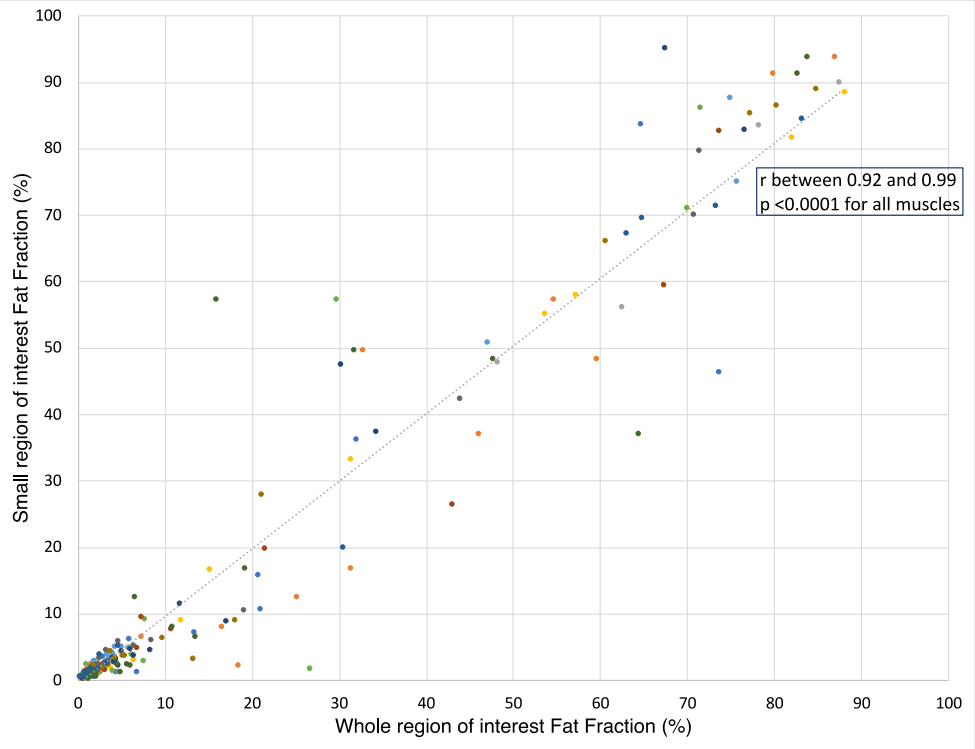
3.4.5 Correlation of baseline data

3.4.5.1 Between MRI measures

3.4.5.1.1 Whole versus small region fat fraction

As may be expected, there were statistically significant correlations between each individual muscle's small and large region of interest. The correlations were highly significant ranging between 0.92 and 0.99 ($p<0.0001$). This is of some interest, given the longer time taken for manual segmentation with large regions of interest, however given the higher standard deviations due to poor reliability, use of small regions of interest for FF assessment is limited, particularly longitudinally. Correlation between small and large regions of interest is shown in Figure 3-3.

Figure 3-3 – Baseline correlation between small and whole muscle regions of interest in CMT1A



Each point represents whole muscle region of interest matched with small region of interest in the same muscle. p is for Spearman's rho. All twelve calf muscle analysed are presented here

3.4.5.1.2 MRI determined fat fraction and Mercuri grading

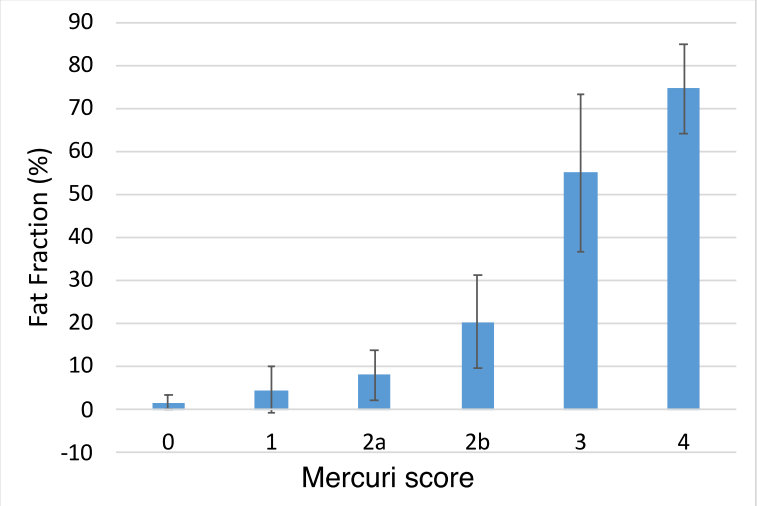
There were strong correlations between baseline Mercuri score and FF in CMT1A patients – a demonstration of convergent validity. All correlations are given in Table 3-7. It should be noted that there was significant overlap in standard deviations between adjacent Mercuri scores (Figure 3-4), particularly for grades 0 to 2b, highlighting the relatively poor sensitivity of the semi-quantitative grading scale to small changes in FF.

Table 3-7 – Correlation between baseline Mercuri grading and fat fraction

Muscle	rho	p
R. TA	0.850	<0.0001
R. PL	0.934	<0.0001
R. LG	0.868	<0.0001
R. MG	0.892	<0.0001
R. Sol	0.938	<0.0001
R. PT	0.754	<0.0001
L. TA	0.838	<0.0001
L. PL	0.795	<0.0001
L. LG	0.760	<0.0001
L. MG	0.794	<0.0001
L. Sol	0.874	<0.0001
L. PT	0.728	<0.0001
Both calves	0.849	<0.0001

Values are Spearman's rho for the stated muscle baseline fat fraction and Mercuri grade. R.=right, L.=left, TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, TP=tibialis posterior

Figure 3-4 – Baseline mean fat fraction for each Mercuri score in CMT1A



Values are mean fat fraction (%) with error bars representing one standard deviation

3.4.5.1.3 MRI fat fraction with cross-sectional area and remaining muscle area

Correlations between FF, CSA and RMA are shown in Table 3-8. Of note, there were no strong correlations between baseline FF and CSA in either group – except weak negative correlation in left medial gastrocnemius (r_s-=0.464, p=0.04) in CMT1A.

Conversely, there were strong correlations however between baseline FF and RMA in the CMT1A group. This may be expected given that the metric RMA appears to be a more sensitive marker of chronic fat atrophy. There were no correlations between

FF and RMA in controls. CSA and RMA correlated together strongly in both groups as expected.

Table 3-8 – Baseline correlation between fat fraction, cross-sectional area and remaining muscle area in CMT1A and controls

	CMT1A			Control		
	FF/CSA	FF/RMA	CSA/RMA	FF/CSA	FF/RMA	CSA/RMA
R.TA	-0.131, p=0.6	-0.697, p=0.001	0.790, p<0.0001	-0.10, p=1.0	-0.32, p=0.9	0.789, p<0.0001
R. PL	-0.388, p=0.09	-0.760, p<0.0001	0.881, p<0.0001	0.068, p=0.8	-0.05, p=1.0	0.997, p <0.0001
R. LG	-0.299, p=0.2	-0.675, p=0.001	0.895, p<0.0001	0.171, p=0.5	0.128, p=0.6	0.999, p<0.0001
R.MG	-0.386, p=0.9	-0.789, p<0.0001	0.841, p<0.0001	0.224, p=0.4	0.180, p=0.5	0.999, p<0.0001
R. Sol	-0.269, p=0.3	-0.623, p=0.003	0.910, p<0.0001	0.092, p=0.7	-0.20, p=0.9	0.993, p<0.0001
R. PT	0.047, p=0.8	0.725, p<0.0001	0.632, p=0.003	0.063, p=0.8	0.32, p=0.9	0.999, p<0.0001
R. triceps	-0.339, p=0.1	-0.794, p<0.0001	0.822, p<0.0001	-0.19, p=0.9	0.197, p=0.7	0.510, p=0.026
R. calf	-0.302, p=0.2	-0.800, p<0.0001	0.800, p<0.0001	0.216, p=0.4	0.151, p=0.5	0.998, p<0.0001
L. TA	0.298, p=0.2	-0.696, p=0.001	0.890, p<0.0001	-0.164, p=0.5	-0.204, p-0.4	0.999, p<0.0001
L.PL	-0.201, p=0.4	-0.740, p<0.0001	0.788, p<0.0001	-0.088, p=0.7	-0.185, p=0.5	0.995, p<0.0001
L.LG	-0.84, p=0.7	-0.782, p<0.0001	0.662, p=0.001	0.007, p=1.0	-0.046, p=0.9	0.999, p<0.0001
L.MG	0.464, p=0.04	-0.857, p<0.0001	0.818, p<0.0001	-0.187, p=0.4	-0.236, p=0.331	0.999, p<0.0001
L.Sol	-0.399, p=0.08	-0.703, p=0.001	0.925, p<0.0001	0.363, p=0.1	0.266, p=0.3	0.995, p<0.0001
L.PT	-0.196, p=0.4	-0.646, -=0.002	0.857, p<0.0001	-0.121, p=0.6	-0.163, p=0.5	0.999, p<0.0001
L.triceps	-0.458, p=0.04	-0.884, p<0.0001	0.803, p<0.0001	-0.134, p=0.6	0.020, p=0.9	0.616, p=0.005
L. calf	-0.373, p=0.1	-0.857, p<<0.0001	0.786, p<0.0001	0.149, p=0.5	0.069, p=0.8	0.997, p<0.0001
Both calves	-0.338, p=0.1	-0.831, p<0.0001	0.792, p<0.0001	0.190, p=0.4	0.119, p=0.6	0.997, p<0.0001

Values are Spearman's rho, p value for r_s. FF=whole region fat fraction, RMA=remaining muscles area, CSA=cross-sectional area, CMT1A=Charcot-Marie-Tooth disease type 1A, R.=right, L.=left, TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, TP=tibialis posterior, triceps=triceps surae group

3.4.5.1.4 Overall MRI determined measures: fat fraction, T2 and MTR

There were strong correlations between all qMRI parameters. As expected, T2 relaxation time increased and MTR decreased as FF increased. Correlations for combined bilateral calf are summarised in Table 3-9. A bivariate linear regression was conducted to examine how well FF could predict T2 time and MTR. A scatterplot showed that the relationship between FF and T2 was positive and linear, and that between FF and MTR was negative and linear. The regression equation for predicting T2 time from FF was:

Equation 3-1 – Regression equation for predicting T2 from fat fraction at baseline in CMT1A

$$\hat{Y} = 47.06 + 1.254x$$

The r^2 for this equation was 0.948, indicating 94.8% of the variance in T2 time was predictable by FF – a strong relationship. The bootstrapped 95% confidence interval for the slope to predict T2 time from FF range from 1.110 to 1.399, thus for every 1% increase in FF, T2 increases by between 1.11 and 1.4 milliseconds. There were no bivariate outliers in either group.

For MTR, the regression equation was:

Equation 3-2 – Regression equation for predicting MTR from fat fraction at baseline in CMT1A

$$\hat{Y} = 31.609 - 0.351x$$

The r^2 for this equation was 0.979 indicating a strong relationship. The bootstrapped 95% confidence interval for the slope to predict MTR from FF range from -0.376 to -0.325, thus for every 1% increase in FF, MTR decreases by between 0.376 and 0.325 p.u.

Table 3-9 – Correlation matrix for baseline whole calf MRI measures

	Fat fraction	T2 time	MTR
Fat fraction		0.974, p<0.0001	-0.989, p<0.0001
T2 time	0.974, p<0.0001		-0.987, p<.0001
MTR	-0.989, p<0.0001	-0.987, p<.0001	

MTR=magnetisation transfer ratio. T2 time=T2 relaxation time

3.4.5.2 MRI determined fat fraction and clinical measures

In CMT1A patients at baseline, age and duration of disease were both highly positively correlated with all baseline FF values and with summary clinical measures except for SF-36.

There were statistically significant correlations in the CMT1A group between overall clinical measures and MRI summary calf FF measures, and between specific myometric and relevant qMRI FF measures

CMT symptom score (CMTSS), CMTES and CMTES-lower limb components (CMTES-LL) all correlated strongly with all individual and combined baseline FF measures (Figure 3-5). Overall SF-36 did not correlate with baseline FF however SF36-PF correlated with baseline combined bilateral calf FF and some individual calf muscle FF, which is in keeping with the site of pathology. There were high negative correlations between MRC scores and FF. In particular correlation between MRC total score and combined bilateral calf FF was -0.726 (p<0.0001) and between MRC-LL and combined bilateral calf FF was -0.748 (p<0.0001). Summary of baseline MRI-clinical correlations is presented in Table 3-11.

Figure 3-5 – Baseline CMT neuropathy score subsets correlate with whole calf fat fraction in CMT1A

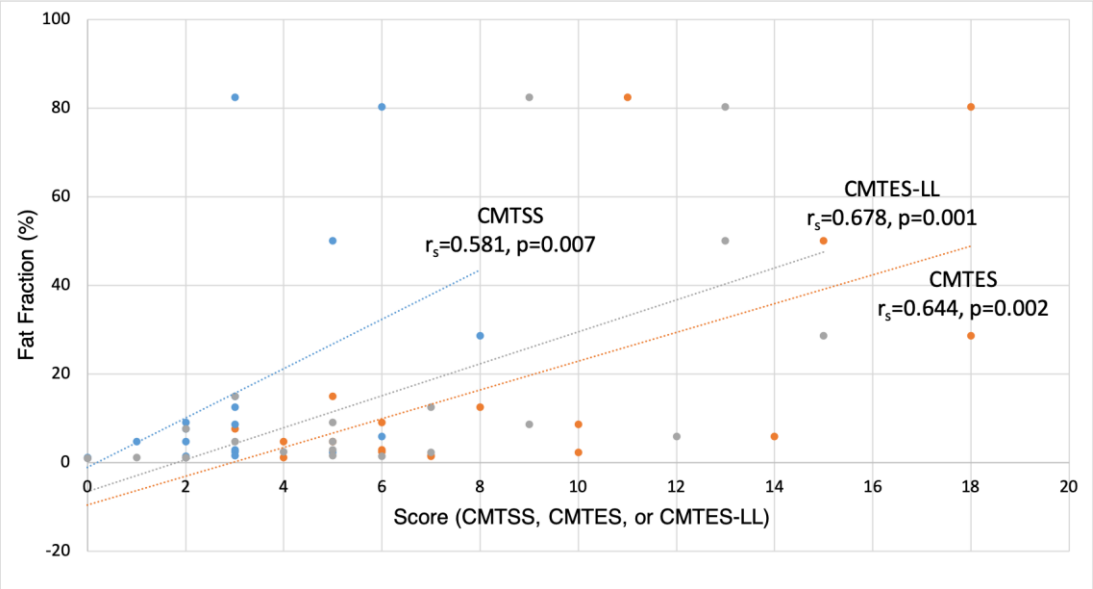


Figure is colour coded for CMTNS subscore (points): orange=CMTES, grey=CMTES-LL, blue=CMTSS. r_S=Spearman rho and associated p value. CMT1A=Charcot-Marie-Tooth disease type 1A, CMTNS=CMT neuropathy score, CMTES=CMT examination score, CMTES-LL=CMTES lower limb components

3.4.5.3 MRI determined measures and myometry

There were strong correlations between baseline FF and myometry in CMT1A patients but not in controls. Correlations were stronger for isometric measures – particularly in plantarflexion – than for isokinetic measures. There were strong correlations in both control and CMT1A between summary myometry measures and CSA and RMA as may be expected. Correlation with RMA was stronger than with CSA across the board. Summary of MRI measure-myometry correlations are given in Table 3-10 and Table 3-12.

Table 3-10 – Baseline correlation between summary myometry, cross-sectional area and remaining muscle area in CMT1A and controls

		CMT1A		Control	
Ankle movement	**Side**	**Triceps CSA**	**Triceps RMA**	**Triceps CSA**	**Triceps RMA**
Plantarflexion	**Right**	0.586, p=0.007	0.726, p<0.0001	0.577, p=0.001	0.804, p<0.0001
	Left	0.640, p=0.002	0.806, p<0.0001	0.453, p=0.008	0.714, p=0.001
		Anterior CSA	**Anterior RMA**	**Anterior CSA**	**Anterior RMA**
Dorsiflexion	**Right**	0.576, p=0.008	0.753, p<0.0001	0.850, p<0.0001	0.834, p<0.0001
	Left	0.548, p=0.012	0.711, p<0.0001	0.598, p=0.007	0.857, p<0.0001

Values and p values are for Pearson correlation coefficient. RMA=remaining muscle area, CSA=cross-sectional area

Table 3-11 – Baseline correlation matrix between summary calf fat fraction and clinical measures in CMT1A

	SF36	SF36-PF	CMTSS	CMTES	CMTES-LL	MRC total	MRC-LL	Age	Duration
Right Calf	-0.236, p=0.35	-0.666, p=0.003	0.578, p=0.008	0.634, p=0.003	0.665, p=0.001	-0.768, p=0.001	-0.777, p<0.0001	0.844, p<0.0001	0.889, p<0.0001
Left Calf	-0.162, p=0.52	-0.678, p=0.002	0.566, p=0.01	0.642, p=0.002	0.679, p=0.001	-0.720, p=0.001	-0.724, p<0.0001	0.823, p<0.0001	0.860, p<0.0001
Both Calves	-0.183, p=0.47	-0.674, p=0.002	0.581, p=0.007	0.644, p=0.002	0.678, p=0.001	-0.726, p<0.0001	-0.748, p<0.0001	0.836, p<0.0001	0.883, p<0.0001
SF36		0.503, p=0.03	-0.412, p=0.09	-0.150, p=0.55	-0.161, p=0.52	0.367, p=0.13	0.267 p=0.28	-0.344, p=0.16	-0.432, p=0.07
SF36-PF			-0.838, p<0.0001	-0.648, p<0.004	-0.650, p<0.004	0.748, p<0.0001	0.647, p=0.004	-0.610, p=0.007	-0.675, p=0.02
CMTSS				0.895, p<0.0001	0.843, p<0.0001	-0.761, p<0.0001	-0.725, p<0.0001	0.485, p=0.03	0.653, p=0.002
CMTES					0.981, p<0.0001	-0.696, p<0.001	-0.678, p<0.001	0.426, p=0.06	0.577, p=0.008
CMTES-LL						-0.695, p=0.001	-0.680, p=0.001	0.436, p=0.05	0.563, p=0.01
MRC total							0.945, p<0.0001	-0.629, p=0.003	-0.710, p<0.0001
MRC-LL								-0.598, p=0.005	-0.680, p=0.001
Age									0.935, p<0.0001
Duration									

Values are Pearson rho correlation coefficient and p value. R.=right, L.=left, TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, TP=tibialis posterior. MRC=Medical Research Council strength assessment, SF36=Short Form Health Survey 36, PF=physical functioning domain, CMTES=CMT Examination Score, CMTSS=CMT symptom score. Cell highlighted in blue where correlation was significant

Table 3-12 – Correlation between baseline myometry and fat fraction in CMT1A

			CMT1A				
Ankle action	**Type**	**Side/angle**	**Both calves**	**Right calf**	**Left calf**	**Right triceps**	**Left triceps**
Plantarflexion	**Isometric**	**R. 10°**	-0.604, p=0.005	-0.597, p=0.005		-0.608, p=0.003	
		L. 10°	-0.649, p=0.002		-0.660, p=0.002		-0.730, p<0.0001
	Isokinetic	**R. 60°/s**	-0.416, p=0.07	-0.411, p=0.07		-0.399, p=0.08	
		L. 60°/s	-0.450, p=0.046		-0.457, p=0.043		-0.514, p=0.02
						Right anterior	**Left anterior**
Dorsiflexion	**Isometric**	**R. 10°**	-0.519, p=0.02	-0.517, p=0.02		-0.546, p=0.01	
		L. 10°	-0.583, p=0.007		-0.585, p=0.007		-0.619, p=0.004
	Isokinetic	**R. 60°/s**	-0.655, p=0.002	-0.655, p=0.002		-0.685, p=0.001	
		L. 60°/s	-0.537, p=0.02		-0.543, p=0.01		-0.567, p=0.009

			Control				
Ankle action	**Type**	**Side/angle**	**Both calves**	**Right calf**	**Left calf**	**Right triceps**	**Left triceps**
Plantarflexion	**Isometric**	**R. 10°**	-0.266, p=0.3	-0.286, p=0.2		-0.243, p=0.3	
		L. 10°	-0.244, p=0.3		-0.225, p=0.4		-0.221, p=0.4
	Isokinetic	**R. 60°/s**	-0.330, p=-0.2	-0.339, p=0.2		-0.320, p=0.2	
		L. 60°/s	-0.192, p=0.4		-0.222, p=0.4		-0.179, p=0.5
						Right anterior	**Left anterior**
Dorsiflexion	**Isometric**	**R. 10°**	0.034, p=0.90	0.013, p=0.96		-0.231, p=0.3	
		L. 10°	0.096, p=0.70		0.081, p=0.7		-0.316, p=0.2
	Isokinetic	**R. 60°/s**	0.163, p=0.5	0.076, p=0.76		-0.207, p=0.4	
		L. 60°/s	0.137, p=0.58		0.171, p=0.5		-0.077, p=0.8

Values and p values are for Pearson or Spearman correlation coefficient as appropriate

3.5 Longitudinal results

3.5.1 Study participants

Of the 40 participants assessed at baseline, 17 CMT1A patients and 20 controls were reassessed at a mean interval ± s.d. of 12.8 ± 1.1 months (visit two). 14 CMT1A patients and 11 controls were reassessed at a mean interval of 28.7 ± 2.5 months (visit three), and finally 10 CMT1A and 8 controls at a mean interval of 47.9 ± 6.3 months (visit four). Participant attrition was attributable to death, illness, inability to travel and self-withdrawal from the study for personal reasons. Controls withdrew either due to personal reasons, or because their patient relative/friend had withdrawn.

There were no significant differences in age, gender, height or weight between the two groups at any visit (Table 3-13).

Table 3-13 – Participant demographic details at each visit

	Baseline			Visit two		
	CMT1A	Controls	p	CMT1A	Controls	p
Gender	M=11, F=9	M=12, F=8	0.76	M=10, F=7	M=12, F=8	0.94
Age	42.8 ± 13.9	52.1 ± 16.1	0.05	44.8 ± 14.8	53.1 ± 16.1	0.12
Weight	70.4 ± 15.7	73.3 ± 16.8	0.24	ND	ND	ND
Height	167.2 ± 11.8	171.1 ± 10.1	0.55	ND	ND	ND
	Visit three			Visit four		
	CMT1A	Controls	p	CMT1A	Controls	p
Gender	M=8, F=6	M=7, F=4	0.75	M=5, F=5	M=5, F=3	0.62
Age	45.6 ± 15.5	58.0 ± 15.3	0.06	49.1 ± 15.7	61.8 ± 9.8	0.06
Weight	ND	ND	ND	ND	ND	ND
Height	ND	ND	ND	ND	ND	ND

Values are mean ± standard deviation. M=male, F=female, ND=no data available. p value for χ^2 test for gender, and otherwise by independent t-test

3.5.2 Raw longitudinal data

3.5.2.1 MRI measures

In the CMT1A group, combined bilateral mean calf FF ± s.d. was: 16.2 ± 25.2% at baseline, with subsequent increase to 18.1 ± 26.4%, 20.4 ± 28.6% and 27.1 ± 32.4% at visits two, three and four respectively. This compared with a combined bilateral mean calf FF of 3.1 ± 1.5% at baseline in controls, and 3.2 ± 1.4%, 2.9 ± 1.8% and 2.8 ± 0.9% at subsequent visits. As expected, mean FF of all individual calf muscles increased over the study period, as too did mean FF in summary calf FF measures: right and left anterior compartment, right and left triceps surae, right and left calf. There was no change in FF in the control group. Table 3-14, Table 3-15 and Figure 3-6 summarise raw longitudinal qMRI results in CMT1A and controls at baseline and averaged to each follow-up visit.

Table 3-14 – Mean quantitative MRI measures averaged to each visit in CMT1A patients and controls

	CMT1A			
	Baseline	**Visit two**	**Visit three**	**Visit four**
FF small (%)	15.5 ± 25.3	19.2 ± 29.0	21.7 ± 30.6	30.1 ± 33.3
FF whole (%)	16.2 ± 25.2	18.1 ± 26.4	20.4 ± 28.6	27.1 ± 32.4
T2 (ms)	66.6 ± 32.5	69.7 ± 34.1	71.4 ± 35.5	79.1 ± 37.8
MTR (p.u)	26.1 ± 9.0	25.1 ± 10.0	24.6 ± 10.1	25.5 ± 14.7
CSA (cm^2)	92.8 ± 24.0	94.5 ± 25.8	90.2 ± 26.3	83.1 ± 24.4

	Control			
	Baseline	**Visit two**	**Visit three**	**Visit four**
FF small (%)	1.7 ± 1.0	1.7 ± 0.9	1.9 ± 1.3	2.1 ± 0.9
FF whole (%)	3.1 ± 1.5	3.2 ± 1.4	2.9 ± 1.8	2.8 ± 0.9
T2 (ms)	44.8 ± 3.6	44.3 ± 2.9	43.1 ± 6.4	45.9 ± 4.0
MTR (p.u)	32.0 ± 0.90	32.4 ± 1.3	31.7 ± 1.2	35.5 ± 2.5
CSA (cm^2)	117.4 ± 27.3	118.9 ± 27.3	113.1 ± 28.9	107.3 ± 32.3

All values are mean ± standard deviation. FF=fat fraction, small=small region of interest, whole=whole muscle region of interest, T2=transverse T2 relaxation time, MTR=magnetisation transfer ratio, CSA=cross-sectional area, p.u=percentage units, ms=milliseconds

Table 3-15 – Individual and summary mean calf fat fraction averaged to each visit in CMT1A and controls

	Baseline		Visit two		Visit three		Visit four	
	CMT1A	Control	CMT1A	Control	CMT1A	Control	CMT1A	Control
Muscle	Fat fraction (%) ± standard deviation							
R. TA	13.1 ± 21.1	1.8 ± 1.0	16.0 ± 23.2	1.8 ± 0.8	17.8 ± 24.7	1.6 ± 0.9	23.7 ± 31.8	1.6 ± 0.8
R. PL	23.7 ± 28.0	3.4 ± 2.5	27.0 ± 31.4	3.7 ± 2.3	29.6 ± 30.1	3.1 ± 3.2	39.8 ± 35.2	5.8 ± 5.8
R. LG	15.4 ± 28.5	2.6 ± 1.5	19.3 ± 32.1	2.8 ± 1.2	20.0 ± 32.8	2.0 ± 1.5	28.2 ± 38.2	2.0 ± 1.2
R. MG	18.4 ± 28.5	2.8 ± 1.2	22.6 ± 31.1	3.1 ± 1.0	24.4 ± 31.8	2.3 ± 1.2	34.0 ± 36.6	2.0 ± 1.0
R. Sol	12.2 ± 23.8	3.2 ± 2.5	14.7 ± 26.2	3.5 ± 2.7	17.3 ± 28.3	4.1 ± 2.6	24.4 ± 33.8	3.0 ± 4.9
R. PT	12.1 ± 21.7	1.7 ± 1.6	15.2 ± 23.8	1.9 ± 1.0	18.4 ± 27.4	1.7 ± 0.8	23.0 ± 32.4	1.7 ± 0.8
L. TA	14.6 ± 21.7	2.3 ± 1.1	16.8 ± 24.1	2.2 ± 0.8	17.8 ± 26.2	2.0 ± 1.1	23.8 ± 31.7	1.7 ± 0.5
L. PL	20.5 ± 27.2	4.4 ± 2.7	24.9 ± 29.6	4.6 ± 3.0	26.1 ± 30.7	3.8 ± 2.5	33.5 ± 34.5	3.8 ± 2.6
L. LG	17.1 ± 28.7	3.0 ± 1.7	20.8 ± 31.4	4.0 ± 2.2	21.1 ± 32.3	2.7 ± 1.1	29.6 ± 35.4	3.0 ± 2.4
L. MG	19.9 ± 29.5	4.2 ± 1.7	24.3 ± 31.9	4.3 ± 1.8	26.7 ± 32.9	3.7 ± 2.1	34.1 ± 36.2	3.9 ± 2.5
L. Sol	12.6 ± 22.3	3.4 ± 1.6	15.6 ± 25.6	3.4 ± 2.3	18.7 ± 28.8	3.4 ± 2.0	25.9 ± 32.5	3.1 ± 2.0
L. PT	12.9 ± 22.1	2.6 ± 1.5	16.9 ± 25.9	2.5 ± 1.2	18.4 ± 28.4	2.1 ± 1.2	25.1 ± 33.6	2.2 ± 1.0
R. CALF	14.5 ± 24.0	2.7 ± 1.6	17.6 ± 26.6	3.0 ± 1.5	19.8 ± 28.4	2.7 ± 2.3	26.9 ± 34.2	1.4 ± 1.4
L. CALF	18.1 ± 26.5	3.8 ± 1.7	18.6 ± 26.1	3.5 ± 1.4	20.9 ± 28.7	3.0 ± 1.4	27.7 ± 32.4	3.0 ± 1.1
BOTH CALVES	16.2 ± 25.2	3.2 ± 1.6	18.1 ± 26.6	3.2 ± 1.5	20.4 ± 28.4	2.9 ± 1.8	27.1 ± 34.2	2.8 ± 0.9

R.=right, L.=left, TA tibialis anterior, PL=peroneus longus, LG=lateral head of gastrocnemius, MG=medial head of gastrocnemius, Sol.=soleus, PT=tibialis posterior, calf=combination of all muscles in that leg. p values not given as visit interval inconsistent (see Figure 3-7)

3.5.2.2 Clinical measures

Longitudinal clinical measures are summarised in Table 3-16. Note is made of the general trend of increasing CMTNS subsets, and reducing MRC scores.

Table 3-16 – Summary clinical measures averaged to each visit in CMT1A

	Baseline	Visit two	Visit three	Visit four
CMTSS (0-12)	3.1 ± 2.0	3.3 ± 2.1	2.8 ± 2.1	3.6 ± 2.5
CMTES (0-28)	8.0 ± 5.1	8.1 ± 5.4	8.6 ± 5.8	10.2 ± 6.1
CMTES-LL (0-20)	6.3 ± 4.3	6.6 ± 4.3	5.3 ± 5.0	8.1 ± 5.2
SF36 (0-100%)	73.9 ± 15.2	72.5 ± 21.2	74.2 ± 18.2	83.5 ± 6.3
SF36 PF (0-100%)	65.0 ± 23.0	66.0 ± 23.0	62.0 ± 22.0	66.1 ± 26.7
MRC-LL (0-110)	95.4 ± 15.4	95.9 ± 15.8	91.5 ± 20.6	80.6 ± 20.9
MRC-UL (0-120)	112.4 ± 7.9	112.5 ± 8.7	108.9 ± 15.4	108.0 ± 18.1
MRC-total (0-230)	207.8 ± 21.8	208.5 ± 23.6	200.5 ± 34.4	193.3 ± 36.6

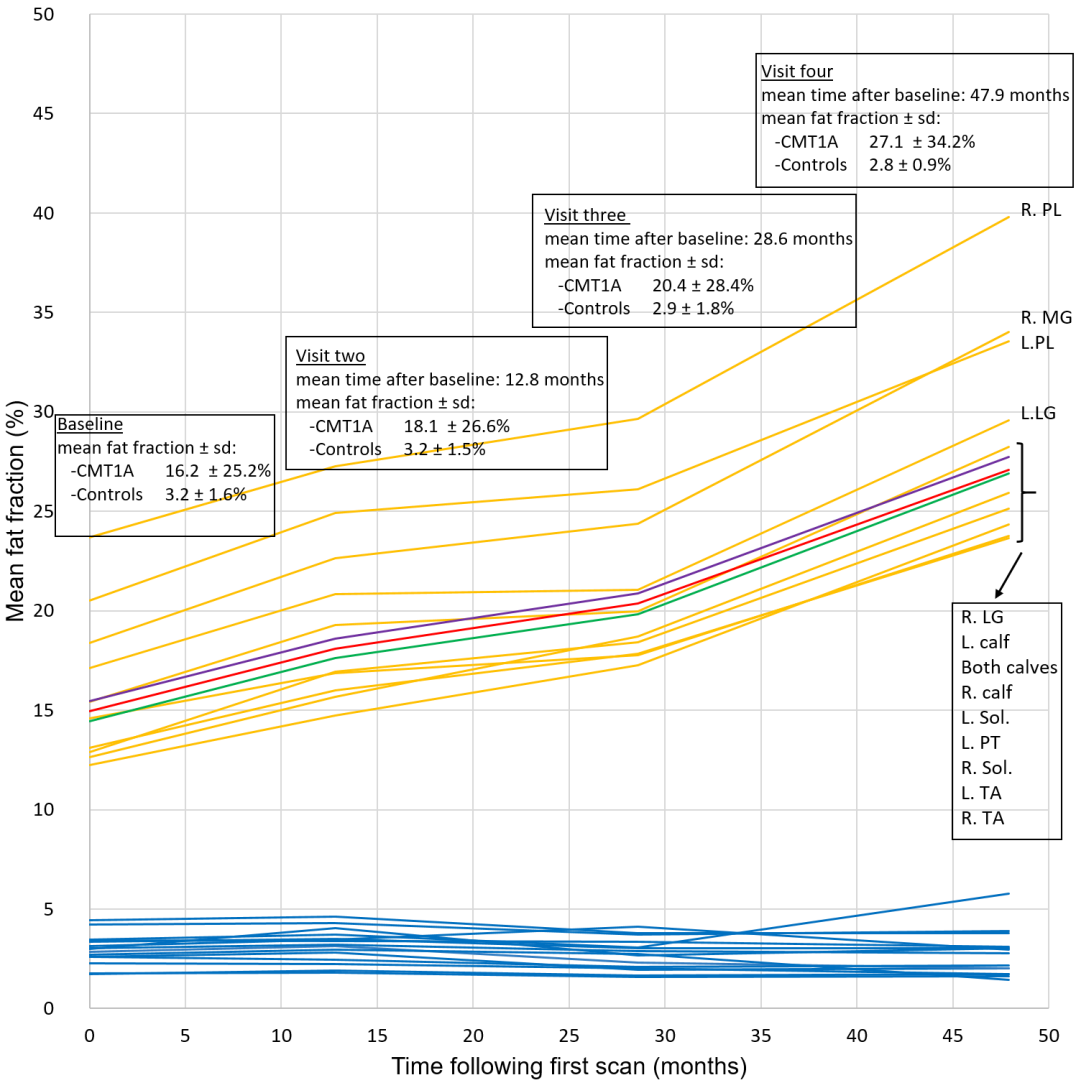
*Values are mean ± standard deviation. *MRC=Medical Research Council strength assessment, SF36=Short Form Health Survey, PF=physical functioning domain, CMTES=CMT examination score, CMTSS=CMT symptom score*

**Muscle strength was assessed using a modified MRC scale (O'Brien, 2000) with 5: Normal strength, 5-: Barely detectable weakness, 4+: Gravity and moderate to maximal resistance, 4: Gravity and moderate resistance, 4-: Gravity and minimal resistance, 3: Full range of motion against gravity only, 2: Movement when gravity is eliminated, 1: Flicker of movement seen or felt, 0: No movement. Upper limb/neck and lower muscle scores were summed to obtain a total upper limb/neck and total lower limb score for each subject. For this purpose, 5- was scored as 4.75, 4+ as 4.25 and 4- as 3.75. The maximum score obtainable was 120 for upper limb/neck and 110 for lower limb*

3.5.2.3 Myometric measures

Table 3-17 summarises longitudinal ankle myometric measures over the four visits in CMT1A and controls. At the time of analysis, it was noted that ankle plantarflexion values in controls at visit four were markedly reduced compared with other visits, however detailed checking of the dynamometer revealed no technical fault, and thus analysis was performed based on these values. Values for CMT1A patients appear unaffected.

Figure 3-6 – qMRI determined mean calf fat fraction in CMT1A and controls at each visit



Each line represents mean fat fraction across the group in a single muscle. Figure is colour coded for visual aid: yellow=CMT1A patients, blue=controls, red=combined bilateral calf CMT1A, purple=left calf CMT1A, green=right calf CMT1A. L.=left, R.=right, PL=peroneus longus, MG=medial head of gastrocnemius, LG=lateral head of gastrocnemius, Sol.=soleus, TA tibialis anterior, sd=standard deviation

Table 3-17 – Ankle myometry values averaged to each visit in CMT1A and controls

			CMT1A				Control			
Ankle action	**Type**	**Side/angle**	**Baseline**	**Visit two**	**Visit three**	**Visit four**	**Baseline**	**Visit two**	**Visit three**	**Visit four**
Plantarflexion	**Isometric**	**R. 10°**	33.7 ± 19.6	39.9 ± 26.4	33.8 ± 21.6	28.3 ± 24.7	68.5 ± 21.6	70.5 ± 29.0	58.9 ± 29.0	47.6 ± 15.9
		L. 10°	32.7 ± 19.8	38.6 ± 24.6	34.5 ± 24.8	21.1 ± 19.3	65.1 ± 17.7	64.6 ± 17.7	60.5 ± 22.8	45.4 ± 23.4
	Isokinetic	**R. 60°/s**	18.7 ± 12.8	23.8 ± 13.5	22.9 ± 15.7	20.2 ± 19.3	58.4 ± 22.5	65.0 ± 30.9	53.0 ± 32.3	49.9 ± 25.4
		L. 60°/s	19.0 ± 12.4	22.5 ± 13.8	20.9 ± 15.4	18.2 ± 22.1	58.4 ± 21.6	63.1 ± 24.0	57.1 ± 22.0	54.3 ± 22.0
Dorsiflexion	**Isometric**	**R. 10°**	9.7 ± 8.1	11.5 ± 8.7	12.5 ± 8.3	11.0 ± 10.3	36.1 ± 13.9	36.9 ± 15.0	34.6 ± 13.0	38.6 ± 18.7
		L. 10°	10.0 ± 7.8	12.7 ± 9.4	12.1 ± 8.5	12.3 ± 11.0	35.3 ± 13.0	37.1 ± 13.6	35.0 ± 11.8	38.6 ± 17.1
	Isokinetic	**R. 60°/s**	11.9 ± 8.4	14.1 ± 7.7	17.0 ± 8.1	12.2 ± 8.4	25.2 ± 8.8	22.4 ± 7.3	23.9 ± 6.3	24.1 ± 8.8
		L. 60°/s	11.9 ± 11.0	15.9 ± 9.8	16.1 ± 5.6	10.6 ± 9.2	25.0 ± 10.0	25.6 ± 7.6	22.8 ± 4.8	24.3 ± 7.7
Inversion	**Isometric**	**R. 10°**	15.6 ± 10.0	18.7 ± 13.3	14.7 ± 12.9	11.8 ± 7.9	19.2 ± 6.5	19.8 ± 7.4	19.1 ± 8.1	19.3 ± 8.5
		L. 10°	14.8 ± 10.4	14.3 ± 9.0	13.2 ± 9.8	10.1 ± 9.2	19.6 ± 7.3	20.8 ± 7.4	20.0 ± 8.8	19.1 ± 9.8
	Isokinetic	**R. 60°/s**	16.0 ± 9.9	19.3 ± 12.1	17.0 ± 14.5	10.8 ± 8.1	24.5 ± 8.2	25.1 ± 8.4	22.7 ± 10.3	23.6 ± 9.2
		L. 60°/s	14.3 ± 9.5	16.1 ± 10.0	14.7 ± 10.1	10.9 ± 9.0	24.4 ± 9.2	26.7 ± 7.9	24.8 ± 10.8	23.5 ± 10.4
Eversion	**Isometric**	**R. 10°**	7.1 ± 4.1	8.7 ± 5.9	7.9 ± 5.8	7.4 ± 4.7	21.4 ± 7.9	22.1 ± 9.3	22.6 ± 14.0	29.3 ± 7.6
		L. 10°	7.7 ± 4.8	9.4 ± 5.5	7.6 ± 5.2	6.7 ± 5.3	20.2 ± 7.7	22.0 ± 8.1	20.0 ± 9.7	20.1 ± 10.7
	Isokinetic	**R. 60°/s**	8.4 ± 4.0	9.3 ± 3.8	9.6 ± 8.2	7.2 ± 5.7	16.3 ± 5.8	17.5 ± 6.9	17.7 ± 7.8	21.5 ± 9.3
		L. 60°/s	7.8 ± 3.1	9.6 ± 4.4	8.9 ± 5.6	6.7 ± 5.4	16.8 ± 6.1	17.9 ± 6.9	16.3 ± 8.2	16.5 ± 7.9

Values are mean torque in newton metres ± standard deviation. Isometric values are the peak torque at the fixed angle, and isokinetic values are the peak torque at the fixed speed noted

3.5.3 Annualised change for accurate longitudinal analysis

Given the non-uniformity of timings of visits three and four (Figure 3-7), longitudinal change in all parameters were converted to an annualised measure such that:

Equation 3-3 – Calculation of annualised change

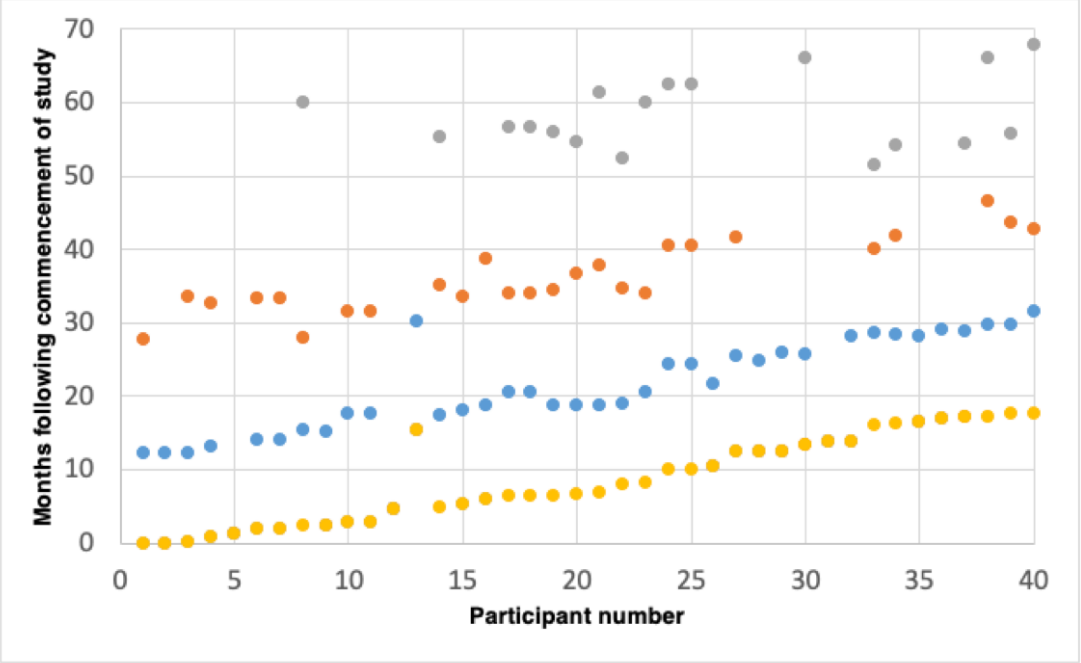
$$annualised\ change\ in\ measure\ (\%\ per\ year) = \frac{raw\ measure}{days\ following\ initial\ assessment}\ x\ 365$$

thus allowing accurate longitudinal statistical analysis at timepoints two, three and four.

Timepoint two refers to annualised data from all second visits, timepoint three to annualised data from all third visits, and timepoint four to annualized data from the final visit of each patient who had ≥ three visits. Longitudinal annualised change in all measured parameters are summarised in Table 3-18 and for MRI determined FF in Table 3-19.

Change in all collected measures for CMT1A patients was assessed for statistically significant difference from baseline with paired t-test or Wilcoxon signed rank test as appropriate and from change in matched controls with student two-tailed t-test or Mann-Whitney U test as appropriate. Longitudinal change in the CMT1A cohort was statistically significant in several of the measured, parameters both qMRI derived and functional.

Figure 3-7 – Study participant assessment timeline



Each point represents a single assessment in a single participant. Figure is colour coded for visit number: visit 1=yellow, visit 2= blue, visit 3=orange, visit 4=grey

Table 3-18 – Summary of annualised longitudinal data in CMT1A and controls

	Timepoint two					Timepoint three					Timepoint four				
	CMT1A	Controls	p1	p2	SRM	CMT1A	Controls	p1	p2	SRM	CMT1A	Controls	p1	p2	SRM
MRI DATA															
FF whole (%)	1.2 ± 1.1	0.2 ± 0.5	<0.001	0.02	1.04	0.9 ± 0.9	-0.0 ± 0.4	0.002	0.006	1.01	0.7 ± 0.6	0.0 ± 0.4	0.002	0.01	1.07
FF small (%)	1.4 ± 2.5	0.0 ± 0.4	0.04	0.02	0.56	0.7 ± 1.3	0.1 ± 0.3	0.06	0.16	0.55	1.0 ± 2.1	0.1 ± 0.3	0.07	0.15	0.51
T2 (ms)	0.5 ± 4.2	-0.6 ± 1.4	0.63	0.30	0.12	0.1 ± 1.4	-0.1 ± 0.8	0.8	0.64	0.07	0.3 ± 1.5	0.1 ± 0.6	0.51	0.83	0.17
MTR (pu)	0.1 ± 1.4	0.1 ± 1.0	0.88	0.90	0.04	-0.2 ± 0.3	0.0 ± 0.4	0.08	0.35	-0.51	0.1 ± 0.5	0.5 ± 0.8	0.55	0.14	0.17
CSA (cm^2)	0.4 ± 3.7	1.1 ± 5.2	0.64	0.67	0.11	-2.0 ± 2.8	0.0 ± 1.7	0.02	0.06	-0.71	-1.2 ± 1.9	-0.3 ± 1.0	0.04	0.20	-0.59
RMA (cm^2)	-0.5 ± 3.8	0.8 ± 5.0	0.59	0.39	-0.13	-2.3 ± 2.6	0.0 ± 1.7	0.006	0.03	-0.87	-2.0 ± 2.2	-0.3 ± 1.2	0.004	0.03	-0.89
CLINICAL DATA															
MRC-UL (points)	0.5 ± 3.2	NA	0.49	NA	0.17	-1.6 ± 7.0	NA	0.29	NA	-0.20	-0.7 ± 3.4	NA	0.45	NA	-0.2
MRC-LL (points)	-0.4 ± 3.8	NA	0.67	NA	-0.10	-1.9 ± 6.3	NA	0.27	NA	-0.31	-1.5 ± 3.5	NA	0.12	NA	-0.43
MRC-overall (points)	0.1 ± 6.6	NA	0.93	NA	0.02	-3.6 ± 13.2	NA	0.3	NA	-0.27	-2.2 ± 6.8	NA	0.23	NA	-0.32
SF36 (points)	-2.4 ±15.0	NA	0.51	NA	-0.16	-0.9 ± 6.2	NA	0.61	NA	-0.14	-0.6 ± 6.0	NA	0.85	NA	-0.10
CMTES (points)	0.3 ± 1.2	NA	0.33	NA	0.24	0.6 ± 0.6	NA	0.003	NA	1	0.5 ± 0.5	NA	0.002	NA	1.12
CMTSS (points)	0.3 ± 0.8	NA	0.16	NA	0.35	0.1 ± 0.3	NA	0.42	NA	0.22	0.1 ± 0.3	NA	0.15	NA	0.41
MYOMETRY															
Ankle PF (nM)	5.7 ± 11.8	0.8 ± 10.2	0.07	0.20	0.44	-0.6 ± 5.1	-1.2 ± 3.6	0.65	0.75	-0.11	-1.5 ± 2.8	-3.3 ± 3.3	0.09	0.16	-0.49
Ankle DF (nM)	1.9 ± 3.9	1.2 ± 3.7	0.07	0.47	0.46	1.3 ± 1.3	1.0 ± 1.1	0.02	0.47	0.82	0.7 ± 0.8	0.8 ± 1.4	0.08	0.86	0.77
Ankle inversion (nM)	1.4 ± 7.7	0.9 ± 4.7	0.50	0.79	0.16	0.3 ± 3.1	0.5 ± 1.1	0.76	0.86	0.08	-0.4 ± 1.9	-0.1 ± 0.8	0.51	0.58	-0.20
Ankle eversion (nM)	1.7 ± 2.9	1.3 ± 3.0	0.046	0.69	0.53	0.4 ± 1.0	0.4 ± 1.6	0.21	0.97	0.30	0.3 ± 0.9	-0.1 ± 1.1	0.26	0.40	0.26

Values are annualised mean change/year in each measurement. p1=p-value for paired t-test in CMT1 group. p2=p-value of two-tailed t-test (CMT1 against controls). CMT1A=Charcot-Marie-Tooth disease type 1A, FF=fat fraction, MTR=magnetisation transfer ratio, RMA=remaining muscle area, CSA=cross-sectional area, MRC-UL/LL=MRC score in upper and lower limbs, SF-36=Short Form Health Survey 36, CMTES/SS=CMT examination/sensory scores, PF=plantarflexion, DF=dorsiflexion. Rows are highlighted in blue if both p values are <0.

Table 3-19 – Annualised change in mean calf muscle fat fraction in CMT1A

	Timepoint two					Timepoint three					Timepoint four				
Muscle	**Change (%/yr)**	**SD; 95% CI**	**p1**	**p2**	**SRM**	**Change (%/yr)**	**SD; 95% CI**	**p1**	**p2**	**SRM**	**Change (%/yr)**	**SD; 95% CI**	**p1**	**p2**	**SRM**
R. TA	1.3	1.8; 0.4 to 2.2	0.003	0.005	0.73	0.8	0.6; 0.4 to 1.2	0.001	<0.0001	1.32	0.9	0.8; 0.4 to 1.3	0.001	<0.0001	1.15
L. TA	0.9	1.8; 0.0 to 1.8	0.04	0.08	0.49	0.6	0.8; 0.1 to 1.1	0.003	0.002	0.74	0.7	0.9; 0.2 to 1.2	0.004	0.002	0.76
R. PL	2.1	2.8; 0.6 to 3.5	0.001	0.11	0.74	1.7	2.2; 0.4 to 2.9	0.002	0.01	0.76	1.8	2.3; 0.5 to 3.0	0.001	0.02	0.78
L. PL	1.8	3.6; -0.2 to 4.1	0.002	0.07	0.51	0.9	1.5; -0.3 to 1.7	0.05	0.31	0.60	1.0	1.4; -0.1 to 1.7	0.01	0.03	0.68
R. LG	1.4	2.5; 0.1 to 2.7	0.002	0.12	0.56	0.9	1.0; 0.3 to 1.5	0.001	0.00	0.84	0.6	0.9; 0.1 to 1.1	0.002	0.003	0.75
L. LG	0.9	1.3; 0.2 to 1.6	<0.0001	0.56	0.69	0.8	1.0; 0.2 to 1.4	0.01	0.048	0.78	0.8	1.2; 0.2 to 1.4	0.01	0.02	0.69
R. MG	1.4	2.5; 0.2 to 2.7	0.001	0.09	0.58	1.3	1.7; 0.4 to 2.3	0.001	<0.0001	0.80	1.1	1.5; 0.2 to 2.0	0.001	<0.0001	0.71
L. MG	1.5	3.0; -0.1 to 3.0	0.02	0.05	0.49	1.9	3.4; -0.1 to 3.9	0.001	0.06	0.56	1.3	1.7; 0.3 to 2.2	0.001	0.004	0.74
R. Sol	0.8	1.0; 0.3 to 1.3	0.003	0.15	0.78	0.5	0.5; 0.2 to 0.8	0.002	0.048	1.03	0.4	0.4; 0.1 to 0.6	0.004	0.31	0.84
L. Sol	1.2	2.6; -0.2 to 2.5	0.03	0.20	0.45	1.1	2.4; -0.3 to 2.5	0.02	0.19	0.47	0.9	1.7; -0.1 to 1.9	0.003	0.02	0.53
R. PT	1.3	2.6; 0.0 to 2.7	0.001	0.01	0.51	1.1	1.4; 0.3 to 1.9	0.001	<0.0001	0.78	0.7	0.6; 0.3 to 1.0	0.001	<0.0001	1.07
L. PT	2.1	4.4; -0.1 to 4.4	0.007	0.004	0.50	1.1	1.7; 0.1 to 2.1	0.001	0.001	0.63	1.0	1.5; 0.1 to 1.8	0.00	<0.0001	0.64
R. Triceps	1.1	1.2; 0.4 to 1.7	0.001	0.09	0.88	0.7	0.6; 0.3 to 1.0	0.001	0.01	1.10	0.5	0.6; 0.2 to 0.9	0.002	0.009	0.91
L. Triceps	1.0	2.1; -0.4 to 1.5	0.02	0.32	0.41	1.2	1.4; 0.1 to 1.9	0.004	0.049	0.74	1.0	1.1; 0.2 to 1.4	0.001	0.002	0.75
R. CALF	1.3	1.4; 0.6 to 2.0	0.001	0.01	0.91	0.9	0.7; 0.5 to 1.3	0.001	0.001	1.20	0.7	0.6; 0.3 to 1.1	0.001	0.007	1.07
L. CALF	1.0	1.0; 0.5 to 1.5	0.001	0.01	0.99	1.0	1.2; 0.3 to 1.7	0.001	0.01	0.85	0.8	0.8; 0.3 to 1.2	0.001	<0.0001	0.97
BOTH CALVES	1.2	1.1; 0.6 to 1.7	<0.0001	0.02	1.04	0.9	0.9; 0.5 to 1.4	0.002	0.006	1.01	0.7	0.6; 0.3 to 1.0	0.002	0.01	1.07

Change is given in % per year. p1=p-value of paired student t-test or Wilcoxon test (as appropriate) in CMT1A group, p2=p-value of Mann Whitney U test or two-tailed t-test (as appropriate) CMT1 against controls. CI=confidence interval, R=right, L=left, triceps=triceps surae group, TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG: medial gastrocnemius, Sol=soleus, PT=tibialis posterior. SRM=standardised response mean. Rows are highlighted in blue when both p values <0.05

3.5.4 Annualised longitudinal data

3.5.4.1 Clinical data

3.5.4.1.1 Summary clinical and functional measures

In the CMT1A group, there was a highly significant annualised mean change in CMTES at timepoints three and four with mean increase (± s.d.) of 0.6 ± 0.6 points/year (p=0.002) and 0.5 ± 0.5 points/year (p=0.001) respectively. Similarly, there was significant change in CMTES-LL at timepoints three and four. No other summary clinical or functional measure (CMTSS, SF-36 or subsets, MRC-UL, MRC-LL and MRC total score) changed significantly over the study duration (Table 3-18 and Table 3-20). There were no significant changes in any clinical measures in the control group over the study duration.

Table 3-20 – Summary of change in other clinical measures in CMT1A

	Timepoint two			Timepoint three			Timepoint four		
Clinical measure	**mean change**	**p**	**SRM**	**mean change**	**p**	**SRM**	**mean change**	**p**	**SRM**
CMTES-LL	0.5 ± 1.1	0.07	0.45	0.5 ± 0.5	0.003	0.97	0.5 ± 0.4	0.002	1.06
MRC-UL	0.5 ± 3.2	0.49	0.17	-1.7 ± 7.0	0.29	-0.24	-0.7 ± 3.4	0.45	-0.20
MRC-LL	-0.4 ± 3.8	0.67	-0.10	-1.9 ± 6.3	0.27	-0.31	-1.5 ± 3.5	0.12	-0.43
SF-36 (PF)	-1.1 ± 11.8	0.72	-0.09	-1.8 ± 3.5	0.09	-0.51	-0.6 ± 3.8	0.59	-0.15

Values are mean change in each of the clinical measures in points/year. Row highlighted in blue if change significant compared with baseline

3.5.4.2 Myometric data

There were no significant changes in ankle muscle strength measured by fixed dynamometry in any movements over the duration of the study. Although there was apparent improvement in ankle eversion strength in the CMT1A group at timepoint two, and of ankle dorsiflexion strength at timepoint three, these improvements were not significant compared with controls and may represent a learning effect. At the time of analysis, it was noted that ankle plantarflexion values in controls at visit four were markedly reduced compared with previous visits. Detailed checking of the dynamometer revealed no technical fault, and thus analysis was performed based on these values. Of note however, there was no significant change in ankle plantarflexion in the CMT1A cohort compared with baseline.

3.5.4.3 Quantitative MRI determined data

3.5.4.3.1 Fat fraction

3.5.4.3.1.1 Small region fat fraction

There was significant change of 1.4 ± 2.5%, (p=0.04 paired t-test, p=0.02 two-tailed t-test versus controls) in small region FF at timepoint two in CMT1A patients (Table 3-18), though this statistical significant difference was not maintained at timepoints three or four, due to smaller change and higher standard deviation/variability at those timepoints, both of which reflect the poor reliability of this technique of muscle segmentation. One notes that the magnitude of FF change was similar between small and large ROI, but with larger variability with the small ROIs. There were no significant changes in small region FF in the control group at any timepoint.

3.5.4.3.1.2 Whole muscle fat fraction

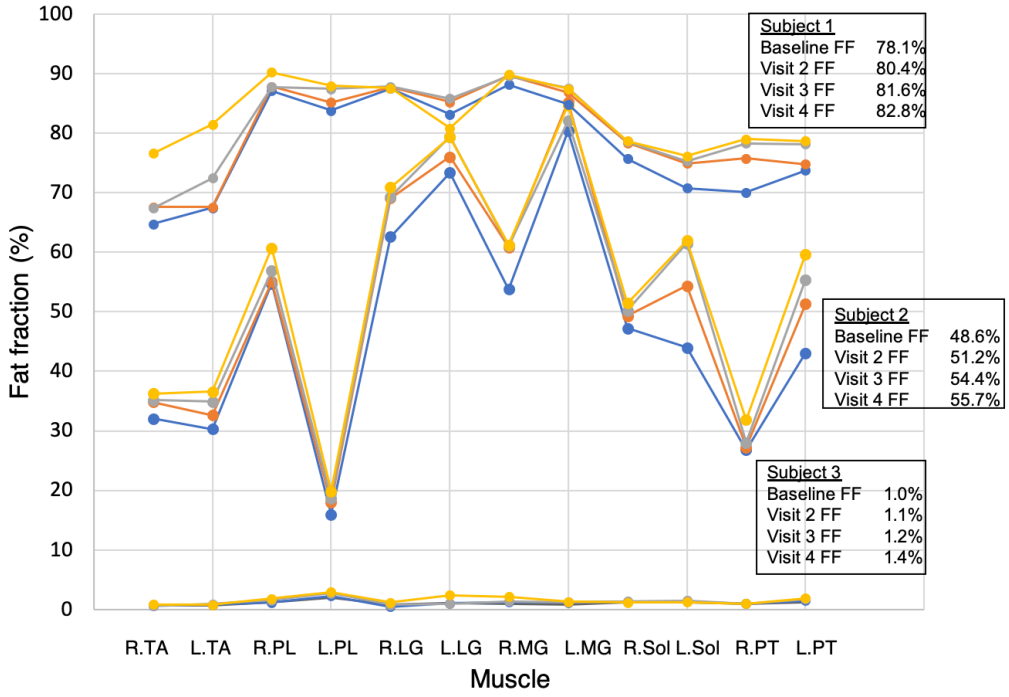
At each of the three follow-up timepoints, there was a significant increase in combined bilateral calf muscle FF: timepoint two FF change 1.2 ± 1.1% (p<0.0001 vs baseline, and p=0.02 vs controls), timepoint three: 0.9 ± 0.9% (p=0.002 vs baseline, and p=0.006 vs controls), timepoint four: 0.7 ± 0.6% (p=0.002 vs baseline, and p=0.01 vs controls) – Figure 3-9.

There were also significant changes in all other summary calf FF measures: right and left calves and each of the triceps surae and anterior groups individually (Figure 3-10). Individual calf muscle FF also changed significantly over the three timepoints with an increasing number of muscles showing significant change at timepoints three and four – at which point there was significant change in all muscles except right soleus – compared with timepoint two. The largest significant FF change in an individual muscle was in right tibialis anterior at timepoint two (FF change 1.3 ± 1.8%) and in right peroneus longus muscle at timepoints three and four with FF change (mean ± s.d.) of 1.7 ± 2.2% per year and 1.8 ± 2.3% per year respectively.

When muscles were grouped together in summary calf measures, FF change was ~1%/year but with lower standard deviation/variability compared with individual muscles, likely due to the larger muscle sample size analysed reducing the relative error of measurement.

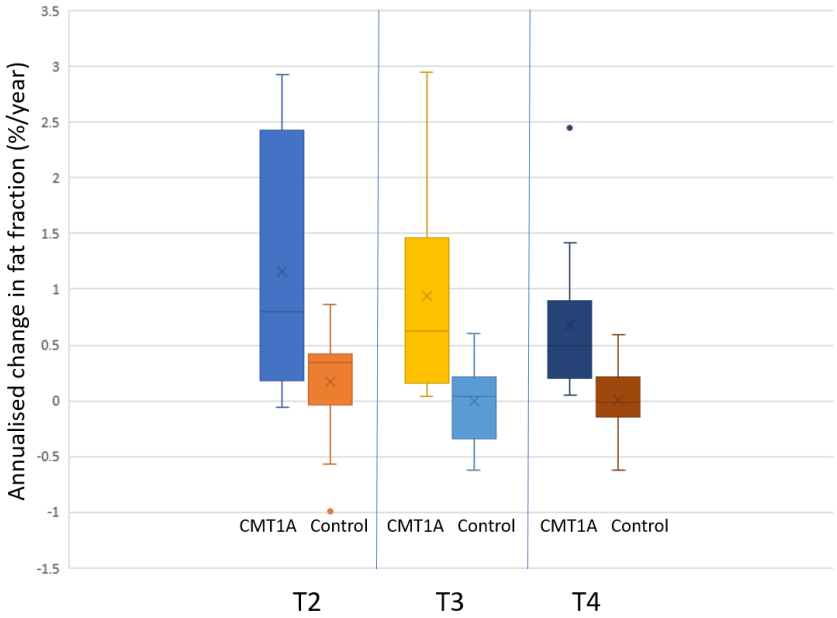
There was no significant FF change in controls in any individual muscle or summary calf FF measure at any timepoint. FF change for all individual muscles and summary measures for the CMT1A group is shown in Table 3-19, and for the control group in Table 3-21. Example line graph of calf fat accumulation in three representative CMT1A patients of differing clinical severity is shown in Figure 3-8.

Figure 3-8 – Progressive combined bilateral mean calf fat accumulation in CMT1A



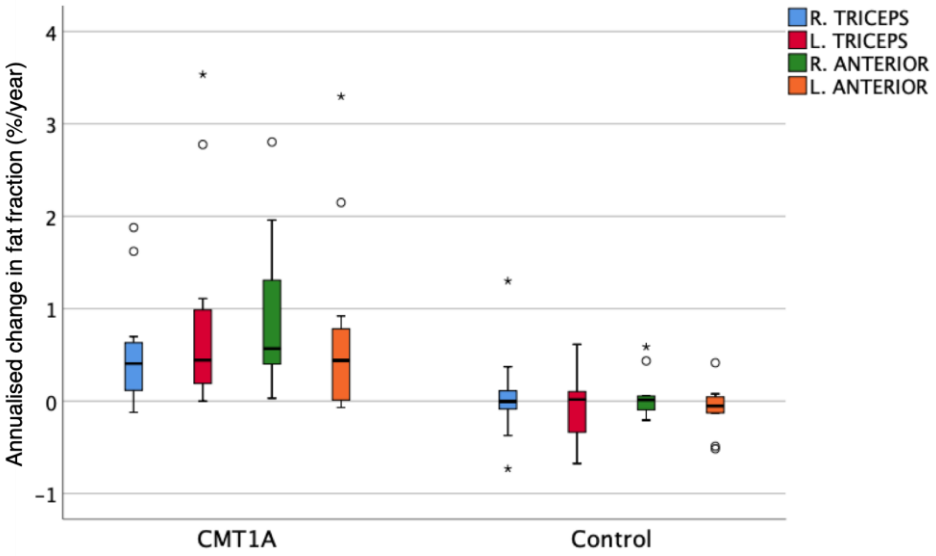
Each point represents a single muscle in a single participant. Lines are for visual aid, and are colour coded for visit number: blue=baseline, orange=visit 2, grey=visit 3, yellow=visit 4

Figure 3-9 – Annualised change in combined bilateral calf fat fraction in CMT1A and controls at timepoints two, three and four



CMT1A-Charcot-Marie-Tooth disease type 1A. T2=timepoint two, T3=timepoint three, T4=timepoint four. Boxes represent median and IQR, whiskers show range, Circles are outliers, x indicates mean.

Figure 3-10 – Fat fraction change in triceps surae and anterior muscle groups at timepoint four in CMT1A and controls



Boxes represent median and IQR, whiskers show range, and circles/asterisks are outlier. All mean differences between CMT1A and control are statistically significant at a level of p< 0.05. R-right, L=left, triceps=triceps surae group, anterior=anterior compartment. CMT1A=Charcot-Marie-Tooth disease type 1A

3.5.4.3.2 Cross-sectional area and remaining muscle area

CSA did not change significantly over the study duration in either the CMT1A or control group. RMA however changed significantly at timepoints three and four in several summary muscle measures: combined bilateral calf RMA change: -2.3 ± 2.6cm^2/year (p=0.006) and -2.0 ± 2.2cm^2/year (p=0.004) at timepoints three and four respectively. There was no significant change in RMA in the control group at any timepoint. Mean RMA change values for CMT1A and control groups are summarised in Table 3-22.

3.5.4.3.3 T2 relaxation time and magnetisation transfer ratio

Although the mean magnetisation transfer ratio (MTR) decreased and T2 time increased in CMT1A over the four timepoints as expected with progressive fat accumulation (Table 3-14), there were no significant changes in individual muscle or summary calf MTR or T2 measures compared with baseline or controls at any timepoint. This lack of change in overall T2 and MTR measures is in large part due to the use of small regions of interest resulting in high variability of these measures in CMT1A patients.

Table 3-21 – Mean calf muscle fat fraction change (% per year) in controls

	Timepoint two		Timepoint three		Timepoint four	
Muscle	**Change (%/yr)**	**SD; 95% CI**	**Change (%/yr)**	**SD; 95% CI**	**Change (%/yr)**	**SD; 95% CI**
R. TA	0.1	0.5; -0.2 to 0.3	0.0	0.3; -0.2 to 0.2	0.1	0.2; -0.1 to 0.2
L. TA	-0.1	0.6; -0.3 to 0.2	-0.1	0.4; -0.3 to 0.2	-0.1	0.3; -0.3 to 0.1
R. PL	0.3	1.2; -0.3 to 0.9	0.0	0.7; -0.5 to 0.5	0.4	1.5; -0.6 to 1.5
L. PL	0.2	1.3; -0.5 to 0.8	0.1	0.6; -0.4 to 0.5	0.0	0.5; -0.4 to 0.3
R. LG	0.2	1.4; -0.5 to 0.9	-0.2	0.9; -0.8 to 0.4	-0.2	0.6; -0.7 to 0.2
L. LG	0.9	1.9; -0.1 to 1.8	0.0	0.8; -0.6 to 0.6	-0.1	0.7; -0.6 to 0.4
R. MG	0.3	0.9; -0.1 to 0.7	-0.2	0.5; -0.5 to 0.2	-0.2	0.4; -0.6 to 0.1
L. MG	0.1	1.3; -0.5 to 0.7	-0.2	0.9; -0.8 to 0.5	-0.2	0.7; -0.7 to 0.3
R. Sol	0.3	0.5; 0.0 to 0.5	0.3	0.9; -0.3 to 1.0	0.4	0.9; -0.3 to 1.0
L. Sol	-0.1	0.9; -0.5 to 0.4	0.1	0.2; -0.1 to 0.2	0.1	0.2; -0.1 to 0.2
R. PT	0.1	0.5; -0.1 to 0.4	-0.1	0.2; -0.2 to 0.1	0.0	0.2; -0.1 to 0.1
L. PT	-0.2	0.8; -0.6 to 0.2	-0.3	0.5; -0.6 to 0.1	-0.2	0.4; -0.5 to 0.0
R. Triceps	0.3	0.7; -0.1 to 0.6	0.0	0.6; -0.4 to 0.4	0.1	0.5; -0.3 to 0.4
L. Triceps	0.3	0.7; -0.1 to 0.6	0.0	0.4; -0.3 to 0.3	-0.1	0.4; -0.3 to 0.2
R. CALF	0.2	0.5; 0.0 to 0.5	0.0	0.5; -0.3 to 0.4	0.1	0.4; -0.2 to 0.4
L. CALF	0.1	0.5; -0.2 to 0.4	0.0	0.4; -0.3 to 0.2	-0.1	0.3; -0.3 to 0.2
BOTH CALVES	0.2	0.5; 0.0 to 0.4	0.0	0.4; -0.2 to 0.2	0.0	0.4; -0.2 to 0.2

SD=standard deviation, CI=confidence interval, R=right, L=left, triceps=triceps surae group, TA=tibialis anterior, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior. All p values non-significant (not shown)

Table 3-22 – Change in remaining muscle area in CMT1A and controls

Timepoint two					
	CMT1A	**Control**			
Muscle	**Change in RMA (cm^2/yr)**		**p1**	**p2**	**SRM**
R. Triceps	-0.4 ± 1.4	0.2 ± 2.0	0.23	0.27	-0.31
L. Triceps	0.2 ± 1.6	0.1 ± 1.6	0.69	0.98	0.10
R. Anterior	-0.1 ± 0.4	-0.1 ± 0.5	0.36	0.89	-0.23
L. Anterior	0.3 ± 0.8	0.0 ± 0.4	0.22	0.32	0.31
R. Calf	-0.6 ± 1.4	0.3 ± 2.5	0.12	0.21	-0.40
L. Calf	0.1 ± 2.6	0.5 ± 2.6	0.90	0.66	0.03
Both Calves	-0.5 ± 3.8	0.8 ± 5.0	0.59	0.39	-0.13

Timepoint three					
	CMT1A	**Control**			
Muscle	**Change in RMA (cm^2/yr)**		**p1**	**p2**	**SRM**
R. Triceps	-1.0 ± 0.8	-0.3 ± 0.6	0.001	0.04	-1.18
L. Triceps	-0.8 ± 1.1	-0.2 ± 0.6	0.02	0.16	-0.71
R. Anterior	-0.1 ± 0.3	0.2 ± 0.3	0.51	0.10	-0.18
L. Anterior	0.0 ± 0.4	0.1 ± 0.2	0.75	0.63	-0.10
R. Calf	-1.4 ± 1.2	-0.1 ± 0.9	0.001	0.006	-1.21
L. Calf	-0.8 ± 1.6	0.1 ± 1.0	0.07	0.12	-0.52
Both Calves	-2.3 ± 2.6	0.0 ± 1.7	0.006	0.03	-0.87

Timepoint four					
	CMT1A	**Control**			
Muscle	**Change in RMA (cm^2/yr)**		**p1**	**p2**	**SRM**
R. Triceps	-1.1 ± 1.3	-0.4 ± 0.4	0.004	0.12	-0.87
L. Triceps	-0.4 ± 0.9	-0.2 ± 0.4	0.12	0.52	-0.40
R. Anterior	-0.1 ± 0.2	0.1 ± 0.2	0.26	0.10	-0.31
L. Anterior	0.0 ± 0.3	0.1 ± 0.2	0.51	0.05	0.18
R. Calf	-1.4 ± 1.7	-0.3 ± 0.6	0.005	0.046	-0.86
L. Calf	-0.5 ± 1.1	0.0 ± 0.7	0.09	0.17	-0.48
Both Calves	-2.0 ±2.2	-0.3 ± 1.2	0.004	0.03	-0.89

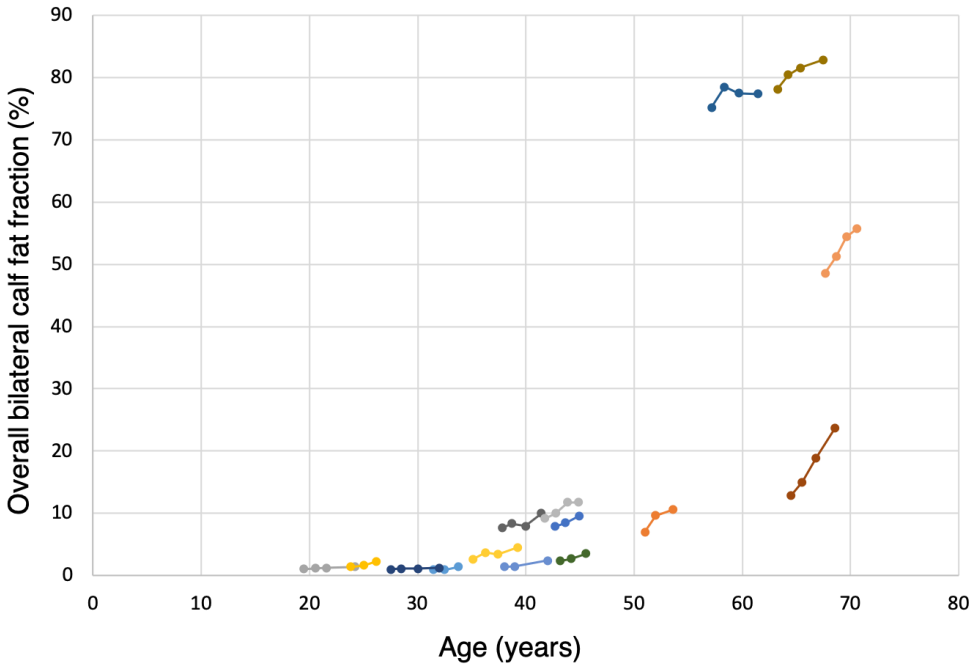
Values are annualised change in remaining muscle area (cm^2/year). p1=paired t-test in CMT1A. p2=two-tailed t-test (CMT1A versus controls). Rows highlighted in blue if both p1 and p2 <0.05. R=right, L=left. SRM=standardised response mean

3.5.5 Longitudinal Correlations

3.5.5.1 Age, duration of disease and combined bilateral calf fat fraction

There were statistically significant moderate positive correlations between baseline age and baseline combined bilateral calf FF (r_s=0.671, p=0.01) as well as high positive correlations between time in years since diagnosis and baseline combined bilateral calf FF (r=0.895, p=<0.0001). These statistically significant correlations persisted at visits 2, 3 and 4. Figure 3-11 demonstrates the relationship between patient age and combined whole calf FF in the CMT1A cohort, which appears not to be a clearly linear relationship, and perhaps better modelled on an exponential scale.

Figure 3-11 – Combined whole calf fat fraction at patient age for CMT1A



Each line represents a single patient with CMT1A, with points representing each visit

3.5.5.2 Changes in summary MRI and clinical measures

3.5.5.2.1.1 Timepoint two

Combined bilateral calf FF change did not correlate with change in any summary clinical or functional measure. There were moderate negative correlations between change in left calf FF and MRC-LL (r_s= -0.512, p=0.04), as well as between change in left triceps surae and total MRC score (r_s= -0.617, p=0.008). There were weak correlations between changes in clinical measures: CMTES-LL with CMTSS

(r_s=0.477, p=0.045) and SF36(PF) with MRC-LL (r_s= 0.540, p=0.04). There were no correlations between change in summary MRI and myometric measures.

3.5.5.2.1.2 Timepoint three

There were high positive correlations between combined bilateral calf FF change and change in CMTES (r_s=0.723, p=0.005) and between combined bilateral calf FF change and change in CMTES-LL (r_s= 0.669, p=0.009). Changes in other summary calf FF measures also showed moderate to high positive correlations with both CMTES and CMTES-LL but not with changes in other clinical or functional measures: right calf with CMTES (r_s= 0.610, p=0.02) and CMTES-LL (r_s= 0.563, p=0.04), left calf with CMTES (r_s=0.769, p=0.001) and CMTES-LL (r_s=0.695, p=0.006), right triceps with CMTES (r_s=0.650, p=0.01) and CMTES-LL (r_s=0.576, p=0.03), left triceps with CMTES (r_s=0.756, p=0.002) and CMTES-LL (r_s=0.748, p=0.002), left anterior compartment with CMTES (r_s= 0.758, p=0.02) and CMTSS (r_s=0.602, p=0.02). There were no correlations between change in summary MRI and myometric measures.

3.5.5.2.1.3 Timepoint four

There were high positive correlations between changes in CMTES and changes in combined bilateral calf muscle FF (r_s=0.706, p=0.005) – Figure 3-12, moderate positive correlations with change in CMTES-LL (r_s=0.638, p=0.01) and moderate negative correlations with change in SF36(PF) (r_s= -0.604, p=0.03). Changes in other summary calf FF measures also correlated with changes in overall clinical measures: left calf with CMTES (r_s=0.761, p=0.002) and CMTES-LL (r_s=0.715, p=0.004), right triceps with CMTES (r_s=0.567, p=0.04) and CMTES-LL (r_s= 0.569, p=0.04), left triceps with CMTES (r_s=0.713, p=0.004) and CMTES-LL (r_s=0.737, p=0.003). Again, there were no correlations between change in summary MRI and myometric measures.

3.5.5.3 Changes in MRI measures

3.5.5.3.1.1 Timepoint two

There were no correlations between changes in summary FF measures and matching RMA or CSA. There were moderate/high negative correlations between change in small region combined bilateral calf FF, left calf FF and right calf FF and the matching change in MTR (r_s= -0.777, p<0.0001; r_s= -0.629, p=0.009 and r_s= -0.646, p=0.009 respectively) but not between FF change and T2 change. Change in muscle T2 and MTR did not correlate.

3.5.5.3.1.2 Timepoint three

There were no correlations between changes in summary FF measures and changes in matching CSA. There were moderate negative correlations between small region

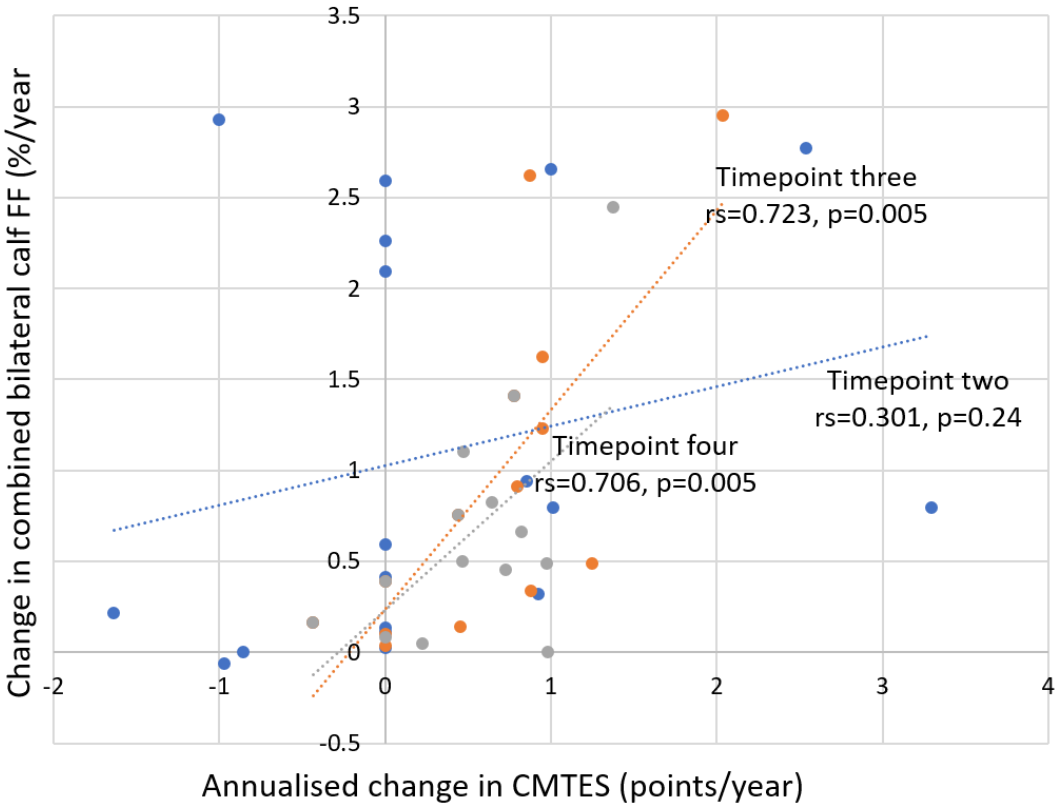
right anterior compartment FF and matching RMA (r_s= -0.569, p=0.03). Small region left anterior compartment FF correlated moderately strongly with change in matching T2 (r_s=0.534, p=0.049). MTR and T2 change did not correlate.

3.5.5.3.1.3 Timepoint four

There were no correlations between changes in summary FF measures and changes in matching CSA. There were moderate negative correlations between changes in whole region combined bilateral calf FF and RMA (r_s= -0.512, p=0.02) – Figure 3-13. Change in small region left anterior compartment FF correlated strongly with matching T2 change (r_s=0.746, p=0.001). Changes in T2 or small region FF did not correlate with MTR.

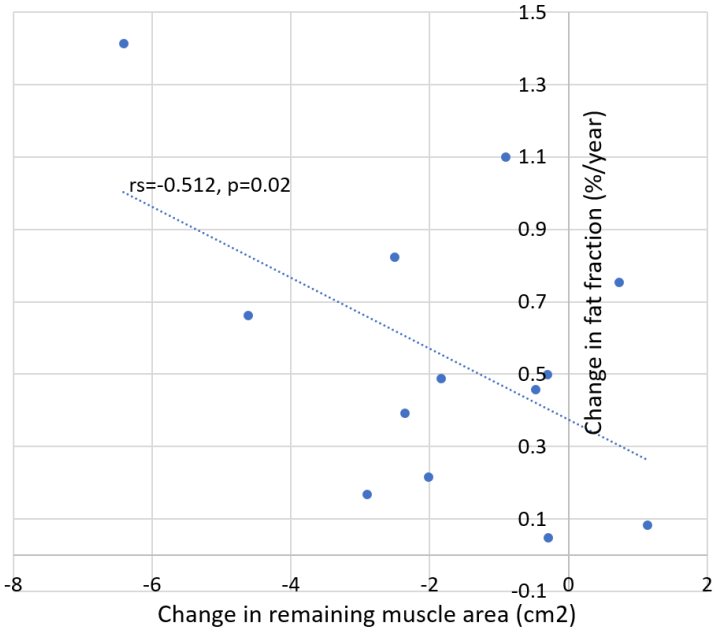
Correlation between MRI determined FF and clinical/functional measures is illustrated in Figure 3-12, Figure 3-13 and Figure 3-14.

Figure 3-12 – Correlation between change in CMTES and combined bilateral calf fat fraction change in CMT1A



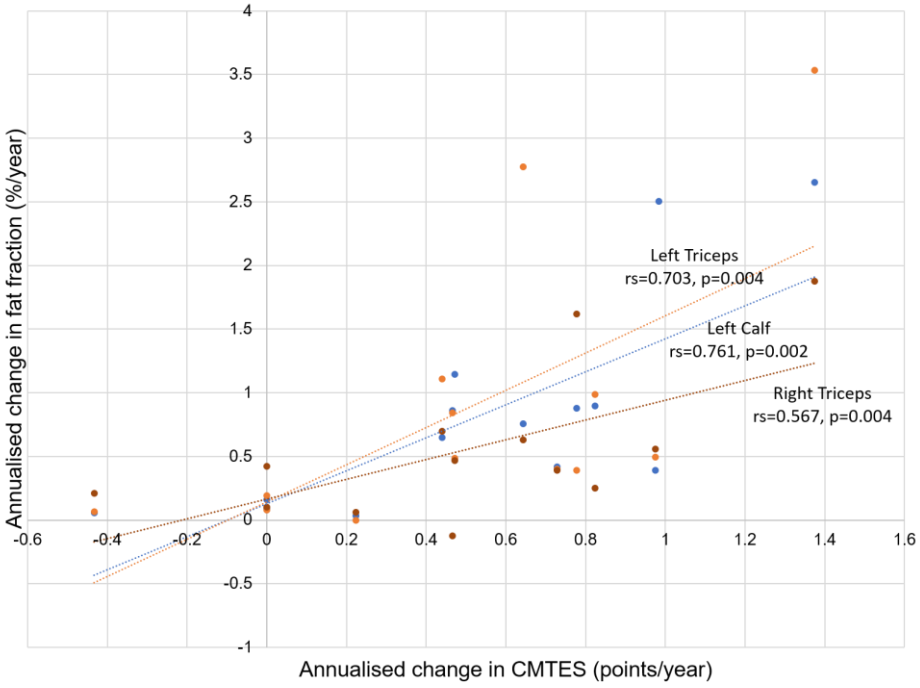
In the CMT1A group, annualised mean change in CMTES correlated strongly with annualised change in combined bilateral calf fat fraction at timepoints three and four, but not at timepoint two. Figure is colour coded for timepoint: blue=two, orange=three, grey=four, CMTES=CMT examination score. Correlation is by Spearman test.

Figure 3-13 – Correlation between change in combined bilateral calf fat fraction and remaining muscle area at timepoint four in CMT1A



At timepoint four in the CMT1A group, annualised mean change in RMA correlated moderately strongly with annualised change in combined bilateral calf fat fraction

Figure 3-14 – Correlation between change in CMTES and summary MRI derived fat fraction measures at timepoint four in CMT1A



CMT1A=Charcot-Marie-Tooth disease type 1A, CMTES=CMT examination score, L.=left, R.=right. Figure is colour coded for muscle group: orange=L. triceps, red=R. triceps, light blue=L. calf. Each point represents a single patient with CMT1A

3.5.6 Predicting longitudinal fat fraction change

At all follow-up timepoints, baseline combined bilateral calf FF predicted annualised FF change. Similarly, baseline STIR predicted FF change, although combining the two did not improve the strength of the prediction at any timepoint. Baseline age correlated strongly with FF change at all timepoints, though the relationship appears to be non-linear.

3.5.6.1 Baseline fat fraction

A scatterplot of CMT1A patient baseline combined bilateral FF against annualised FF change is shown in Figure 3-15 and for all CMT1A muscles in Figure 3-17. An apparent relationship was noted between baseline FF and FF change at each timepoint, with greatest change seen in muscles with an intermediate FF. This relationship was thus explored further by subgroup analysis based on baseline FF.

At baseline, mean combined bilateral calf FF in the control group was 3.2 ± 1.6%, thus defining the upper limit for normal combined bilateral calf FF in CMT1A as ≤6.4% (95th centile). The twenty patients with CMT1A were thus divided into three groups based on baseline combined bilateral mean calf FF: 'normal' if FF ≤6.4% (n=11), 'intermediate' if FF between 6.4% and 60% (n=7), and 'high' if FF >60% (n=2).

Baseline combined bilateral mean calf FF was 2.0 ± 1.3%, 18.8 ± 15.6% and 81.3 ± 1.5% in the 'normal', 'intermediate' and 'high' baseline FF groups respectively. Mean change at timepoint four was 0.3 ± 0.2% (p=0.008 for paired t-test, p<0.001 for two-tailed t-test), 1.2 ± 0.7% (p=0.02 for paired t-test. Inadequate number for two-tailed t-test against controls) and 0.8 ± 0.4% (p=0.23 paired t-test. Inadequate number for two-tailed t-test against controls) for the three subgroups respectively. Mean change was significant at all timepoints in the 'normal' and 'intermediate' FF groups, but not significant in the 'high' group at any timepoint – a reflection of fewness of numbers. Maximum FF change was seen in patients with baseline combined bilateral FF >6.4%. At timepoint two, FF change was greatest when combining 'intermediate' and 'high' baseline FF groups (FF change 1.9 ± 0.9%), and at timepoints three and four, FF change in the 'intermediate' group was greatest.

When patients with 'intermediate' and 'high' baseline FF were combined (n=9), mean baseline FF was 32.7 ± 30.7% with highly significant FF change of 1.9 ± 0.9% (p=0.0003), 1.5 ± 0.9% (p=0.003) and 1.1 ± 0.7% (p=0.005) at timepoints two, three and four respectively. Mean FF change for each CMT1A subgroup is shown in Table 3-23.

There was a significant difference in FF change between those patients with 'normal' and those with 'intermediate' baseline FF (p=0.001, p=0.02 and p=0.008 for timepoints two, three and four respectively) and between those with 'normal' and those with a combination of 'intermediate' and 'high' baseline FF at all timepoints (p<0.0001, p=0.01 and p=0.01 for timepoints two, three and four respectively). NB – the number of patients in the 'high' group too little to conduct one-way ANOVA comparing the three group means, thus the 'high' group was excluded from analysis.

Further subdividing the 'intermediate' group of patients is likely to lead to further refinement, though numbers in this case (n=7) were inadequate for this analysis to be done.

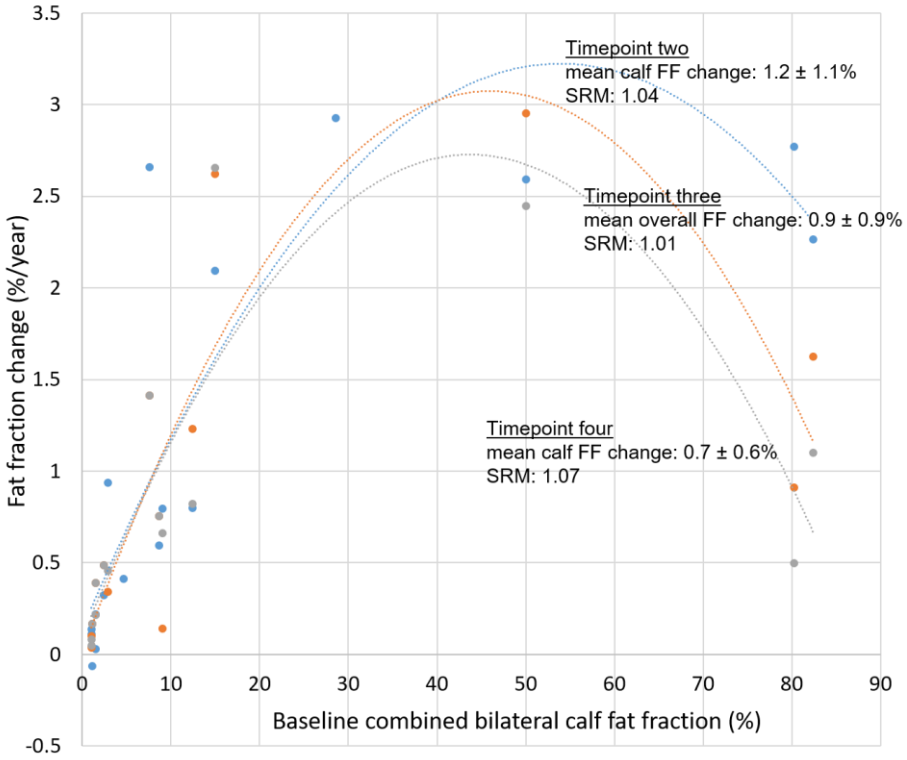
Mean change ± standard deviation and standardised response mean for these baseline FF groups is summarised in Figure 3-16 and Table 3-23.

Table 3-23 – Mean fat fraction change in CMT1A based on baseline fat fraction

	Timepoint two			Timepoint three			Timepoint four		
Group	**mean FF change**	**p**	**SRM**	**mean FF change**	**p**	**SRM**	**mean FF change**	**p**	**SRM**
Normal	0.3 ± 0.3	0.049	0.84	0.3 ± 0.2	0.04	1.42	0.3 ± 0.2	0.008	1.46
Intermediate	1.8 ± 1.0	0.004	1.75	1.5 ± 1.1	0.02	1.40.	1.2 ± 0.7	0.02	1.63
High	2.5 ± 0.4	0.06	7.00	1.3 ± 0.5	0.17	2.51	0.8 ± 0.4	0.23	1.88
Int. and High	1.9 ± 0.9	0.0003	2.05	1.5 ± 0.9	0.003	1.55	1.1 ±0.7	0.005	1.65

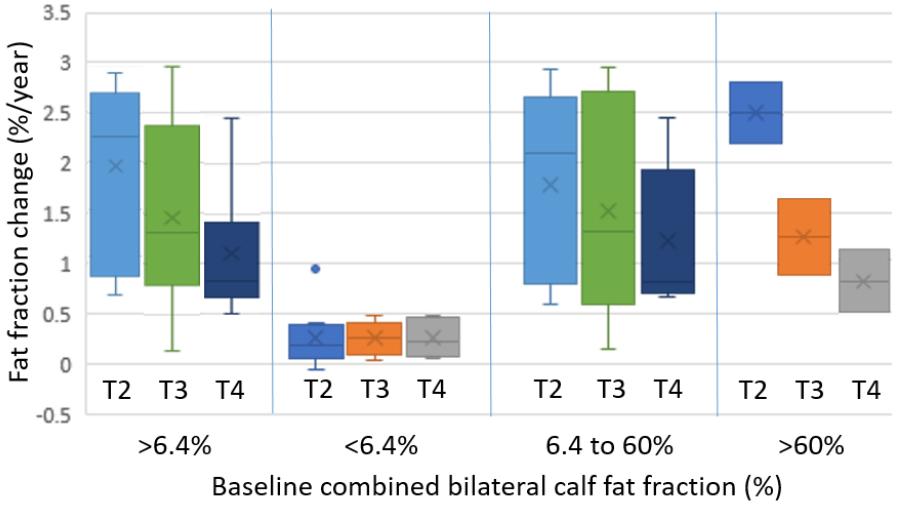
Values are change in combined bilateral fat fraction ± standard deviation. (%/year). Normal=FF<6.4%, moderate=6.4%<FF<60%, High=FF>60%. P value is for paired t-test against baseline in CMT1A group. SRM=standardised response mean. Row highlighted in blue if paired t-test p value <0.05

Figure 3-15 – Fat fraction change against baseline combined bilateral calf fat fraction, at timepoints two, three and four, in CMT1A



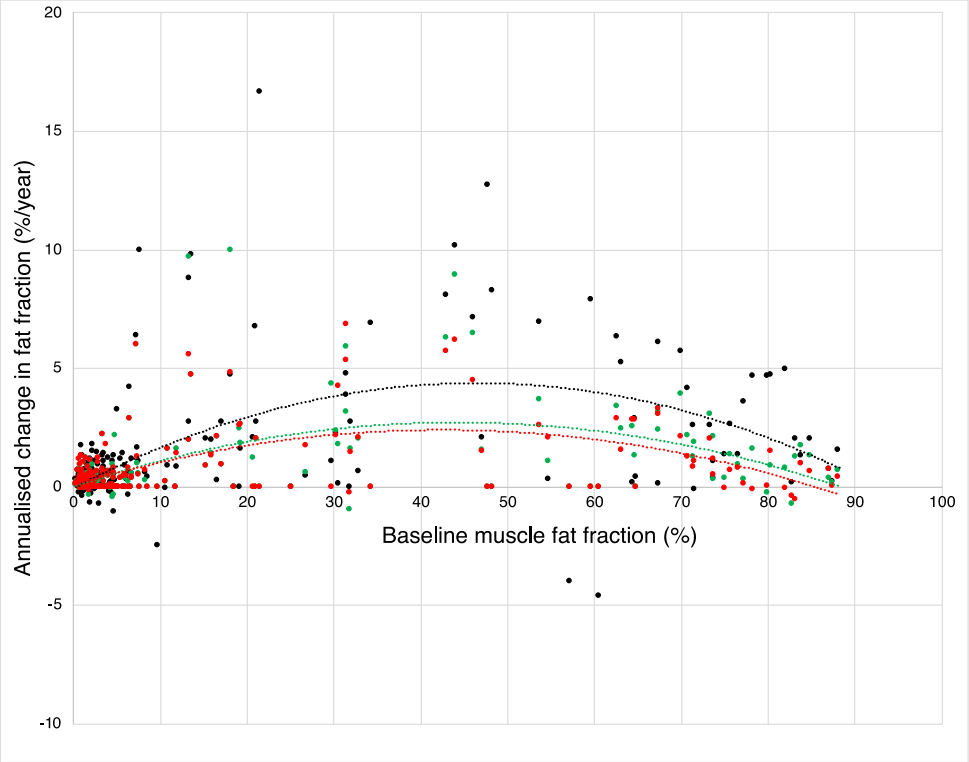
Each dot represents a single patient. Figure is colour coded for timepoint: blue= timepoint two, orange=timepoint three, grey=timepoint four. Superimposed lines are 2 order polynomial trendlines for visual aid

Figure 3-16 – Fat fraction change in CMT1A patients divided by baseline fat fraction



Boxes represent median and IQR, whiskers show range, and circle is an outlier. X=mean. Difference between means is statistically significant (excluding FF>60% group). T2=timepoint two, T3=timepoint three, T4=timepoint four.

Figure 3-17 – Fat fraction change in individual CMT1A muscles plotted against baseline combined bilateral calf fat fraction



All three follow-up visits are combined. Each dot represents a single muscle. Figure is colour coded for timepoint: black = timepoint two, green=timepoint three, red=timepoint four. Lines are 2 order polynomial trend lines for visual aid

3.5.6.2 Baseline STIR hyperintensity

There was a statistically significant difference between groups with differing baseline STIR hyperintensity by one-way ANOVA. A Tukey post hoc test revealed that at timepoint two, mean FF change was statistically significantly higher in CMT1A patients' muscles with baseline mild STIR changes (mean FF change: 2.0 ± 2.4%) compared with those without STIR changes (mean FF change: 1.0 ± 1.8%, p=0.02). Similarly at timepoints three and four, annualised mean FF change was statistically significantly higher in muscles with baseline mild STIR changes (1.6 ± 2.2% and 1.4 ± 1.7% for timepoints three and four respectively) compared with those without STIR changes (0.8 ± 1.4%, p=0.01; 0.8 ± 1.2%, p=0.03). Greatest change was seen in those muscles with baseline mild STIR hyperintensity (Table 3-24). In baseline FF CMT1A subgroups (normal, intermediate and high), there was no significant difference in annualised FF change, related to the presence or absence of STIR hyperintensity (one-way ANOVA/Tukey post hoc).

Table 3-24 – Fat fraction change divided by STIR signal in CMT1A and controls

		CMT1A	Controls
Timepoint	**STIR hyperintensity**	**mean change (%/year) ± sd**	
Two	Absent	1.0 ± 1.8	0.2 ± 1.1
	Mild	2.0 ± 2.4*	0.1 ± 1.0
	Marked	0.9 ± 2.1	ND
Three	Absent	0.8 ± 1.4	-0.1 ± 0.6
	Mild	1.6 ± 2.2*	0.3 ± 0.7
	Marked	1.3 ± 1.9	ND
Four	Absent	0.8 ± 1.2	-0.0 ± 0.7
	Mild	1.4 ± 1.7*	0.3 ± 0.7
	Marked	1.3 ± 1.3	ND

*STIR=short tau inversion recovery, sd=standard deviation, *significant at a level of p <0.05 for one-way ANOVA/Tukey post hoc test, ND=no data*

3.5.6.3 Baseline T2 time

To negate the marked effect of progressive fat accumulation on T2 relaxation time (Carlier, 2014; Hollingsworth, 2014), we examined T2 time in muscles with normal baseline FF, as defined by the 95th centile in baseline combined bilateral FF by small region ROIs in controls (baseline FF <3.7%).

Baseline FF in the CMT1A group was 1.5 ± 0.7%, and in the control group 1.3 ± 0.8% (p=0.02). Baseline T2 time in these two groups was 46.2 ± 6.1ms in CMT1A and 44.1 ± 3.7ms in controls (p=0.0002) demonstrating that even in CMT1A muscles with 'normal' FF, T2 time is significantly raised, which is in keeping with a period of water accumulation occurring prior to fat infiltration.

T2 change in the normal baseline FF CMT1A group was 1.1 ± 2.6ms/year, 0.7 ± 2.0ms/year and 0.7 ± 1.7ms/year respectively for timepoints two, three and four (all changes were significant compared with baseline values) compared with -0.5 ± 2.7ms/year, -0.05 ± 1.3ms/year and 0.2 ± 1.3ms/year in controls (no significant change with baseline).

It is of some interest that CMT1A muscles with normal baseline FF (FF<3.5%) showed significant T2 change over time (with low responsiveness, SRM= 0.42, 0.35 and 0.41 at timepoints two, three and four respectively), whereas in the overall cohort, there was no significant T2 change. This deserves further investigation by separating $T2_{water}$ from $T2_{fat}$ at point of acquisition. T2 change results for normal muscles in CMT1 and control groups are summarised in Table 3-25.

Table 3-25 – T2 change in muscles with baseline FF <3.5% at all timepoints in CMT1A and controls

	CMT1A	Control	p
Mean baseline FF (%)	1.5 ± 0.7	1.3 ± 0.8	0.02
T2 change t.p.2 (ms/yr)	1.1 ± 2.6	-0.5 ± 2.7	<0.0001
T2 change t.p.3 (ms/yr)	0.7 ± 2.0	-0.1 ± 1.3	0.003
T2 change t.p.4 (ms/yr)	0.7 ±1.7	0.2 ± 1.3	0.01

FF=fat fraction, p for unpaired two-tailed t-test with control group. t.p.=timepoint

3.5.7 Standardised response mean

In the CMT1A group, significant annualised change was seen at all timepoints in qMRI determined FF measures, and at timepoints three and four in RMA and CMTES/CMTES-LL score.

Combined bilateral calf FF showed significant change with large responsiveness at all timepoints: SRM of 1.04, 1.01 and 1.07 for timepoints two, three and four. Similarly, all other summary calf FF measures showed significant change with large responsiveness (Table 3-19).

Without subgroup analysis, maximum standardised response mean (SRM) across all outcome measures was 1.32 for qMRI measured right anterior group FF at timepoint three. At all timepoints, both mean change and standard deviation for summary FF measures was less than for individual muscles with a ratio favouring increased SRM. This is likely due to lower variability/measurement error with analysis of larger muscle area.

Stratification of CMT1A patients based on baseline FF resulted in marked improvement in outcome measure SRM (Table 3-23) across all homogenised subgroups, with maximum SRM seen when excluding 'normal' CMT1A muscles at all three timepoints (i.e. when combining muscles with 'intermediate' and 'high' baseline FF). Improvement in SRM was also noted however in the 'normal' FF group compared with whole group analysis. Excluding patients with 'normal' baseline FF from analysis led to largest study responsiveness: FF change 1.9 ± 0.9%/year, SRM=2.05 at timepoint two.

RMA showed significant change at timepoints three (right triceps, right calf, combined bilateral calf) and four (right calf, combined bilateral calf) with large responsiveness in these measures: SRM between -0.86 and -1.21 (Table 3-22).

T2 change in CMT1A muscles with ‘normal’ FF was significant at each timepoint, though with low responsiveness.

The only clinical parameters with SRM > 1 were CMTES and CMTES-LL at timepoints three and four with large responsiveness at both timepoints: SRM of 1.00 and 1.12 respectively for CMTES, and 0.97 and 1.06 for CMTES-LL.

SRM of MRI measures, clinical data and myometry at timepoints two, three and four are summarised in Table 3-18 and for individual muscles in Table 3-19.

3.6 Discussion

As drug trials for CMT1A and other NMD fast approach, clinically meaningful and highly responsive outcome measures are urgently needed.

This natural history study is the longest and most comprehensive ongoing single centre study of patients with CMT1A, having followed patients and matched controls for a period of up to five years to date. We have gathered detailed and valuable natural history data, in addition to novel results regarding a battery of potential outcome measures, strengthening our previously published 12 month data in this cohort (Morrow et al., 2016).

In brief, results demonstrate the 3-point Dixon method of FF calculation to be a highly responsive and longitudinally valid outcome measure over up to five years in CMT1A. Further, we demonstrate several findings over extended follow up which were not apparent at 12 months, including significant changes with associated large responsiveness in other summary qMRI calf FF measures and functional measures.

3.6.1 Cross-sectional assessment

Cross-sectional results are in keeping with our previously published work, and are briefly discussed here with mention of several new findings.

3.6.1.1 Baseline qMRI measures and pattern of muscle fat infiltration

Baseline bilateral combined calf FF in CMT1A was significantly increased, and CSA and RMA significantly reduced in all summary calf measures and most individual muscles compared with controls, in keeping with overall chronic fatty atrophy of calf muscles.

At baseline, peroneus longus and medial gastrocnemius were most affected, and tibialis posterior least affected by fat infiltration. This pattern, is well described in the literature (Berciano et al., 2006; Chung et al., 2008), however we report a novel finding in the CMT1A cohort of significantly increased tibialis posterior CSA (right posterior tibial CSA: 4.7 ± 1.2cm^2 in CMT1A versus 3.8 ± 1.3cm^2 in controls, two-tailed t-test. p=0.04) in the presence of significantly reduced CSA in all other muscles. Although pseudohypertrophy is recognised in CMT1A (de Freitas et al., 2015), this is not the case here given the relatively lower tibialis posterior muscle FF compared with other calf muscles, suggesting rather that there is a degree of early deep posterior group muscle hypertrophy, perhaps partially compensating for earlier superficial ankle plantarflexor atrophy. Relating to this, is that in both calves, the soleus muscle showed reduced CSA without accompanying significant FF increase, consistent

perhaps with a degree of atrophy of these large ankle plantarflexors. This hypothesis requires further assessment. On the other hand, the dual processes of fat infiltration and muscle atrophy appear to be temporally distinct in the peroneus longus and medial gastrocnemius (MG) muscles, both of which are the most affected by fat infiltration in this and other published cohorts. In both of these muscles, a significant increase in FF compared with controls, is not accompanied by a significant reduction in CSA, demonstrating a temporal disconnect between these processes, at least in these two muscles.

T2 time was significantly longer and MTR significantly lower in CMT1A than in controls, which is a reflection of the effect of fat accumulation on these measures, a well recognised finding which is again confirmed in this thesis by linear regression analysis. T2 time was shown to be significantly longer compared with controls in 'normal' CMT1A muscles consistent with there being an inflammatory process measurable by quantitative T2 imaging, prior to fat accumulation in these muscles, as has been described in other NMD (Friedman et al., 2014; Kan et al., 2009; Studýnková et al., 2007). This may be important in future in targeting treatments in early disease, and it would be worthwhile assessing this further in future studies by separating $T2_{water}$ from $T2_{fat}$ across all FF (Janiczek et al., 2011).

3.6.1.2 Cross sectional correlations

Importantly, cross-sectional results confirmed high criterion validity of MRI measures of fatty atrophy, by strong correlations between qMRI determined FF, strength and CMTES. Construct validity was also demonstrated for qMRI FF by strong correlations with Mercuri scores in individual muscles, through the finding of significant differences between FF, T2 and MTR measures in CMT1A, as well as by demonstration of calf muscle atrophy, in keeping with clinical understanding of the disease.

Demonstration of cross-sectional qMRI validity although necessary, is inadequate however to be certain of the significance of a longitudinal response for the patient. Longitudinal correlation with appropriate gold standards is also needed. As an illustrative example, in a study in patients with osteoporosis, Riggs and colleagues showed that although sodium fluoride significantly increased overall bone mass by up to 12% (which was elsewhere strongly correlated with reduced fracture rate), it actually resulted in increased skeletal fragility and fractures due to formation of brittle bone (Riggs et al., 1990).

3.6.2 Longitudinal assessment

3.6.2.1 Quantitative MRI outcome measures

The most impressive and clinically relevant longitudinal findings from this study are that qMRI determined FF by 3-point Dixon is shown to have persistent large internal responsiveness over prolonged study duration, and equally importantly to have excellent longitudinal criterion validity by correlation with longitudinal changes in CMTES/CMTES-LL and RMA.

3.6.2.1.1 Fat fraction

Quantitative MRI measured significant FF change at all timepoints in CMT1A patients: combined bilateral calf mean FF change of 1.2 ± 1.1%, 0.9 ± 0.9% and 0.7 ± 0.6% at timepoints two, three and four respectively with large responsiveness at each timepoint, SRM: 1.04, 1.01 and 1.07. Similarly in other summary calf FF measures, qMRI measured significant change at all three timepoints, with moderate to large responsiveness in all measures at timepoints three and four. Maximum qMRI responsiveness at timepoint two was in combined bilateral calf (SRM 1.04), and at timepoints three and four in the right anterior muscle group with significant FF change of 0.8 ± 0.6%/year (SRM=1.32) at timepoint three, and 0.9 ± 0.8%/year (SRM=1.15) at timepoint four.

In individual muscles, only right tibialis posterior was shown by qMRI to change significantly at timepoint two. This difference between individual and summary measures is likely explained by a difference in muscle sample size compared with summary measures, with the larger summary measures showing less sampling error and variability. As the duration of the study increased however, qMRI was able to measure significant FF change with moderate to large responsiveness in almost all individual muscles (Table 3-19).

For both individual muscle and summary calf FF measures, there was a clear step up in qMRI responsiveness from timepoint two to three, due in almost all cases to a large fall in standard deviation of mean change, rather than an increase in mean change itself.

3.6.2.1.2 Cross-sectional area and remaining muscle area

Whilst there were no significant changes in CSA at any timepoint, RMA changed significantly at timepoints three and four: combined bilateral calf RMA change: -2.3 ± 2.6cm^2/year and -2.0 ± 2.2cm^2/year respectively with large responsiveness: SRM=-0.87 and -0.89. Further, several other summary calf RMA measures also showed significant change at timepoints three and four, including right triceps at timepoint

three and right calf at timepoints three and four, again with large responsiveness at each timepoint. The disparity between changes in CSA and RMA is indicative of RMA being a more sensitive marker of muscle fatty atrophy (by virtue of its derivation from a combination of CSA and FF), and hence being more sensitive to fat accumulation than CSA. This is in keeping with findings from a recent study in patients with LGMDR9 (Murphy et al., 2019), though has not been demonstrated previously in CMT1A.

For each CMT1A calf muscle group in which there was significant RMA change, the SRM for the change in RMA was less than the corresponding SRM for FF change (except in right triceps at timepoint three). This finding suggests that variation in the CSA measurements reduces responsiveness of RMA over time (Table 3-26), thus making RMA a less responsive outcome measure than FF, albeit marginally in this case.

Table 3-26 – Comparison of fat fraction and matching remaining muscle area SRM in CMT1A

	Timepoint three				Timepoint four			
	RMA		FF		RMA		FF	
Muscle group	SRM	p	SRM	p	SRM	p	SRM	p
R. Triceps	-1.21	0.001	1.10	0.001				
R. Calf	-1.18	0.001	1.20	0.001	-0.82	0.005	1.07	0.007
Both Calves	-0.87	0.006	1.01	0.002	-0.89	0.004	1.07	0.002

R.=right, RMA=remaining muscle area, FF=fat fraction, SRM=standardised response mean

3.6.2.2 Clinical and functional outcome measures

Surprisingly, in light of findings from a number of previous studies (Chapter 1 – Burns et al., 2009; J. et al., 2009; Lewis et al., 2013; Pareyson et al., 2011; Verhamme et al., 2009), change in CMTES which was not significant at timepoint two (0.3 ± 1.2 points/year), showed significant change at longer follow up with mean change of 0.6 ± 0.6 points/year (p=0.002) and 0.5 ± 0.5 points/year (p=0.001) at timepoints three and four respectively. Compared with published data from previous natural history studies, the magnitude of change in CMTES was comparable, though the standard deviation of change was considerably smaller reflecting a more homogenous patient population. Similarly the CMTES-LL changed significantly at timepoints three and four: mean change of 0.5 ± 0.5 points/year and 0.5 ± 0.4 points/year respectively. These two subsets of the CMTNS thus showed large responsiveness at extended follow up (SRM of 1.06 and 1.12 respectively for CMTES-LL and CMTES at ~five years) and are without doubt worth combining alongside qMRI measures in studies of longer duration.

There were no significant changes in any other clinical or functional measures at any timepoint, making these less important to include in future studies. It is unfortunately not possible for the author to comment conclusively on ankle plantarflexion by myometry at timepoint four given that values in controls were systematically smaller than previous timepoints across the cohort – likely reflecting measurement error, however importantly there were no significant changes in ankle plantarflexion seen in CMT1A compared with baseline at any timepoint.

3.6.2.3 Longitudinal correlations

There were strong and significant correlations seen between qMRI and clinical/functional measures at all three follow-up timepoints.

Perhaps the most important longitudinal correlation demonstrated was between change in CMTES and change in combined bilateral calf FF at timepoints three (r_s= 0.723, p=0.005) and four (r_s=0.706, p=0.005). There were also moderate positive correlations between combined bilateral calf FF change and CMTES-LL change at timepoints three (r_s=0.669, p=0.009) and four (r_s=0.638, p=0.01), as well as between other summary FF measures and CMTES/ES-LL and CMTSS at timepoints three and four. This demonstrates for the first time in a CMT1A cohort, strong longitudinal criterion validity of qMRI determined FF. As discussed previously, this opens the door for qMRI to be applied to other similar diseases.

There were also moderate negative correlations between change in combined bilateral calf FF and change in RMA only at timepoint four (r_s=-0.512, p=0.02) which provides further evidence of longitudinal validity in this cohort.

3.6.2.4 Optimisation of quantitative MRI outcome measure responsiveness

This study has demonstrated large responsiveness of qMRI determined FF over prolonged follow up in CMT1A. Here I discuss relevant factors which may be used to further optimise responsiveness.

3.6.2.4.1 Trial duration

One of the indirect methods by which outcome measure responsiveness may be increased is by extending trial duration, thus maximising change of the specific measure over time. However if the group is not homogenous, standard deviation of change may also increase which is counterproductive.

In this study, as summarised in Table 3-27, there was a clear increase in responsiveness of individual muscle and most summary measures between timepoints two and three, and also between timepoints two and four, however

between timepoints three and four, the increase was less marked and less uniform, with responsiveness of some measures improving, whilst others did not. In almost all cases, the increase in SRM from timepoint two to three was predominantly due to a fall in standard deviation as mean annualised change remained static or indeed fell in some cases. Examining responsiveness of the stratified groups (Table 3-23) garnered no further support for increased trial duration.

Results, although mixed, suggest therefore that extending trial duration beyond timepoint three (two-year follow up) does not necessarily ensure a clear increase in qMRI responsiveness. This is due to the fact that participants are heterogenous in respect of the primary outcome. The ideal situation for increasing outcome measure responsiveness would be a group of homogenous patients in whom the expected change is very similar, resulting in low standard deviation and high outcome measure responsiveness. In such a situation, extending trial duration would be beneficial, but in reality such a cohort is unlikely unless strict enrolment criteria are instituted which in turn would limit external validity of results.

Based on the current study, the author concludes that regardless of the precise qMRI outcome measure used, it is certainly worth considering a trial of ≥24 months if using qMRI as the primary outcome measure, though the incremental benefit on qMRI responsiveness of extending duration beyond 24 months is not established. The decision regarding trial duration is of course multifaceted, but these findings are useful given the large number of compounds which need to be put through rigorous trials in quick succession.

Table 3-27 – Standardised response mean for summary muscle measures

	Standardised response mean		
Measure	**Timepoint two**	**Timepoint three**	**Timepoint four**
Both Calves	1.04	1.01	1.07
R. Calf	0.91	1.20	1.07
L. Calf	0.99	0.85	0.97
R. Anterior	0.73	1.32	1.15
L. Anterior	0.49	0.74	0.76
R. Triceps	0.88	1.10	0.91
L. Triceps	0.41	0.74	0.75

SRM=standardised response mean. R.=right, L.=left. Cells are highlighted in blue if SRM associated with significant change in fat fraction

3.6.2.4.2 Single or both lower limbs

When considering how much muscle to segment for longitudinal analysis, one may imagine that the larger volume of muscle afforded by segmenting both calves instead of a single calf may reduce measurement error by increasing muscle 'sample size', thus reducing standard deviation due to variation and increasing responsiveness. This study has demonstrated that this is not obviously the case. Across all timepoints, difference in responsiveness between using both legs versus one leg was marginal, and indeed at timepoints three and four, single leg FF proved equally if not more responsive than combined bilateral calf FF (Table 3-28).

Table 3-28 – Comparison of SRM for summary measures at all timepoints

	Standardised response mean		
Measure	**Timepoint two**	**Timepoint three**	**Timepoint four**
Both Calves	1.04	1.01	1.07
R. Calf	0.91	1.20	1.07
L. Calf	0.99	0.85	0.97
R. Anterior	0.73	1.32	1.15
L. Anterior	0.49	0.74	0.76
R. Triceps	0.88	1.10	0.91
L. Triceps	0.41	0.74	0.75

Rows are colour coded for visual aid. Orange=triceps surae group, blue=anterior group, green=whole calf. R=right, L=left

Balancing this, is the finding of right-left asymmetry in responsiveness of some summary measures, which is unrelated to baseline FF. This is also present in individual muscles (Table 3-19). We have observed the right leg to show larger responsiveness in almost all muscles, though this is not likely to be of major significance.

In the current climate of manual muscle segmentation, the author suggests single calf segmentation in CMT1A, unless marked clinical asymmetry is apparent.

3.6.2.4.3 Single muscle or composite group of muscles

This study demonstrates that if trial duration is to be short (~12 months), FF of a combination of muscles (combined bilateral calf, right calf of left calf) should be used as the qMRI outcome measure. There was only a single individual muscle with significant FF change at timepoint two, with only small/moderate SRM. If study duration is extended ≥24 months however, although still favoured, a combination of muscles is less clearly more responsive than individual muscles. Summary muscle

SRM was between 0.85 and 1.20 at timepoint three, and between 0.97 and 1.07 at timepoint four, whereas individual muscle SRM was between 0.63 and 1.03 at timepoint three and 0.53 and 1.07 and timepoint four.

If a single or several muscles are to be chosen, the question is which one/s? The muscle should show significant change over time, with moderate FF (a baseline FF which is too high may predict a lower therapeutic response in trials, and makes reliable ROI placement more difficult) and with little variability. This study demonstrates that certain muscles show relatively lower responsiveness (e.g. medial and lateral gastrocnemius) due to small mean change relative to standard deviation, and should probably be avoided in isolation, however these finer points of individual muscle selection need further corroboration. The main finding from this study is that almost all individual muscles have similar baseline FF, show significant FF change and demonstrate at least moderate responsiveness at prolonged follow up, though a combination of muscles is still favoured due to generally lower p values and less variability in results.

3.6.2.4.4 Stratification of patients based on fat fraction at baseline

Across the duration of this longitudinal study, the largest SRM were seen when patients were stratified according to baseline FF. Patients with intermediate baseline FF (6-4% to 60%) showed FF change of 1.8 ± 1.0%/year with SRM of 1.75 at timepoint two, 1.5 ± 1.1%/year with SRM of 1.40 at timepoint three, and 1.2 ± 0.7%/year with SRM of 1.63 at timepoint four. When patients with 'normal' FF were excluded from analysis, FF change at timepoint two was 1.9 ± 0.9%/year with SRM of 2.05, the largest in the study by any outcome measure at any timepoint.

This finding of maximum progression in intermediately fat infiltrated muscles has been noted several times previously in other neuromuscular diseases, for example in a cohort of patients with FSHD, Andersen et al. noted that progression in absolute FF was less in all muscles examined with FF >60% (Andersen *et al.*, 2017). Similarly, in CMT1A, we have demonstrated this concept several years ago in a cohort of patients from two different sites (Morrow et al., 2018).

This then raises the question of a composite qMRI derived outcome measure, one in which each patient is assessed at the anatomical point in the lower limbs where most FF change is expected over trial duration. In the cohort studied in this thesis, CMT1A patients with 'normal' baseline calf FF made up ~50% of participants, and so this 50% with low FF are essentially a burden on the overall responsiveness of qMRI FF, not only by reducing the mean overall change, but also increasing the standard deviation

of change. Harnessing maximal FF change in this ~50% of patients would be expected to have the same effect as analysing the intermediate group alone, whilst still allowing universal participation and ensuring external validity of results.

It should also be noted that even in analysis of the subgroup of patients with 'normal' baseline calf FF, qMRI responsiveness was greatly improved compared with whole group analysis, despite much lower FF change. FF change and SRM for the baseline 'normal' CMT1A group at timepoints three and four were 0.3 ± 0.2%/year, SRM=1.42 and 0.3 ± 0.2%/year, SRM=1.46, compared with 0.9 ± 0.9%/year, SRM=1.01 and 0.7 ± 0.6, SRM=1.07 in the whole group. This is of some significance, as it may be that future treatments are most effective in muscles before they gather a moderate FF, in which case it would be ideal to select patients with low baseline FF, in whom even though the expected change is small, homogenising of that change will lead to improved responsiveness (albeit not as responsive as seen in the moderate baseline fat fraction group).

The concept of a composite qMRI outcome measure is further addressed in Chapter 5.

3.6.2.5 Other outcome measures

T2 time showed significant change in 'normal' CMT1A muscles, and although responsiveness was low in this study, T2 should be further examined as a biomarker in early disease. Finally, when considering composite outcome measures in CMT1A; in this study, CMTES and CMTES-LL showed large responsiveness over longer trial duration and should be included alongside qMRI measures in any trial of prolonged duration.

3.6.2.6 Final considerations

3.6.2.6.1 Anatomical

The current study uses axial MRI slices which are 10mm thick with 10mm gap between slices. One wonders whether the distance between slices has any bearing on longitudinal results. With only ten axial slices to analyse in this protocol, the longitudinal imaging block may be positioned in such a way that slices are not perfectly aligned longitudinally, with up to 10mm craniocaudal difference in the central point of each slice. If there are even small changes in muscle FF between slices, this may affect analysis. This question of muscle fat gradients is addressed in detail in Chapter 4 of this thesis.

3.6.2.6.2 Single or multiple axial slices

A vital question which is addressed in detail in Chapter 4 of this thesis, is whether segmentation of several axial slices instead of a single slice improves responsiveness of qMRI. Intuitively this seems possible given that as the volume or area of muscle analysed increases, relative measurement error should reduce. Analysis of multiple slices may also account for heterogeneity in disease progression.

3.6.2.6.3 Translation of findings to interventional studies in NMD

Although intuitive, it is noted that the demonstrated large SRM for qMRI in this natural history study would only be applicable to an interventional study, if any medication used was expected to affect lower limb FF.

3.7 Conclusion

Over extended follow up, qMRI by the 3-point Dixon method measured significant ongoing fat accumulation with large responsiveness in calf muscles of patients with CMT1A. Significant change was also measured in calf RMA and CMT examination score both of which also showed large responsiveness. Over five years, qMRI determined FF assessment demonstrated higher SRM than RMA, in keeping with greater responsiveness of FF to disease progression. Excellent longitudinal qMRI FF validity is demonstrated by strong and statistically significant correlation with change in CMT examination score and RMA.

Quantitative MRI measure responsiveness can be significantly increased by careful attention to study design and outcome measure selection. Important factors include optimising trial duration (qMRI responsiveness was shown to plateau after three years of follow up), giving careful consideration to the specific combination of calf muscles to be analysed depending on trial duration, and most importantly by severity based patient stratification. Accurate slice selection, number of axial slices analysed and baseline severity based slice selection may all be important factors to consider and are examined further in Chapters 4 and 5.

Results of this study strongly support the application of qMRI FF alongside functional assessments as a primary outcome measure for clinical trials in CMT1A and similar neuromuscular condition.

4 Charcot-Marie-Tooth type 1A – calf muscle fat fraction gradient analysis

4.1 Introduction

Charcot-Marie-Tooth disease type 1A (CMT1A) is a length-dependent neuropathy, meaning that distal lower limb sensation and movement are affected first, with upper limbs only becoming affected when symptoms and signs reach ~knee level. As may be expected both from clinical practice (Reilly, Murphy and Laura, 2011), and in keeping with clinicopathological studies (Berciano et al., 2000), qualitative MRI assessment of lower limb muscles in patients with CMT1A reveals that intrinsic foot muscles are affected in early disease (Chung et al., 2008; Gallardo et al., 2006; Gallardo et al., 2009; Pelayo-Negro et al., 2014), often in clinically asymptomatic patients, after which variable degrees of calf muscle fatty atrophy develops. Thigh muscles are commonly normal despite often quite extensive MR involvement of the calf musculature (Gallardo et al., 2009), though subclinical thigh muscle involvement may be seen in advanced disease prior to the development of overt weakness, as was reported in a large pedigree (Pelayo-Negro et al., 2014) and more recently in a small cohort of patients (Kim et al., 2019). Beyond CMT1A, qualitative proximal-distal muscle fat gradients have been reported in calf muscles of patients with CMT2A (Chung et al., 2008), CMT2F (Gaeta et al., 2012), CMT2M (Gallardo et al., 2008), and CMT2J (Gallardo et al., 2009).

More recently, several groups have described quantitative lower limb fat fraction (FF) patterns in two inherited muscle diseases: facioscapulohumeral dystrophy (FSHD) (Dahlqvist et al., 2014) and Duchenne Muscular Dystrophy (DMD) (Chan and Liu, 2002), however muscle fat gradients have not been examined quantitatively in patients with CMT.

If present with a significant effect size in CMT1A, proximal-distal muscle fat gradients would have serious implications for imaging test-retest reproducibility, and hence responsiveness, of quantitative MRI (qMRI) as an outcome measure. This problem could be additionally exacerbated for studies in which MRI block positioning is based on surface landmarks, which can result in large inter-scan difference in longitudinal slice positioning.

In this chapter, proximal-distal calf muscle fat gradients in CMT1A and controls are examined both cross-sectionally and longitudinally in time by qMRI. The results are used to further refine qMRI as an outcome measure.

4.2 Background literature

Here I review all published literature, both quantitative and qualitative relating to lower limb muscle fat gradients in neuromuscular disease before presenting our findings.

4.2.1 Qualitative studies

In CMT1A, Berciano et al. examined a large pedigree of eighteen patients, of which four (three mildly affected, and one severely affected – by CMTNS) underwent thigh and calf T1 and T2 weighted MRI at 1.5 Tesla. One mildly affected patient also had foot imaging. Imaging of all three patients with a mild phenotype showed subtle and subclinical fatty infiltration of anterolateral calf muscles, and minimal thigh muscle involvement in one patient. There was extensive fatty atrophy of intrinsic foot muscles in the single patient who underwent foot imaging. In the patient with severe phenotype, imaging showed extensive fatty atrophy of all foot and calf muscles, posteromedial thigh muscle compartments, and internal and external hip rotator muscles (Berciano et al., 2006). The authors concluded the existence of a distal-proximal gradient, in that the foot was affected more than calf, which was affected more than the thigh, but did not comment on a gradient of involvement within the calf itself.

Stilwell et al. examined a series of 23 patients with CMT1A with T1-weighted MRI at 1.5 Tesla, analysing three axial T1w slices in each calf: one each at proximal, mid and distal calf. (Stilwell, Kilcoyne and Sherman, 1995). The authors found a distal-predominant qualitative muscle fat gradient in all calf muscle compartments except for the deep posterior compartment in which there was no appreciable qualitative gradient despite radiological involvement.

Gallardo and colleagues examined a cohort of eleven patients with CMT1A aged between 8y and 61y (median age 24y) with thigh and calf MRI (T1 weighted imaging and T2 weighted spin echo sequences). Seven patients also had foot imaging. The authors reported that in six patients with mild disease (normal examination and mild CMTNS), there was isolated intrinsic foot muscle fat involvement mainly involving the lumbricals, with preservation of calf muscles. In the remaining five patients who had more severe disease clinically including variable degrees of foot drop, in addition to intrinsic foot muscle involvement, MRI showed variable and distally accentuated fatty infiltration of the lateral, anterior and superficial posterior calf muscle compartments with relative sparing of the deep posterior compartment. The authors noted that the most distal foot muscles were most severely affected in keeping with length-

dependent degeneration of motor axons, and that foot muscle involvement was present in asymptomatic muscles (Gallardo et al., 2008).

In their two-year natural history study, Pelayo-Negro et al. assessed a cohort of 14 patients with CMT1A and 14 matched controls, with T1 and T2 weighted lower limb MRI at 1.5 Tesla which revealed foot muscle atrophy in all patients, with variable degrees of calf atrophy – predominantly anterior with only mild posterior compartment involvement – mostly medial gastrocnemius and soleus with minimal involvement of tibialis posterior and lateral gastrocnemius (Pelayo-Negro et al., 2014). The authors noted that there was greater fatty infiltration distally in both the calf and thigh muscles, in addition to greater fatty infiltration in the feet than calves, and calves than thighs.

Finally, Gallardo et al. examined thigh and calf by T1 weighted MRI at 1.5 Tesla, in a pedigree of three CMT2B patients, reporting a proximal-to-distal gradient of fat infiltration in 19/66 (29%) muscles. In 53% of muscles no gradient was observed and interestingly, in the remaining muscles (18%), the gradient was reversed (Gallardo et al., 2008).

4.2.2 Quantitative studies

In their natural history study of 41 patients with FSHD, Janssen et al. (Janssen et al., 2014) studied the natural history of FSHD with lower limb MRI determined fat quantification (multiple spin-echo) at 3 Tesla, in a cohort of 41 patients, 14 of whom underwent repeat imaging after four months. The authors found that although most lower limb muscles had either high or low FF, those with intermediate fatty infiltration had a 'fat gradient', with more severe fatty infiltration distally with mean ± SEM of 7.0 ± 1%/cm, ($p < 0.001$). This value was higher compared with muscles that were normal or mildly fat infiltrated (1.3 ± 0.3%/cm) and those that were heavily infiltrated by fat (1.1 ± 0.1%/cm) This corroborates earlier work by Kan et al. who examined seven patients with FSHD and seven matched controls with multiple spin echo MRI at 3 Tesla. The authors found that in five patients, there was non-uniform muscle fat infiltration even within a single muscle, with predominantly distal involvement, and a FF change of up to 37% in some muscles, though it was not stated over what distance (Kan et al., 2009).

Lareau-Trudel et al. examined both thighs and calves with qMRI at 1.5 Tesla in a cohort of 35 patients with FSHD. A gradient of intramuscular FF was observed between the proximal and the distal slices: mean intramuscular FF was 11.6 ± 8.6% (proximal) and 27.2 ± 12.4% (distal), however there was no assessment between these points (Lareau-Trudel et al., 2015).

In DMD, Hoojimans et al. examined muscle fat gradients using the Dixon technique at 3 Tesla in 22 patients and 12 matched controls. Interestingly, the group noted higher FF in distal and proximal muscle segments compared to the muscle belly in all DMD muscles ($p < 0.001$). The group felt this may in part be due to higher mechanical strain at muscle end regions. Importantly, the authors reported that a shift of 15 mm in slice position resulted in a difference in mean FF of ~1-2%, and up to 12% ($p < 0.01$) (Hooijmans et al., 2017).

4.3 Aims

The primary aim of this study was to examine cross-sectional proximal-distal muscle fat gradients in calf muscles of patients with Charcot-Marie-Tooth type 1A by quantitative MRI, and detail the effects of these gradients on longitudinal analysis. These results are then translated into methods by which internal responsiveness of quantitative MRI can be further enhanced.

4.4 Results

4.4.1 Participants

Thirteen patients with CMT1A and eight matched controls were assessed at baseline and after 12 months. There were no significant differences in age, gender, height, weight or body mass index between the groups. Baseline CMT1A and control group demographic and clinical details are summarised in Table 4-1.

Table 4-1 – Baseline demographic details

	CMT1A	Control	p
Demographic details			
Gender	7M, 6F	6M, 2F	0.36
Age (years)	46	53	0.31
Height (cm)	1.65	1.73	0.14
Weight (kg)	67.2	76.8	0.10
BMI (kg/m2)	24.8	25.6	0.67
Clinical characteristics			
Age at onset (years)	7.2 ± 18.4	NA	NA
CMTES (0-28)	8.2 ± 5.8	NA	NA

For gender, the p value is for the χ^2 test. For other tests, p value is for the two-tailed t-test

4.4.2 Cross-sectional results

All cross-sectional results for this thesis have been calculated by method c (Figure 2-2).

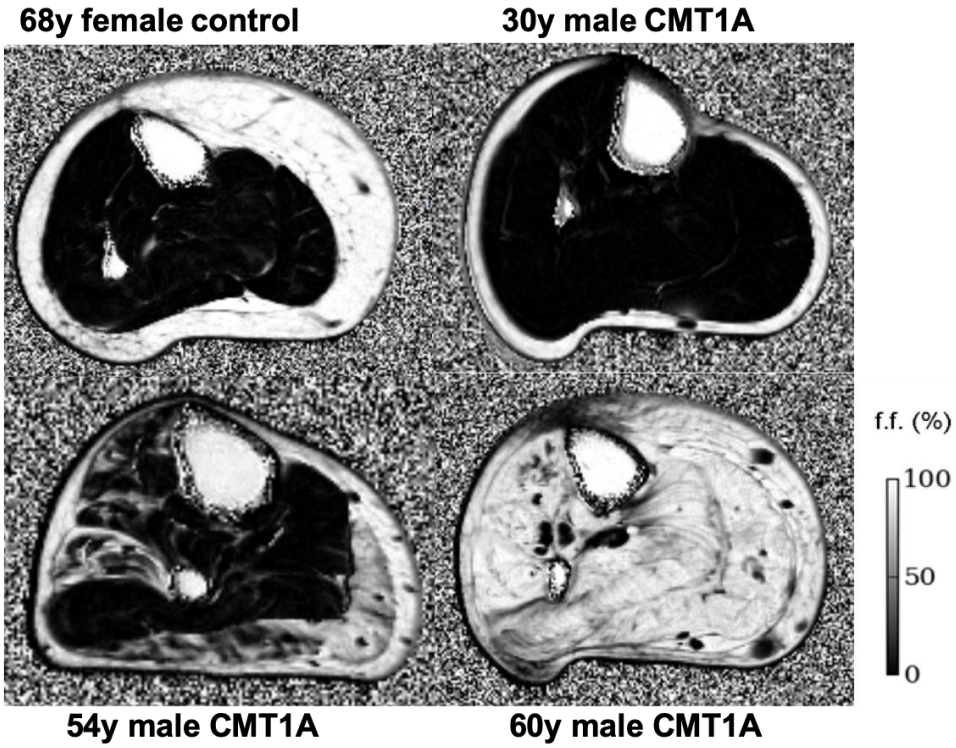
4.4.2.1 MRI data

Representative 3-point Dixon FF maps in a healthy control and three CMT1A patients with differing severities of disease are shown in Figure 4-1. Cross-sectional right calf FF results in CMT1A and controls are given in Table 4-2.

4.4.2.1.1 Calf fat fraction

In the CMT1A group at baseline, whole calf mean FF ± standard deviation was 20.4 ± 28.4% versus 2.1 ± 1.3% in controls (p=0.03). As expected, all muscles in CMT1A patients were more heavily fat infiltrated compared with controls – this difference reached statistical significance in all combined muscle groups (peroneal, tibial and whole calf), and in the individual peroneal innervated muscles: tibialis anterior, extensor hallucis longus and peroneus longus. In CMT1A patients, peroneal innervated muscles were more heavily fat infiltrated (26.4 ± 29.0%) than tibial innervated muscles (18.8 ± 28.3%) though this difference was not statistically significant in this cohort (p=0.51).

Figure 4-1 – Representative axial MRI Dixon fat maps in CMT1A and controls



Axial images are acquired at right mid-calf level in a healthy control, and three patients with differing severities of CMT1A. With disease progression, the normal muscle which appears black is replaced by fat which appears white (see scale)

Table 4-2 – Baseline mean fat fraction in CMT1A and controls

	CMT1A	Control	
Muscle	**Mean fat fraction ± s.d**		**p**
TA	22.6 ± 30.2	0.9 ± 0.5	0.001
EHL	23.8 ± 30.8	2.0 ± 1.4	0.02
PL	33.5 ± 31.6	2.3 ± 1.5	0.005
MG	26.9 ± 31.9	1.8 ± 0.9	0.08
LG	22.6 ± 33.0	1.5 ± 1.2	0.12
Sol	16.8 ± 28.1	2.9 ± 2.4	0.14
PT	16.4 ± 27.4	1.7 ± 1.3	0.06
Peroneal	26.4 ± 29.0	1.5 ± 0.9	<0.0001
Tibial	18.8 ± 28.3	2.3 ± 1.5	0.045
Whole calf	20.4 ± 28.3	2.1 ± 1.3	0.03

Values are mean fat fraction (%) ± standard deviation. MRI values are derived from whole muscle region-of-interest means by method c. The p value is for Mann Whitney U test (CMT1A vs control). TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, MG=medial gastrocnemius, LG=lateral gastrocnemius, Sol=soleus, PT=posterior tibial. Rows highlighted in blue if p<0.05

4.4.2.1.2 Cross-sectional area and remaining muscle area

At baseline, there were significant differences in cross-sectional area (CSA) between CMT1A and control groups in all combined muscles. In whole calf, mean CSA was 38.7 ± 11.6cm^2 in CMT1A versus 56.6 ± 10.2cm^2 in control, (p=0.002). In peroneal innervated muscles, mean CSA was 9.9 ± 3.1cm^2 versus 15.6 ± 2.8cm^2 in controls, (p=0.0005), and in posterior tibial innervated muscles, CSA was 28.8 ± 8.8cm^2 versus 41.1 ± 7.7cm^2 in controls (p=0.004). CSA difference was also significant in some individual muscles (Table 4-3). There were also significant baseline differences in remaining muscle area (RMA) between CMT1A and control groups (Table 4-4).

Figure 4-2 illustrates progressive right calf fat atrophy in three CMT1A patients with similar body mass index and different severities of disease, compared with a healthy control. Once again, as in the larger CMT1A cohort, note is made of the dichotomy between significant fat infiltration without statistically significant CSA reduction in peroneus longus – consistent with a temporal disconnect between these two processes in this muscle (Table 4-2, Table 4-3).

These dual findings of fat infiltration alongside muscle atrophy are in keeping with progressive fatty atrophy in the lower limbs of this cohort of CMT1A patients.

Table 4-3 – Baseline cross-sectional area in CMT1A and controls

	CMT1A	Control	
Muscle	**mean CSA (cm^2) ± s.d**		**p**
TA	4.0 ± 1.7	7.7 ± 2.0	<0.0001
EHL	1.9 ± 0.6	2.8 ± 0.6	0.003
PL	4.0 ± 1.5	5.0 ± 1.4	0.12
LG	2.8 ± 1.8	3.5 ± 2.4	0.50
MG	6.8 ± 3.4	10.0 ± 2.5	0.02
Sol	15.8 ± 6.8	19.0 ± 4.5	0.25
PT	6.2 ± 2.2	11.5 ± 2.4	<0.0001
Peroneal	9.9 ± 3.1	15.6 ± 2.8	<0.0001
Tibial	28.8 ± 8.8	41.1 ± 7.7	0.004
Whole calf	38.7 ± 11.6	56.6 ± 10.2	0.002

Cross-sectional area is given in cm^2. p value is for the independent sample two tailed t-test with controls using method c. s.d=standard deviation, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior, peroneal=common peroneal innervated muscles, tibial=posterior tibial innervated muscles. Rows are highlighted in blue if p<0.05

Table 4-4 – Baseline remaining muscle area in CMT1A and controls

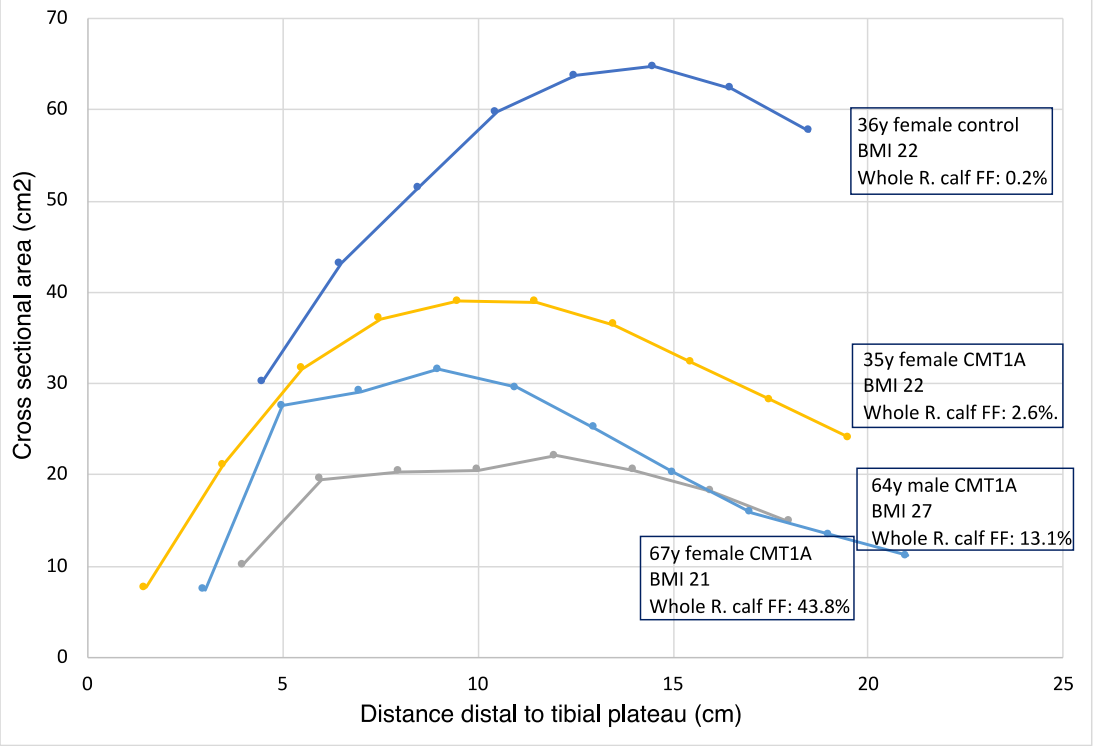
	CMT1A	Control	
Muscle	**mean RMA (cm^2) ± s.d**		**p**
TA	3.3 ± 2.3	7.7 ± 2.0	<0.0001
EHL	1.5 ± 0.9	2.8 ± 0.6	0.003
PL	2.9 ± 1.9	5.0 ± 1.4	0.02
LG	2.4 ± 2.1	3.5 ± 2.4	0.29
MG	5.5 ± 4.2	10.0 ± 2.5	0.01
Sol	13.7 ± 8.2	19.0 ± 4.5	0.15
PT	5.1 ± 2.7	11.5 ± 2.4	<0.0001
Peroneal	7.7 ± 4.5	15.3 ± 2.8	<0.0001
Tibial	26.7 ± 13.9	43.4 ± 8.1	<0.0001
Whole calf	34.5 ± 18.2	58.7 ± 10.7	0.003

Remaining muscle area is given in cm^2. p value is for the independent sample two tailed t-test with controls using method c. s.d=standard deviation, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior, peroneal=common peroneal innervated muscles, tibial=posterior tibial innervated muscles. Rows are highlighted in blue if p<0.05

4.4.2.1.3 MRI distribution of fat infiltration

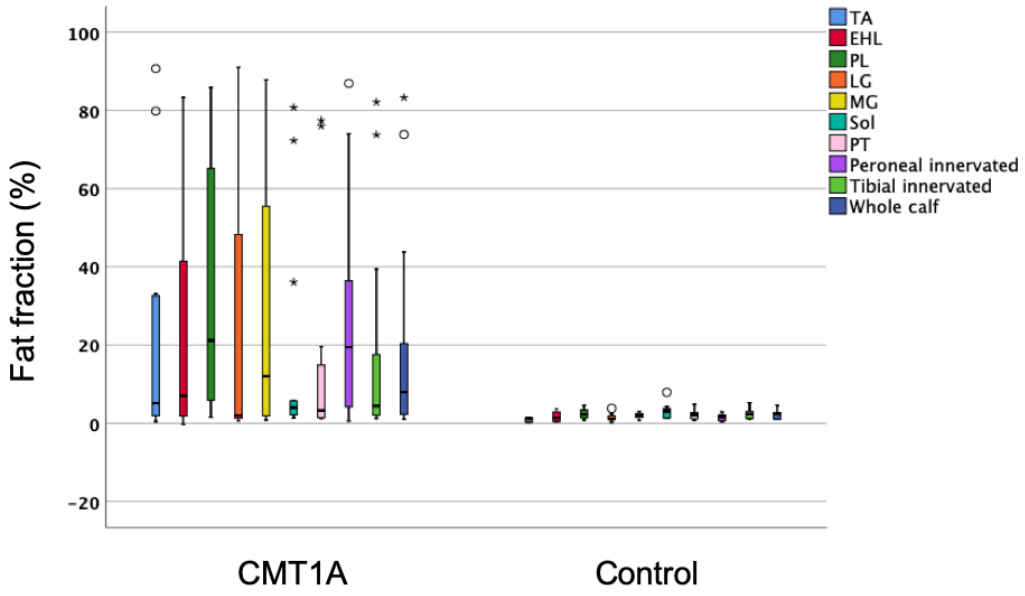
In CMT1A patients, the pattern of muscle involvement was variable (Figure 4-3) with skewed distribution. Peroneus longus and medial head of gastrocnemius were most heavily fat infiltrated (mean FF 33.5 ± 31.6% and 26.9 ± 33.0% respectively) and tibialis posterior and soleus muscles least affected (mean FF: 16.4 ± 27.4% and 16.8 ± 28.1% respectively), however there were no significant differences between mean individual muscle FF in CMT1A patients at baseline.

Figure 4-2 – Right calf cross-sectional area in three CMT1A patients and one control



Each line represents one participant. FF=fat fraction, R.=right

Figure 4-3 – Baseline fat fraction in individual and grouped muscles in CMT1A and controls

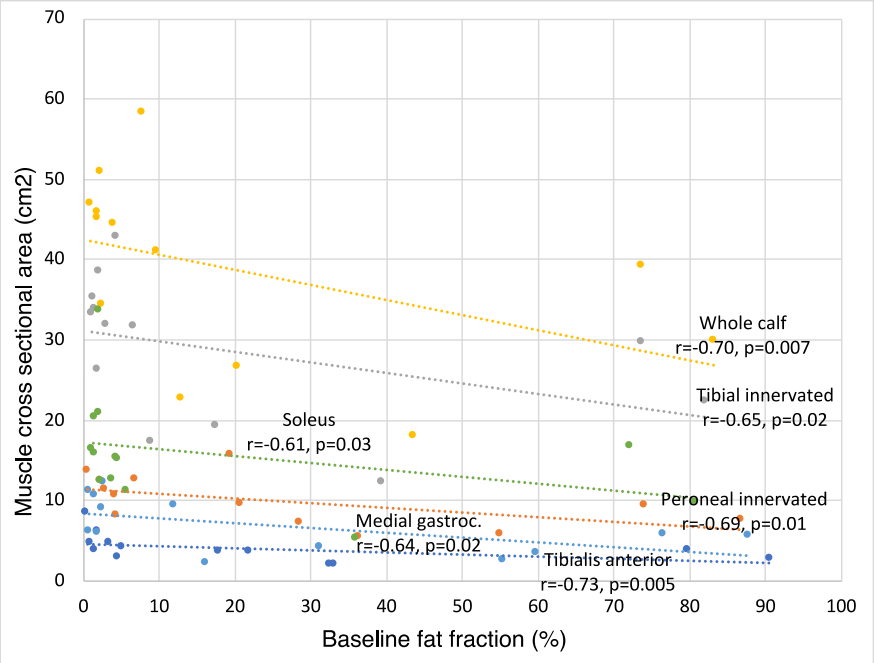


Boxes represent median fat fraction (%) and IQR, whiskers show range, and circles/asterisks are outliers. TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior, peroneal=common peroneal innervated muscles, tibial=posterior tibial innervated muscles

4.4.2.2 Cross-sectional correlation between MRI measures

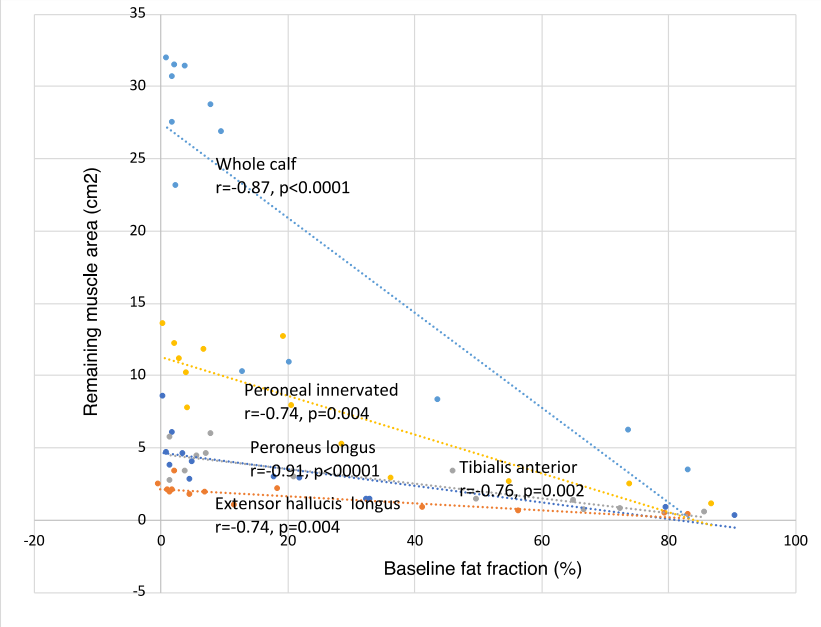
At baseline in CMT1A patients, there were strong negative correlations between baseline FF and CSA in all muscle combinations and some individual muscles (soleus, tibialis anterior and medial gastrocnemius muscles). Whole calf (r=-0.70, p=0.007), peroneal innervated group (r=-0.69, p=0.01), posterior tibial innervated group (r=-0.65, p=0.02). There were also strong negative correlations between baseline FF and baseline RMA across the spectrum of individual and combined muscles, except for medial head of gastrocnemius. In controls, there were no correlations between baseline FF and baseline RMA or CSA. Figure 4-4, Figure 4-5 and Figure 4-6 summarise these correlations graphically.

Figure 4-4 – Baseline fat fraction correlates with baseline cross-sectional area in CMT1A



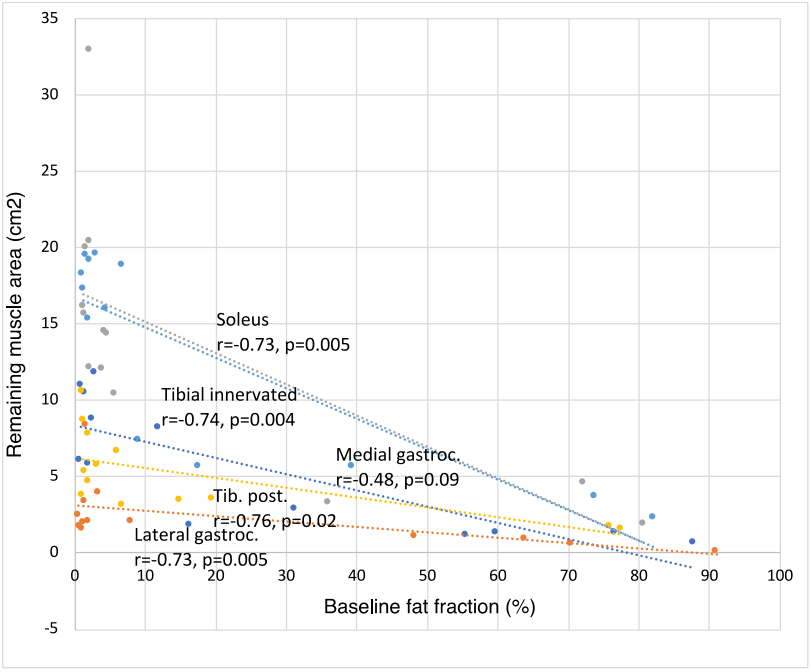
Graph is colour coded for muscle/group of muscles: blue=medial gastrocnemius, yellow=whole calf, navy blue=tibialis anterior, grey=tibial innervated group, green=soleus, orange=peroneal innervated group

Figure 4-5 – Baseline fat fraction correlates with baseline remaining muscle area in anterior compartment muscles in CMT1A



Graph is colour coded for muscle/group of muscles: blue=whole calf, yellow=peroneal innervated muscles, navy blue=tibialis anterior, grey=peroneus longus, orange=extensor hallucis longus

Figure 4-6 – Baseline fat fraction correlates with baseline remaining muscle area in posterior compartment muscles in CMT1A



Graph is colour coded for muscle/group of muscles: blue=posterior tibial innervated muscles combined, yellow=posterior tibial, navy blue=medial gastrocnemius, grey=soleus, orange=lateral gastrocnemius

4.4.2.3 Muscle stratification based on baseline fat fraction

For the control group at baseline, the FF 95th centile was 4.8% thus defining an upper threshold for 'normal' FF in the CMT1A cohort. Individual muscles and groups of muscles in the CMT1A cohort were thus divided into three groups based on this value: 'normal' if mean FF ≤4.8% across all slices, 'intermediate' if mean FF between 4.8% and 70%, and 'high' if mean FF >70%. In addition, given that no controls had a distal slice FF of >7.5%, muscles or groups of muscles were also defined as intermediate if the FF on any single slice was >7.5% (despite an overall mean FF of ≤4.8%). This was done to account for the length-dependent nature of the disease which may not have been adequately characterised by analysing central slices alone.

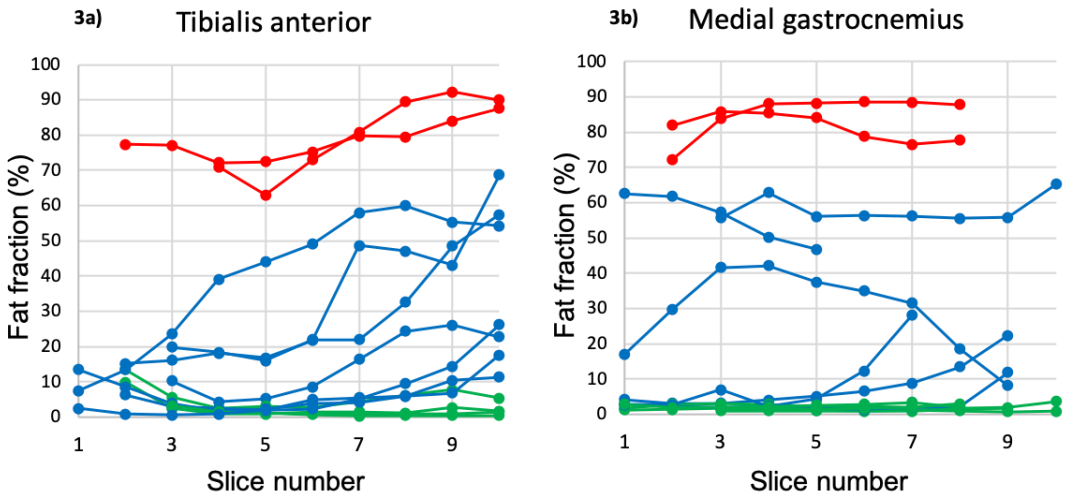
Of the 91 muscles examined at baseline in the CMT1A group, 38 (41.8%) had normal FF, 39 (42.9%) had intermediate FF, and 14 (15.4%) had high FF. 23/39 (59.0%) of peroneal innervated muscles had intermediate FF, compared with 16/52 (30.8%) of posterior tibial innervated muscles (p=0.007). Baseline mean FF for tibialis anterior muscle and medial head of gastrocnemius muscle in all CMT1A patients is shown in Figure 4-7. Mean FF for all muscles is summarised in Table 4-5.

Table 4-5 – Whole calf mean fat fraction in CMT1A

	Muscle									
Patient	TA	EHL	PL	MG	LG	Sol	PT	Peroneal	Tibial	Whole calf
1	22.0	18.4	21.2	12.0.	3.5	4.7	6.0	20.8	6.5	9.7
2	0.5	-0.3	1.5	0.8	1.2	1.3	1.2	0.5	1.1	1.0
3	1.5	1.5	7.3	0.9	0.6	1.7	1.1	4.2	1.4	1.9
4	4.7	4.8	4.1	1.9	0.8	2.3	2.0.	4.4	2.1	2.6
5	33.1	41.5	66.8	16.3	7.9	5.8	14.9	36.4	9.1	13.1
6	90.7	83.3	85.8	87.9	91.0.	80.7	77.4	86.9	82.4	83.3
7	79.8	79.5	65.1	76.6	70.4	72.2	76	74.1	75.4	73.8
8	5.1	7.0.	8.0.	1.5*	1.5	4.4	3.3	6.9	3.2	4.1*
9	3.5*	1.9*	1.5*	2.8	2.0.	2.2	1.4	2.3*	2.1	2.3
10	18.0	11.7	49.8	59.8	48.3	4.0*	6.8	28.7	17.5	20.4
11	0.9	1.2	5.9	2.5	1.3	1.5	1.3	3.0	1.7	2.0.
12	32.6	56.6	72.6	55.5	63.9	36.1	19.6	55.1	40.9	43.8
13	1.9*	2.2	46.1	31.3	1.6	2.1	2.1	19.4	4.3	7.9

mean calf muscle fat fraction (FF) calculated by method c in patients with CMT1A. Each row summarises fat fraction for an individual CMT1A patient. Values are colour coded for FF: green=normal FF (≤4.8%), blue=intermediate FF (4.8%≥70%) or distal slice >7.5%, red=high FF (>70%). TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, MG=medial gastrocnemius, LG=lateral gastrocnemius, Sol=soleus, PT=posterior tibial, peroneal=peroneal innervated muscles combined, tibial=posterior tibial innervated muscles combined*

Figure 4-7 – Baseline mean fat fraction in tibialis anterior (a) and medial gastrocnemius (b) muscles in CMT1A

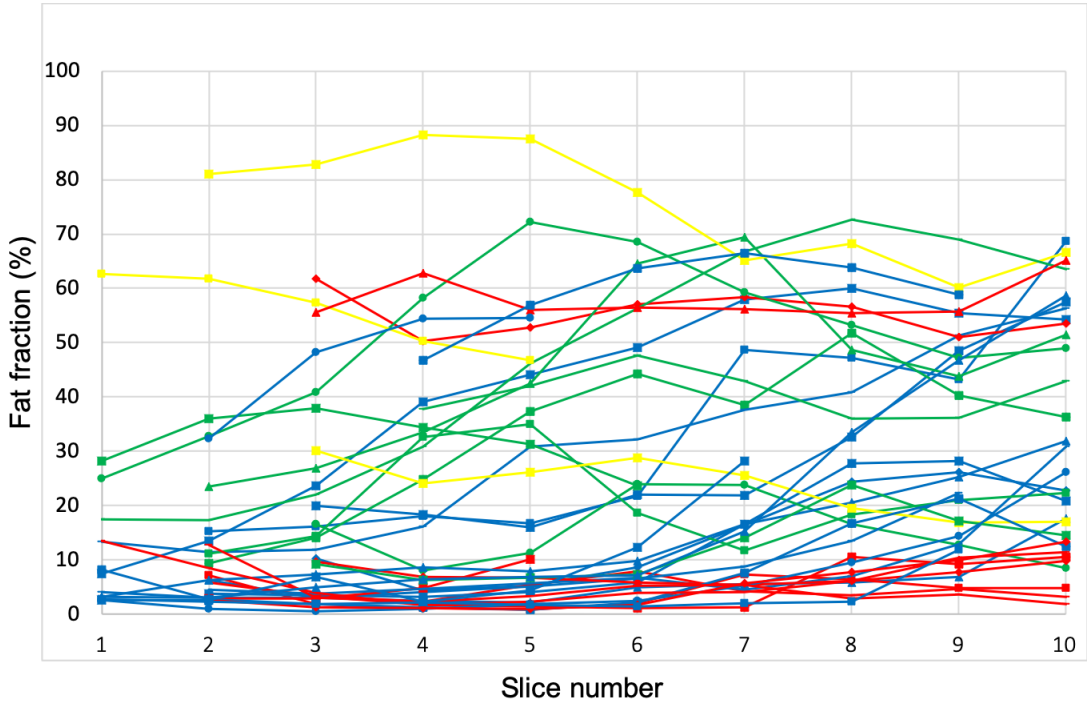


Each line represents a single patient. Lines are colour coded for whole muscle fat fraction (FF): High FF in red (>70%), Intermediate FF (4.8≥70%) or distal slice FF>7.5% in blue, normal FF (≤4.8%) in green

4.4.2.4 Cross-sectional gradient analysis

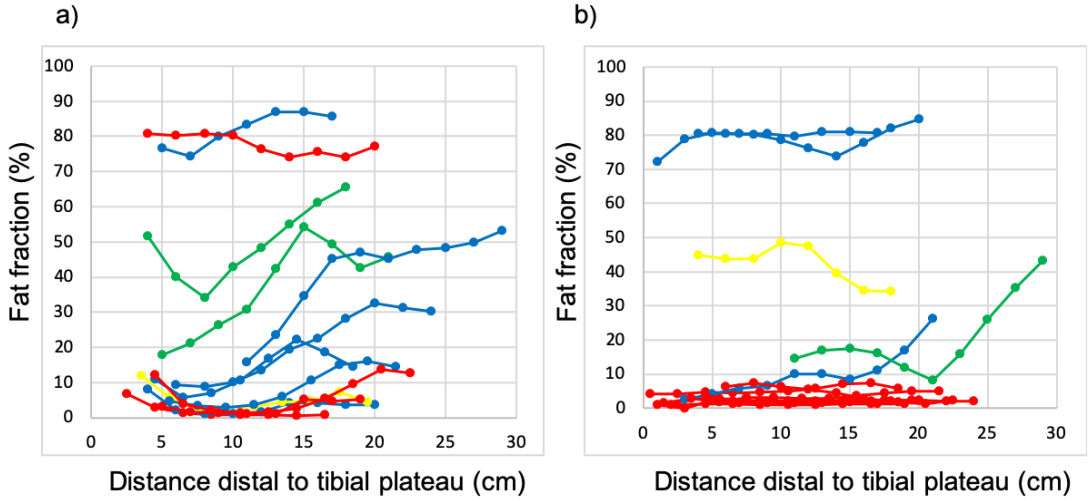
Proximal-distal right calf FF gradients were present in 41.8% (38/91) of CMT1A muscles overall, occurring most commonly in muscles with intermediate FF: 29/39 (74.4%), in which they were positive, negative or a combination of both – Figure 4-8. FF gradients were most frequent in the anterolateral compartment: 9/13 (69.2%) CMT1A patients had a FF gradient in peroneal innervated muscles, whereas 4/13 (30.8%) had a FF gradient in tibial innervated muscles (Table 4-6, Figure 4-9). These FF gradients were present in all muscles and could be very steep, with a maximum positive gradient of 12.8%/cm, and negative gradient of -10.6%/cm in the CMT1A group. Positive gradients >2%/cm were not seen in controls, although negative gradients were seen proximally (maximum FF gradient of -6.7%/cm). All FF gradient values in CMT1A and controls are summarised in Table 4-6. Representative proximal-distal right calf axial fat maps in a CMT1A patient and control are shown in Figure 4-10.

Figure 4-8 – Individual muscle fat fraction in all 'intermediate' muscles in CMT1A



Axial MRI slice is numbered from proximal (1) to distal (10). Each line (n=39) represents a single muscle in a single patient. Lines are colour coded for gradient: Blue=positive gradient (n=17), yellow=negative gradient (n=3), green=both positive and negative gradient (n=9), red=no gradient (n=10)

Figure 4-9 – Baseline calf fat fraction in peroneal (a) and tibial (b) innervated muscles in CMT1A



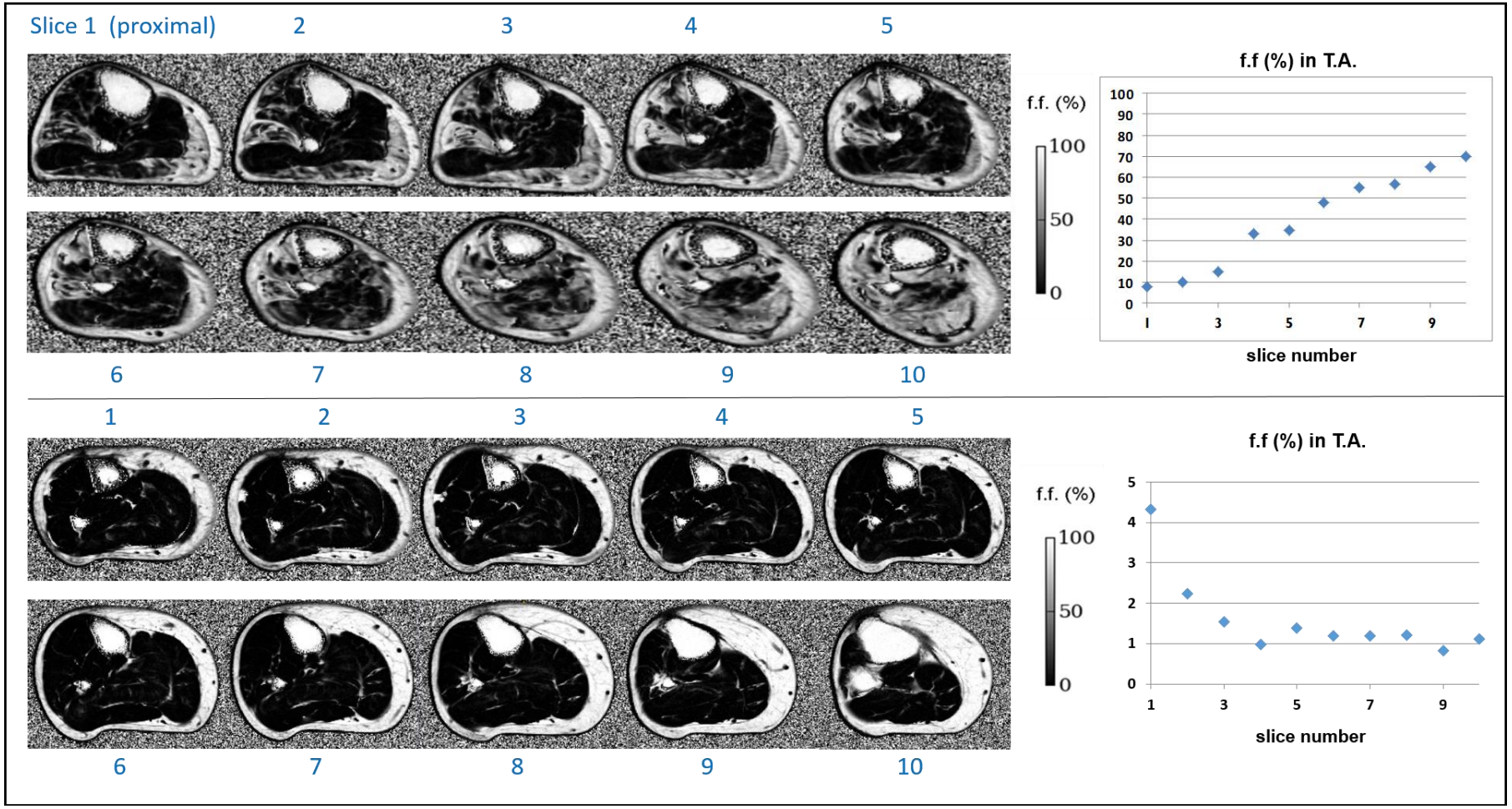
Each line represents a single patient. Lines are colour coded for gradient: blue=positive gradient, yellow=negative gradient, green=both positive and negative gradient, red=no gradient

Table 4-6 – Baseline distribution of muscle fat fraction gradients in CMT1A and controls

	CMT1A					Control				
	Positive gradient		Negative gradient		No gradient	Positive gradient		Negative gradient		No gradient
Muscle	n (%)	max.(%/cm)	n (%)	max.(%/cm)	n (%)	n (%)	max.(%/cm)	n (%)	max.(%/cm)	n (%)
TA	8 (61.5)	12.8	0	-3.0	5 (38.5)	0	0.8	2 (25.0)	-5.1	6 (75.0)
EHL	5 (38.5)	7.4	2 (15.4)	-5.8	8 (61.5)	0	1.1	1 (12.5)	-5.1	7 (87.5)
PL	6 (46.2)	11.0	6 (46.2)	-10.6	6 (46.2)	0	0.9	1 (12.5)	-6.7	7 (87.5)
LG	3 (23.1)	3.0	0	-4.5	10 (76.9)	0	2.4	0	-2.1	8 (100)
MG	5 (38.5)	8.0	2 (15.4)	-3.8	7 (53.8)	0	2.3	0	-0.6	8 (100)
Sol	5 (38.5)	9.1	1 (7.7)	-3.4	8 (61.5)	0	1.0	0	-1.2	8 (100)
PT	2 (15.4)	9.1	3 (23.1)	-8.2	9 (69.2)	0	0.9	0	-2.1	8 (100)
Peroneal	9 (69.2)	5.6	4 (30.8)	-5.9	3 (23.1)	0	1.3	3 (37.5)	-5.9	8 (100)
Tibial	4 (30.8)	5.1	2 (15.4)	-4.0	8 (61.5)	0	1.2	0	-2.0	8 (100)
Whole calf	2 (15.4)	4.0	0	-2.0	11 (84.6)	0	1.2	0	-1.5	8 (100)

Maximum gradient refers to the maximum gradient in that muscle across all participants. TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, MG=medial gastrocnemius, LG=lateral gastrocnemius, Sol=soleus, PT=posterior tibial

Figure 4-10 – Lower limb fat gradients in CMT1A and controls



Each panel of ten axial 3D Dixon MR images is from a single participant (CMT1A top, control bottom): slice 1=proximal, to slice 10=distal right calf. Adjacent graphs plot fat fraction in tibialis anterior muscle against slice number. f.f.= fat fraction. T.A= tibialis anterior

4.4.3 Longitudinal results

Repeat assessments were performed in all participants after a mean of 387.9 (± 34.8) days. All follow-up value changes were annualised for precise analysis.

4.4.3.1 Fat fraction change

In CMT1A patients at follow up, mean right calf FF increased significantly by all methods of analysis except method a, increasing by (mean ± s.d.): 2.4 ± 2.6, 2.6 ± 3.0, 2.1 ± 2.2 and 1.8 ± 2.0 percentage units by method b, c, d and e respectively. In the peroneal groups, FF change was significant by methods b, c, d and e: FF change in peroneal group: 3.0 ± 6.3%, 3.8 ± 4.2%, 2.6 ± 2.0% and 2.4 ± 1.7% respectively for methods b, c, d and e. In the posterior tibial group FF change was significant: 1.9 ± 2.7%, 1.9 ± 2.3% and 1.7 ± 1.9% by methods c, d and e. In individual muscles, FF increased significantly over the follow-up period by the more sophisticated methods of slice selection. Of note, there was no significant change in FF over 12 months in any muscle or group of muscles when FF was calculated by method a, and only in TA, PL, peroneal group and whole calf by method b. In the control group, no significant FF change was seen over 12 months in any individual or combination of muscles by any method of slice selection. FF change in CMT1A and control by different methods of calculation are summarised in Table 4-7.

4.4.3.2 Cross-sectional area and remaining muscle area change

Whole calf RMA reduced by 2.4 ± 4.9cm^2 in the CMT1A cohort, though this was not a significant change when compared with the control group. There were no significant changes in RMA at 12 months in any group of muscles or summary measures except for soleus by method a: (RMA change -1.2 ± 1.9cm^2), though this was not a significant change when compared with controls (p=0.70). Similarly, there was no significant change in CSA in individual muscles or summary measures in CMT1A patients when compared with controls – by any method of calculation.

Table 4-7 – Change in mean fat fraction in CMT1A patients and controls by different methods of calculation

	Method														
	a					b					c				
	CMT1A	Control				CMT1A	Control				CMT1A	Control			
Muscle	mean FF change ± sd (95% CI)		p1	p2	SRM	mean FF change ± sd (95% CI)		p1	p2	SRM	mean FF change ± sd (95% CI)		p1	p2	SRM
TA	4.3 ± 13.1 (2.9 to 11.4)	1.9 ± 4.0 (-1.4 to 5.2)	0.06	0.60	0.32	3.8 ± 6.3 (0.4 to 7.2)	0.1 ± 0.5 (-0.3 to 0.5)	0.01	0.02	0.60	2.5 ± 4.6 (-0.0 to 5.0)	0.1 ± 0.3 (-0.3 to 0.5)	0.005	0.01	0.54
EHL	1.2 ± 5.6 (-1.9 to 4.2)	1.1 ± 2.0 (-0.5 to 2.8)	0.92	0.47	0.21	3.1 ± 4.1 (0.9 to 5.3)	0.1 ± 0.9 (-0.7 to 0.9)	0.02	0.09	0.76	2.9 ± 3.3 (1.1 to 4.7)	0.1 ± 0.7 (-0.4 to 0.7)	0.005	0.04	0.88
PL	3.8 ± 10.6 (-1.9 to 9.6)	0.9 ± 1.1 (-0.1 to 1.8)	0.02	0.46	0.36	4.2 ± 4.6 (1.7 to 6.7)	0.2 ± 0.6 (-0.3 to 0.7)	0.002	0.006	0.92	3.1 ± 2.4 (1.8 to 4.4)	0.2 ± 0.6 (-0.3 to 0.7)	0.001	<0.0001	1.29
LG	2.1 ± 3.8 (0.01 to 4.2)	0.3 ± 1.2 (-0.7 to 1.3)	0.07	0.045	0.55	0.7 ± 2.3 (-0.6 to 2.0)	0.1 ± 0.9 (-0.6 to 0.7)	0.32	0.01	0.29	1.4 ± 2.3 (0.2 to 2.7)	0.0 ± 0.7 (-0.5 to 0.5)	0.04	0.03	0.63
MG	3.0 ± 3.4 (1.2-4.8)	0.4 ± 0.7 (-0.3 to 1.0)	0.004	0.21	0.89	3.3 ± 4.9 (0.6 to 5.9)	0.3 ± 0.5 (-0.2 to 0.7)	0.003	0.30	0.67	3.3 ± 4.1 (1.1 to 5.5)	0.1 ± 0.5 (-0.3 to 0.5)	0.007	0.09	0.81
Sol	0.5 ± 0.8 (0.07 to 0.9)	0.6 ± 1.2 (-0.4 to 1.6)	0.06	0.80.	0.63	0.9 ± 3.0; (-0.7 to 2.6)	0.2 ± 0.4; (-0.2 to 0.5)	0.13	0.41	0.31	0.9 ± 3.3 (-0.8 to 2.7)	0.1 ± 0.3 (-0.2 to 0.3)	0.12	0.24	0.29
PT	1.3 ± 2.4 (-0.02 to 2.6)	1.0 ± 1.7 (-0.5 to 2.4)	0.05	0.55	0.53	1.1 ± 1.9 (0.07 to 2.1)	0.4 ± 0.7 (-0.1 to 1.0)	0.03	0.41	0.58	1.5 ± 1.7 (0.6 to 2.5)	0.3 ± 0.5 (-0.1 to 0.7)	0.003	0.06	0.91
Peroneal	3.4 ± 9.5 (-1.7 to 8.6)	1.4 ± 2.6 (-0.8 to 3.6)	0.03	0.70	0.36	3.0 ± 6.3 (-0.5 to 6.4)	0.2 ± 0.5 (-0.3 to 0.6)	0.02	0.002	0.47	3.8 ± 4.2 (1.6 to 6.1)	0.1 ± 0.4 (-0.2 to 0.5)	0.001	<0.0001	0.92
Tibial	1.6 ± 2.8 (0.1 to 3.1)	0.6 ± 1.0 (-0.3 to 1.4)	0.004	0.27	0.59	1.8 ± 2.6 (0.4 to 3.2)	0.2 ± 0.5 (-0.2 to 0.6)	0.006	0.05	0.69	1.9 ± 2.7 (0.5 to 3.4)	0.1 ± 0.4 (-0.2 to 0.4)	0.007	0.03	0.72
Whole calf	2.3 ± 3.0 (0.6 to 3.9)	0.7 ± 1.3 (-0.4 to 1.7)	0.002	0.16	0.76	2.4 ± 2.6 (0.9 to 3.8)	0.2 ± 0.5 (-0.2 to 0.6)	0.002	0.003	0.90	2.6 ± 3.0 (0.9 to 4.2)	0.1 ± 0.4 (-0.2 to 0.4)	0.002	0.006	0.84

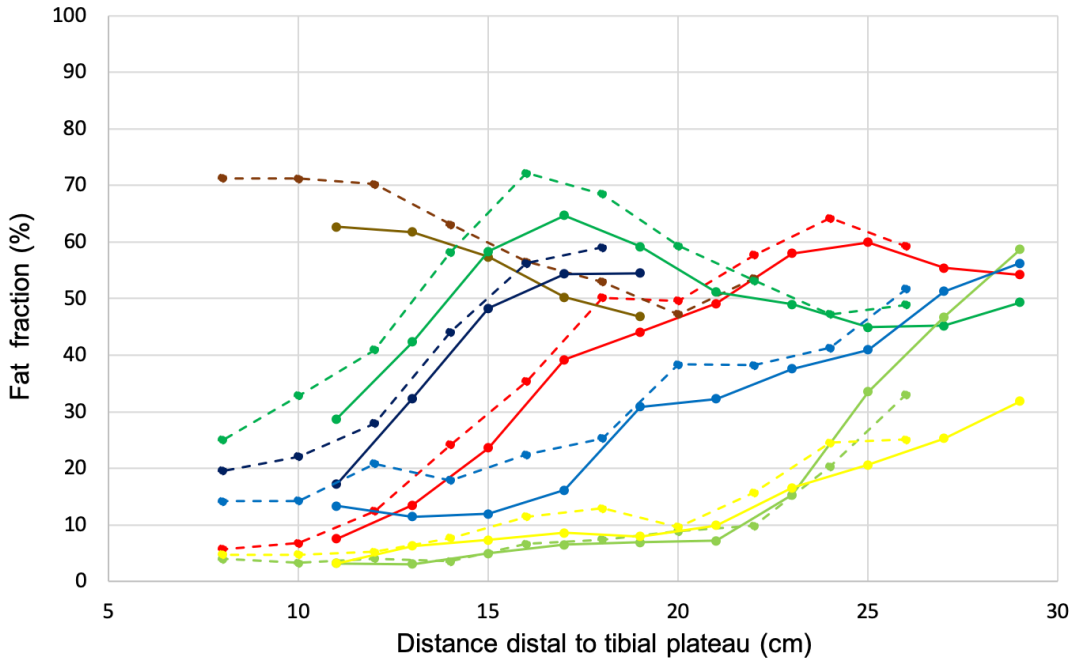
	d					e				
	CMT1A	Control				CMT1A	Control			
Muscle	mean FF change ± sd (95% CI)		p1	p2	SRM	mean FF change ± sd (95% CI)		p1	p2	SRM
TA	2.1 ± 3.2 (0.3 to 3.8)	0.0 ± 0.4 (-0.2 to 0.3)	0.001	0.003	0.65	2.2 ± 2.2 (1.0 to 3.4)	0.1 ± 0.3 (-0.2 to 2.4)	0.002	0.001	1.00
EHL	2.8 ±2.6 (1.4 to 4.2)	0.3 ± 0.4 (-0.1 to 0.6)	0.001	0.001	1.06	2.9 ± 2.5 (1.5 to 4.2)	0.3 ± 0.6 (-0.3 to 0.8)	0.001	0.001	1.17
PL	2.8 ± 2.5 (1.4 to 4.1)	0.0 ± 1.0 (-0.7 to 0.7)	0.002	0.001	1.10	2.3 ± 1.8 (1.3 to 3.2)	0.0 ± 0.6 (-0.5 to 0.5)	0.001	<0.0001	1.26
LG	1.5 ± 2.3 (0.2 to 2.7)	ND	0.04	-	0.65	1.4 ± 2.0 (0.3 to 2.4)	-0.1 ± 0.4 (-0.4 to 2.3)	0.03	0.005	0.68
MG	3.4 ± 3.8 (1.3 to 5.4)	1.1 ± 2.8 (-1.3 to 3.4)	0.003	0.045	0.88	2.2 ± 2.5 (0.9 to 3.6)	0.2 ± 0.4 (-0.1 to 0.5)	0.004	0.049	0.88
Sol	0.9 ± 2.4 (-0.4 to 2.2)	0.1 ± 0.3 (-0.1 to 0.4)	0.04	0.08	0.36	0.9 ± 1.6 (0.02 to 1.8)	0.1 ± 0.3 (-0.1 to 0.4)	0.02	0.12	0.55
PT	1.5 ± 1.6 (0.6 to 2.5)	0.3 ± 0.6 (-0.2 to 0.8)	0.002	0.01	0.95	1.7 ± 1.6 (0.83 to 2.6)	0.2 ± 0.5 (-0.2 to 0.5)	0.002	0.005	1.07
Peroneal	2.6 ± 2.0 (1.5 to 3.6)	0.1 ± 0.3 (-0.2 to 0.4)	0.001	<0.0001	1.29	2.4 ± 1.7 (1.5 to 3.4)	0.1 ± 0.4 (-0.3 to 0.4)	0.001	<0.0001	1.46
Tibial	1.9 ± 2.3 (0.7 to 3.2)	0.1 ± 0.3 (-0.1 to 0.3)	0.002	0.005	0.84	1.7 ± 1.9 (0.7 to 2.7)	0.2 ± 0.3 (-0.1 to 0.4)	0.002	0.003	0.89
Whole calf	2.1 ± 2.2 (0.9 to 4.2)	0.2 ± 0.3 (-0.1 to 0.5)	0.001	0.003	0.97	1.8 ± 2.0 (0.9 to 2.9)	0.1 ± 0.3 (-0.1 to 0.4)	0.001	0.001	0.91

Data are mean change in fat fraction (%) ± standard deviation (95% confidence interval), p1=paired t-test or Wilcoxon test as appropriate (CMT1A), p2=Mann Whitney U test or independent samples t-test as appropriate (CMT1A vs control), SRM=standardised response mean, FF=fat fraction. Rows are highlighted in blue when both t-tests are significant. ND=no data point

4.4.3.3 Longitudinal gradient analysis

At 12 months in CMT1A patients, there was no change in the overall number of muscles with FF gradients >2%/cm. Figure 4-11 demonstrates the change in FF over 12 months in a single patient with CMT1A.

Figure 4-11 – Fat fraction in all muscles at baseline and follow-up in one CMT1A patient



Solid line=baseline, dashed line=follow up. Lines are colour coded for muscle: tibialis anterior=red, Extensor hallucis longus=blue, peroneus longus=green, lateral gastrocnemius=black, medial gastrocnemius=brown, soleus=lime, deep group=yellow

4.4.4 Standardised response mean

In the CMT1A group, significant FF change was seen at 12 months in qMRI-determined FF by methods b, c, d and e. Although in EHL, PT and the peroneal innervated group, the SRM increased progressively with more sophisticated methods of FF calculation: for EHL, SRM=0.76, 0.88, 1.06 and 1,17 for methods b, c, d and e respectively; for PT, SRM=0.91, 0.95 and 1.07 for methods c, d and e, and for peroneal group SRM= 0.92, 1.29 and 1.46 by methods c, d and e respectively, an increase in SRM with increasing slice-wise FF calculation sophistication was not the rule. In whole calf, SRM=0.90, 0.84, 0.97 and 0.91 by methods b, c, d and e respectively, and for PL muscle maximum SRM was seen by method c (SRM=1.29). The highest SRM of 1.46, was seen in the peroneal innervated group FF, calculated by method e. FF change, p value, 95% confidence interval and SRM for all individual and groups of muscles in CMT1A and controls are summarised in Table 4-7.

4.4.5 Group homogenisation – an approach to improve responsiveness

When CMT1A patient individual muscles and muscle groups were stratified based on baseline FF into low, intermediate and high FF groups (Table 4-5), maximum FF change was seen in the intermediate group across essentially all individual muscles and muscle groups by all methods of calculation, compared with the low or high baseline FF groups. This increased mean FF change alongside reduction in standard deviation of the change, due to group homogenisation, resulted in increased SRMs. For example, SRM of whole calf FF by method c increased from 0.84 when all patients were included in analysis, to 1.31 when only those with intermediate baseline FF were included. Highest SRM was 3.50 in the tibialis posterior muscle by method e (previous SRM of 1.07 when all tibialis posterior muscles included in analysis).

It is also of interest that when those muscles with 'low' FF were analysed separately, although the mean change was smaller than seen in the intermediate group, the reduced standard deviation resulted in an overall improvement in responsiveness. As an example, the SRM for whole calf FF increased from 0.91 to 1.20 by method e. FF change based on baseline FF, by all methods of calculation is summarised in Table 4-8.

Table 4-8 – Mean fat fraction change stratified by baseline fat fraction in CMT1A

		TA			EHL			PL			LG			MG		
method	fat infiltration	mean change±sd	p	SRM	mean change±sd	p	SRM	mean change±sd	p	SRM	mean change±sd	p	SRM	mean change±sd	p	SRM
a	Low	0.7 ± 0.8	0.20	0.82	0.02 ± 0.6	0.69	0.03	1.8 ± 1.1	0.27	1.54	0.5 ± 0.8	0.05	0.63	0.3 ± 0.5	0.30	0.54
	Int.	5.7 ± 17.8	0.50	0.32	0.2 ± 4.4	0.92	-0.04	0.8 ± 5.9	0.70	0.13	5.4 ± 7.2	0.32	0.75	5.3 ± 6.4	0.05	0.83
	High	6.7 ± 7.2	0.41	0.94	7.9 ± 12.2	0.53	0.65	19.4 ± 21.8	0.43	0.89	3.4 ± 3.9	0.43	0.88	2.5 ± 1.9	0.31	1.33
b	Low	0.8 ± 0.8	0.15	0.95	0.9 ± 1.0	0.13	0.86	1.5 ± 1.1	0.29	1.45	0.3 ± 0.5	0.41	0.31	0.5 ± 0.4	0.06	1.15
	Int.	6.2 ± 8.0	0.09	0.77	4.8 ± 4.8	0.04	1.01	5.2 ± 5.2	0.02	0.99	0.2 ± 4.1	0.94	0.00	6.3 ± 5.9	0.05	1.06
	High	1.4 ± 1.9	0.48	0.75	1.2 ± 3.6	0.71	0.35	2.6 ± 1.8	0.39	1.46	3.0 ± 3.6	0.45	0.82	1.0 ± 1.9	0.59	0.54
c	Low	1.0 ± 1.1	0.18	0.88	0.7 ± 0.6	0.06	1.16	2.4 ± 2.1	0.36	1.11	0.3 ± 0.6	0.21	0.49	0.5 ± 0.7	0.23	0.62
	Int.	3.8 ± 4.1	0.04	0.62	4.7 ± 3.6	0.02	1.18	3.4 ± 2.7	0.005	1.26	2.7 ± 1.4	0.08	1.93	6.2 ± 4.3	0.02	1.42
	High	1.0 ± 1.4	0.48	0.75	1.6 ± 2.1	0.50	0.73	2.3 ± 1.4	0.26	1.67	4.0 ± 5.1	0.47	0.79	1.7 ± 2.9	0.56	0.59
d	Low	0.5 ± 0.5	0.13	1.03	0.7 ± 0.4	0.02	1.60	2.3 ± 2.0	0.36	1.12	0.4 ± 0.5	0.06	0.77	0.5 ± 0.7	0.29	0.67
	Int.	3.2 ± 4.0	0.02	0.78	4.2 ± 2.8	0.01	1.50	3.4 ± 2.6	0.004	1.30	3.6 ± 3.1	0.18	1.15	6.0 ± 4.0	0.02	1.51
	High	1.2 ± 1.7	0.49	0.72	2.1 2.2	0.41	0.95	0.4 ± 1.0	0.69	0.38	2.7 ± 3.8	0.50	0.71	2.5 ± 2.1	0.29	1.13
e	Low	1.0 ± 1.1	0.18	0.87	0.8 ± 0.5	0.03	1.50	1.7 ± 1.1	0.28	1.52	0.3 ± 0.4	0.03	0.93	0.4 ± 0.6	0.36	0.78
	Int.	3.1 ± 2.7	0.02	1.17	3.4 ± 2.2	0.01	1.54	2.5 ± 2.0	0.005	1.28	3.4 ± 3.1	0.19	1.11	4.0 ± 2.6	0.01	1.52
	High	1.7 ± 0.4	0.11	4.01	5.3 ± 3.5	0.28	1.52	2.0 ± 0.8	0.45	0.82	2.3 ± 2.1	0.37	1.07	1.4 ± 2.1	0.26	0.67

		Sol			PT			Peroneal			Tibial			Whole calf		
method	fat infiltration	mean change±sd	p	SRM	mean change±sd	p	SRM	mean change±sd	p	SRM	mean change±sd	p	SRM	mean change±sd	p	SRM
a	Low	0.4 ± 0.4	0.02	1.04	1.0 ± 0.9	0.03	1.10	0.7 ± 0.9	0.24	0.77	0.4 ± 0.4	0.04	1.10	0.4 ± 0.5	0.11	0.92
	Int.	-0.1 ± 0.7	0.76	-0.20	2.0 ± 4.5	0.43	0.45	5.0 ± 13.0	0.34	0.39	3.1 ± 4.9	0.29	0.64	3.6 ± 3.9	0.08	0.91
	High	1.7 ± 1.0	0.25	1.71	0.9 ± 1.3	0.50	0.70	3.3 ± 3.7	0.42	0.90	2.1 ± 0.7	0.15	2.97	2.8 ± 0.9	0.14	3.03
b	Low	0.4 ± 0.6	0.10	0.66	0.4 ± 0.6	0.12	0.68	1.6 ± 0.8	0.03	2.00	0.4 ± 0.4	0.06	0.99	0.6 ± 0.3	0.02	1.77
	Int.	3.4 ± 5.8	0.42	0.58	1.2 ± 2.9	0.48	0.40	6.2 ± 5.0	0.02	1.24	4.4 ± 3.2	0.07	1.35	4.2 ± 2.8	0.01	1.50
	High	0.6 ± 4.1	0.87	-0.15	3.3 ± 1.2	0.16	2.67	-5.6 ± 10.2	0.58	-0.55	1.1 ± 2.5	0.64	0.45	1.2 ± 2.0	0.53	0.63
c	Low	0.4 ± 0.6	0.09	0.70	0.4 ± 0.4	0.05	0.90	1.4 ± 1.1	0.08	1.28	0.4 ± 0.4	0.07	0.94	0.6 ± 0.6	0.08	1.05
	Int.	3.5 ± 5.7	0.40	0.61	2.4 ± 1.9	0.08	1.30	3.6 ± 2.1	0.02	1.16	4.7 ± 3.0	0.05	1.58	3.5 ± 2.3	0.02	1.31
	High	-0.9 ± 5.7	0.85	-0.16	3.8 ± 1.1	0.13	3.41	2.1 ± 0.80.	0.16	2.69	1.6 ± 3.4	0.63	0.46	1.7 ± 2.7	0.54	0.61
d	Low	0.4 ± 0.5	0.07	0.76	0.4 ± 0.4	0.04	0.99	1.2 ± 0.8	0.05	1.59	0.4 ± 0.4	0.08	0.88	0.5 ± 0.5	0.07	1.13
	Int.	3.0 ± 3.8	0.3	0.80	2.4 ± 1.4	0.04	1.78	3.6 ± 2.1	0.004	1.73	4.4 ± 2.3	0.03	1.88	3.5 ± 2.4	0.01	1.49
	High	-0.3 ± 4.4	0.95	-0.06	3.4 ± 1.8	0.24	1.81	1.5 ± 1.3	0.35	1.15	2.1 ± 2.3	0.42	0.92	2.0 ± 2.1	0.41	0.95
e	Low	0.4 ± 0.5	0.04	0.87	0.6 ± 0.5	0.02	1.20	1.2 ± 0.8	0.06	1.46	0.4 ± 0.4	0.05	1.08	0.5 ± 0.4	0.06	1.20
	Int.	2.6 ± 2.8	0.26	0.91	2.5 ± 0.7	0.01	3.51	3.2 ± 1.8	0.003	1.78	3.7 ± 2.3	0.05	1.62	3.0 ± 2.0	0.01	1.50
	High	0.3 ± 1.8	0.84	0.18	3.9 ± 2.4	0.26	1.66	2.5 ± 1.4	0.25	1.72	1.7 ± 0.8	0.20	2.12	1.9 ± 0.2	0.06	7.73

Values are mean change in fat fraction (%) ± standard deviation. Low=FF≤4.8%, Int.=FF between 4.8 and 70%, High=fat fraction >70%. FF=fat fraction, s.d=standard deviation, SRM=standardised response mean. P value for paired t-test or Wilcoxon test as appropriate. TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, MG=medial gastrocnemius, LG=lateral gastrocnemius, Sol=soleus, PT=tibialis posterior, peroneal=peroneal innervated group, tibial=posterior tibial innervated group. Rows highlighted in blue when p value <0.05

4.5 Discussion

The development of responsive outcome measures for clinical trials in slowly progressive neuromuscular disorders such as CMT1A has gained significant momentum, with qMRI emerging as the clear front runner. In Chapter 3 of this thesis, we have demonstrated in CMT1A, that qMRI is highly responsive over extended follow-up, measuring a significant increase in FF at maximum trial duration of 1.0 ± 0.9%/year, (p=0.001) with SRM of 1.07. This outcome measure responsiveness is far superior to previously used outcome measures in CMT1A, both in our one-year study, and the wider literature (Piscosquito et al., 2015). We have also demonstrated in Chapter 3, several methods by which qMRI responsiveness could be further augmented to increase study power.

The current study is timely in revealing several further important strategies for finessing the use of qMRI as an outcome measure for rare neuromuscular diseases (NMD).

4.5.1 Cross-sectional calf muscle fat gradients

We have demonstrated that there are significant superior-inferior calf muscle fat gradients in patients with CMT1A, and that these gradients can be accurately quantified by MRI. Despite several publications reporting semi-quantitative assessments of calf (Pelayo-Negro et al., 2014) and thigh (Gaeta et al., 2012) muscle fat infiltration in CMT1A, this is the first study which quantifies the degree and pattern of calf muscle fat gradients in this disease.

Lower limb muscle fat gradients in CMT1A may be expected intuitively to be length-dependent (i.e. maximal fatty infiltration distally) given the clinical nature of the disease (Reilly, de Jonghe and Pareyson, 2006). Although it has been reported that foot muscles are affected by fat infiltration before calf muscles, which in turn are affected before thigh muscles, our study has revealed that within the calf muscle itself, fatty infiltration is not consistently predictable, not always length-dependent, and may be either positive, negative, both a combination of positive and negative, or indeed without gradient at all. Interestingly, a reverse gradient was noted in 18% of muscles in a series of three patients reported with CMT2 due to DNM2 mutation (Gallardo et al., 2008).

4.5.1.1 Individual muscle and combined muscle gradients

In CMT1A, the muscle with the most number of positive FF gradients was tibialis anterior (8/13 patients), followed by peroneus longus (6/13 patients), then EHL, MG and soleus (each 5/13 patients). The fewest number of muscle FF gradients was seen in tibialis posterior (2/13 patients) and lateral gastrocnemius (3/13 patients). The muscle with the most negative gradients was peroneus longus (6/13 patients). No patients had negative

gradients in tibialis anterior and lateral gastrocnemius. Lateral gastrocnemius had the largest number of patients with no gradient (10/13).

When peroneal innervated muscles were combined, 9/13 patients (69.2%) had a positive (length-dependent) muscle fat gradient, whereas in the posterior tibial group this was not the case, with the majority of patients demonstrating no gradient (61.5%), and only 30.8% of patients a positive gradient. This is in keeping with findings from a series of 23 CMT1A patients in whom qualitative MRI revealed a gradient in all calf muscles except for the deep posterior compartment (Stilwell, Kilcoyne and Sherman, 1995). Within common peroneal innervated muscles, negative muscle fat gradients were also seen in 30.8% of patients, even in muscles with predominantly positive muscle fat gradients. Combining all muscles resulted in only 2 patients with a positive gradient, and 11 without a gradient.

4.5.1.2 Effect of fat percentage on muscle fat gradients

The disparity in muscle fat gradient frequency between common peroneal and posterior innervated groups appears to be a function of the number of 'intermediate' fat infiltrated muscles in each group. Of the common peroneal innervated muscles, 23/39 (59.0%) were graded 'intermediate', almost double that seen amongst the posterior tibial innervated muscles in which only 16/52 (30.8%) muscles were graded 'intermediate'. Paradoxically, the percentage of low/normal fat infiltrated muscles was 10/39 (25.6%) and 28/52 (53.8%) in peroneal and tibial muscles respectively. Within each group of intermediate muscles, a similar proportion had gradients: 16/23 (69.6%) peroneal innervated and 13/16 (81.3%) tibial innervated.

This study demonstrates that muscle fat gradients occur more commonly in muscles with intermediate FF (29/39 intermediate muscles had gradients). What is not yet clear is why muscles with 'intermediate' FF occur more commonly anterolaterally, than posteriorly. Each muscle will pass from minimally fat infiltrated to almost completely fat infiltrated at some point over the patient's lifetime, but the trigger for this, and the speed at which it occurs is not yet clear. One possibility is that common peroneal innervated muscles pass through the fatty atrophy process at a slower rate and are therefore more easily 'seen' in the intermediate phase when assessed over a relatively short period (as in 12 months here), whereas the posterior tibial innervated muscles complete the process of fatty atrophy much more rapidly, and are thus harder to spot in a 12 month period when intermediate. This process may also occur later in posterior tibial muscles – it is noted in this study that in all patients, peroneal muscles were more heavily fat infiltrated than tibial muscles. This remains to be further examined in future studies.

4.5.2 Magnitude of muscle fat gradients

In the CMT1A group, we determined the largest gradients to be in the common peroneal innervated muscles, with a maximum positive gradient of 12.8%/cm in tibialis anterior

muscle, and a maximum negative gradient of -10.6%/cm in peroneus longus. In the posterior tibial group of muscles, maximum positive gradient was 9.1%/cm in soleus and tibialis posterior, with maximum negative gradient of -8.2%/cm, again in tibialis posterior. As may be expected, grouping the muscles together resulted in mean gradients reduced in magnitude to a degree, with a maximum gradient in whole calf of 4.0%/cm and negative gradient of -2.0%/cm. The peroneal innervated group had a maximum gradient of 5.6%/cm compared with 5.1%/cm in the posterior tibial innervated group.

In the control group, there were no participants with positive gradients, though there were maximum positive gradients of >2%/cm in some muscles. On the other hand, we measured maximum negative gradients of -6.7%/cm in controls – at the proximal end of the imaging block. In both controls and CMT1A patients, there was a tendency for the most proximal slice of the imaging block to have higher FF than the muscle belly, which is likely to be either an MR measurement artefact or an anatomical effect related to the inclusion of tendons and other non-muscle tissue, rather than this being a reflection of the muscle itself.

The magnitude of these muscle fat gradients is of particular concern in longitudinal imaging studies. In Chapter 3, we demonstrated a FF change in the order of 1-2%/year for patients with CMT1A. If care is not taken to account for these proven muscle fat gradients, there is a real risk of Type II error in clinical trials going forward.

4.5.3 Longitudinal fat fraction change and qMRI responsiveness

In this study, qMRI-determined whole calf FF changed significantly at 12 months by all methods of FF calculation except method a (central slice). The latter is not surprising given that longitudinal patient positioning in this study was based on surface anatomy, and the centre of the imaging block differed by up to 10cm between baseline and follow-up scans in some patients. Indeed no individual muscle or group of muscles changed significantly by method a.

By method b (slice closest to a fixed point 14cm distal to the tibial plateau), which is the method used in our longitudinal natural history study (Chapter 3), qMRI measured significant change in whole calf FF of 2.2 ± 2.6% (p=0.002) with large responsiveness: SRM = 0.90. By this method, significant FF change was also seen in tibialis anterior (3.8 ± 6.3%), peroneus longus (4.2 ± 4.6%) and the peroneal innervated group (3.0 ± 6.3%). Maximum SRM was 0.92 for peroneus longus. Note is made that the responsiveness is almost identical to the 0.91 (right calf) seen at 12 months in the longitudinal natural history study (Chapter 3), though the CMT1A cohort is not identical. Posterior tibial muscles showed no significant change by method b.

By method c, (weighted average of two adjacent slices about a fixed point 14 cm distal to the tibial plateau), FF change in all three combined measures was significant, as well

as in all individual peroneal muscles and lateral gastrocnemius. In whole calf, qMRI measured significant change of 2.6 ± 3.0% with SRM of 0.84. Several further individual muscles also demonstrated significant FF change by this method. Almost across the spectrum of muscles, there was an increase in qMRI responsiveness between method b and method c, in keeping with an increase in accuracy of baseline and longitudinal slice positioning.

By method d (weighted average of 3 or 4 slices centred 14 cm distal to tibial plateau), almost all individual muscles and combined muscles showed significant change, and by method e (maximum available slices), only soleus showed no significant FF change. The largest SRM overall was 1.46 for the peroneal innervated muscle group by method e.

In all peroneal muscles and the peroneal group, there was a progressive increase in responsiveness measured by SRM, as the method of slice selection sampled more muscle. The most dramatic improvement was in the peroneal innervated muscle group which showed significant FF change by methods b, c, d and e with an SRM which increased from 0.47 → 0.92 → 1.29 and 1.46 by methods b, c, d and e respectively. Responsiveness of the posterior tibial innervated group also increased with method of selection, though not as dramatically: from 0.72 (method c) to 0.89 (method e). On the other hand, for whole calf, FF change was significant by methods b, c, d and e, but SRM was flat: 0.90, 0.84, 0.97 and 0.91 by these methods respectively.

In individual muscles, those muscles innervated by the common peroneal nerve, showed greatest rise in SRM with increasing slice-selection method sophistication, reflecting the results of their combined-measure SRMs. Tibialis anterior from 0.60 (method b) to 1.00 (method e), EHL from 0.88 (method c) to 1.17 (method e), and peroneus longus from 0.92 (method b) to 1.26 (method e). In posterior tibial innervated muscles however, although there was significant change in almost all muscles by method e, SRM remained lower and flat across the board.

The factor that changes to bring about this increased responsiveness is reduction in measurement error as reflected in the standard deviation, whereas mean change remains more constant. As an example, in the peroneal innervated group, mean FF change by methods b, c, d and e was 3.0%, 3.8%, 2.6% and 2.4%. Standard deviation however consistently reduced: 6.3% → 4.2% → 2.0% → 1.7%. This reduction in standard deviation was seen across all muscles and groups of muscles.

4.5.4 Effect of baseline stratification

As was seen in the CMT1A natural history study, stratification of muscles based on baseline FF resulted in marked improvement in outcome measure responsiveness as a result of maximising and homogenising change. Almost without exception, maximum FF change, and responsiveness were seen in the cohort of patients with baseline intermediate fat infiltration. The method of fat calculation again had some effect on

responsiveness with a general increase in outcome measure responsiveness as the method of slice selection assessed more muscle. It is noted that due to small numbers of patients, many of the FF changes in stratified groups were not significant compared with baseline.

4.5.5 Implications for trial design

The findings of this study have several important implications for future trial designs in rare neuromuscular diseases.

4.5.5.1 Accurate positioning

In choosing a particular slice for longitudinal analysis, the presence of significant muscle fat gradients, underlines the critical importance of accurate slice positioning and selection for analysis. With a maximum muscle fat gradient of 12.8%/cm in CMT1A patients and 6.7%/cm in healthy controls, and an expected FF change over 12 months in the region of 1.5-4.0%, it is clear that longitudinal slice positioning and selection must be precise. Any small positioning error in longitudinal slice selection risks diluting the small and real change in FF with the much larger measurement error due to axial slice misalignment, resulting in type II error. This point is particularly relevant for centres using 2D Dixon MRI, in which protocols adjacent axial slices may be up to 2cm apart, implying slices analysed may be up to 1cm out of alignment if slices are perfectly interleaved. This issue may be somewhat mitigated by choosing to analyse a muscle in which there is less chance of significant muscle fat gradients (e.g. MG, LG, PT), however it should be noted that muscle fat gradients are not predictable, and these muscles are less responsive to change.

4.5.5.2 Single or multiple slices

In terms of the method of slice selection for longitudinal analysis, this study has clearly demonstrated increasing responsiveness (measured by SRM) by analysing a larger 'block' of muscle rather than a single slice. In EHL for example, SRM increased from 0.76 (method b) to 1.17 (method e). This is particularly relevant for muscles with fat gradients (peroneal innervated muscles) in which small errors of placement have larger repercussions, whereas a small disparity in longitudinal slice position is less likely to be relevant in a muscle with no or little gradient. However as noted above, there are drawbacks to choosing a muscle with less likelihood of a fat gradient as well. Similarly, in the whole calf however, SRM stayed essentially stable which is an effect of dilution of all the individual muscle FF changes.

The increase demonstrated in outcome measure responsiveness by sampling a larger volume of muscle is due to a combination of correction of mispositioning (improvement from method b to method c) as well as to reduction in the 'noise' by larger sample size, thus reducing standard deviation of change (from method c → d → e).

It should be noted at this point that simply analysing more slices without correct positioning does not improve responsiveness – as the amount of overlapping muscle remains proportionally unchanged, and this may in fact worsen responsiveness given the somewhat random muscle fat gradients which have been demonstrated in this chapter. The ideal therefore is highly accurate positioning to maximise overlapping muscle longitudinally, which going forward would be facilitated by using imaging with less interslice distance (as with 3D 3-point Dixon). A larger 'volumetric' block of muscle then will reduce noise further improving responsiveness, and also guards against small errors in positioning if they do still occur (Figure 6-1).

4.5.5.3 Single or combinations of muscles

This study shows the superiority of analysing combinations of muscles, rather than a single muscle over 12 months. Analysis of posterior tibial innervated muscles (SRM 1.07 by method e in PT) or combinations of muscles (0.89 by method e) is less beneficial than common peroneal innervated muscles (SRM 1.26 by method e in PL and 1.46 in peroneal muscles by method e). Whole calf analysis lies as expected between the two (SRM 0.91 by method e). As previously discussed, the markedly higher responsiveness in peroneal innervated muscles may well relate to the increased proportion of intermediate muscles in that group, compared with tibial innervated muscles.

This incremental benefit in terms of outcome measure responsiveness however, must be balanced against the time required to manually segment muscle regions of interest, although developing automated techniques may mitigate this.

4.5.5.4 Subpopulation analysis

We have again shown, as demonstrated in Chapter 3, and now seen in a context with even greater impact, that choosing a subpopulation of patients with intermediate FF will ensure maximum FF change over time, though this must be tempered with the need for adequate patient-severity representation. In this study, all combined muscle SRM markedly increased when including intermediate patients only. By method e: Peroneal group from 1.46 to 1.75; Tibial group from 0.90 to 1.62, and whole calf from 0.91 to 1.50. This was the case with all methods of FF calculation. Maximum SRM across the whole study as seen in tibialis posterior (intermediate group) by method e, with highly significant FF change of 2.5 ± 0.7% (p=0.01) and SRM of 3.51.

What this study adds to the natural history study is the comparison between method 'b' which was used in the natural history study, and other more sophisticated methods used here. In all individual muscles and summary FF measures, method e results in greater responsiveness in the intermediate group compared with method b, which again demonstrates the benefit to responsiveness of analysing a larger 'volume' of muscle.

4.5.5.5 Why are there more intermediate peroneal than tibial innervated muscles?

The author suggests that this disparity may relate to the distribution of motor endplates (MEP) in skeletal muscle. Pathological studies have revealed that the distribution of motor endplates varies between muscles. Aquilonis et al. demonstrated that in the tibialis anterior (TA) muscle, most of the motor endplates were superficially distributed along the length of the whole muscle. (Aquilonius et al., 1984). In the posterior tibial muscle however, Oddy and colleagues report that the motor endplates were ~22.1% of the distance from the fibular head to the malleoli – in a horizontal strip across the muscle (Oddy et al., 2006). If there is such a difference between MEP location in peroneal versus tibial innervated muscles, it may stand to reason that the longer distance to the most distal MEPs in the tibialis anterior muscle (or other common peroneal innervated muscle) causes them to be affected before the length-dependent pathological process reaches the muscle belly where the MEPs for the tibialis posterior muscle are situated. In TA, MEPs are affected sequentially (from distal to proximal), resulting in a slower progression of fat collection, whereas in TP, all MEPs are affected together, but later (by which point the peroneal muscles are already moderately affected), resulting in a more rapid progression of fat infiltration. This of course is a hypothesis requiring further study.

4.6 Conclusions

We have demonstrated that lower limb proximal-distal muscle fat gradients exist in all calf muscles of patients with CMT1A, and that these gradients can be accurately quantified by MRI. Muscle fat gradients are not reliably length-dependent, and when present may be positive, negative, or a combination of both. These muscle fat gradients occur more commonly in intermediate fat infiltrated muscles, and can be very steep.

The presence, and extent of muscle fat gradients informs further refinement of qMRI outcome measure responsiveness. Most importantly, precise longitudinal axial slice selection is critical given that small errors in slice placement may result in large fat-gradient-induced measurement error diluting the true longitudinal FF change, making it potentially undetectable.

Quantitative MRI-measured FF responsiveness can be further enhanced by analysing larger muscle 'volume', by analysing specific individual muscles or muscle groups, and by selecting muscles or patients with intermediate fat infiltration, in which mean FF change may be expected to be greatest over study duration. Combining these three methods leads to biggest gains.

In this study, CMT1A patients with normal or high calf fat infiltration at baseline showed little change over the study period, and it is expected that analysis of foot muscles in the

former, and distal thigh muscles in the latter may identify muscles with an intermediate FF in which change over 12 months may be optimised. This concept is further examined in Chapter 5.

We have since added 3D 3-point Dixon imaging to our imaging protocols. The thinner slice thickness with no slice separation, and more extensive coverage, will facilitate more precise longitudinal slice selection, and pseudo-volumetric analysis, both of which are expected to further improve qMRI responsiveness.

5 Longitudinal quantitative MRI in Hereditary Sensory Neuropathy type 1

5.1 Introduction

Hereditary Sensory and Autonomic Neuropathy type 1 (HSAN1 or HSN1) due to *SPTLC1* and *SPTLC2* mutations, is a predominantly sensory peripheral neuropathy which is inherited in an autosomal dominant manner. The neuropathy results in mutilating sensory complications including non-healing ulcers and osteomyelitis which when severe may necessitate lower limb amputation. (Houlden et al., 2006). Patients very often have associated neuropathic pain due to concomitant small fibre sensory nerve involvement. Somewhat surprisingly, weakness develops in a large proportion, particularly later in the disease course and can be a source of marked morbidity, whereas autonomic symptoms occur in a minority. Mortality in HSN1 is earlier than in the general population (Houlden et al., 2006).

The *SPTLC1* and *SPTLC2* genes encode the first two subunits of the enzyme serine palmitoyltransferase (SPT), which catalyses the first and rate limiting step of sphingolipid synthesis by conjugating palmitoyl-CoA and L-serine. Mutations in *SPTLC1* and *SPTLC2* cause the neuropathy of HSN by a gain of function mechanism, as mutations alter the substrate specificity of SPT whereby alanine and glycine are preferred over the canonical serine, resulting in the production of neurotoxic deoxysphingolipids (Eichler et al., 2009; Penno et al., 2010). Exogenous L-serine is a potential treatment based on its ability to reduce neurotoxic plasma 1-deoxysphingolipid levels (Garofalo et al., 2011). There is an urgency to assess drug efficacy in clinical trials.

Of major importance to clinical trials in HSN1 is that its prevalence is only several hundred cases worldwide, and so even with international collaboration, it is not possible to conduct large randomised controlled trials in this disease. Short of significantly extending study duration or developing medications which reverse axonal loss, the current need is a highly responsive outcome measure to allow adequately powered clinical trials of feasible duration.

5.2 Background

A small randomised placebo-controlled crossover trial in 18 patients with symptomatic HSN1 showed that treatment with 400mg/kg/day of L-serine is safe in humans, and may potentially slow disease progression (Fridman et al., 2019) however this trial was primarily a drug safety trial. The trial used change in the Charcot-Marie-Tooth neuropathy score version 2 (CMTNSv2) as the primary outcome measure – (> 1 point over 12 months). A battery of secondary outcome measures were also used including plasma sphingolipids and amino acids, epidermal nerve fibre density, electrophysiology and patient reported measures in the form of the SF-36 and the Neuropathy Pain Scale. The primary outcome measure was reached by one patient in the treatment group, and two in the placebo group, hence the trial was negative on the pre-specified primary end-point. The CMTNSv2 improved over two years in the L-serine group with change: -1.14 ± 0.76 points (mean ± SEM) though not significantly so (p=0.15). The placebo group continued to get worse (CMTNSv2 change: +1.1 ± 0.47 points over 12 months, p=0.02). At 12 months, there was a relative difference in CMTNSv2 between the groups. The Charcot-Marie-Tooth examination score (CMTES) showed similar findings, with relative change of -1.2 points between groups at 12 months (p=0.05). Although not specified in the paper, the standardised response mean for the placebo group calculated as: mean change/[SEM x √sample size] = 0.78 – a marked improvement in responsiveness of Charcot-Marie-Tooth neuropathy score (CMTNS) compared with previous publications. Except for a reduction in 1-deoxysphingolipid, none of the other outcomes measures showed significant change. Critically however, quantitative MRI was not included as an outcome measure.

This recent study demonstrates potential issues arising when small participant numbers are joined with an outcome measure with even moderate responsiveness. The major subsequent risk is of missing a significant treatment benefit (type II error).

As there is a British founder mutation (C133W mutation in *SPTLC1*) (Nicholson et al., 2001), the cohort of patients in the UK is amongst the largest worldwide, putting the UK in a good position to lead a trial in this disease, but first a highly responsive outcome measure is needed.

5.3 Aim

The primary aim of this study is to determine whether MRI can provide a quantitative assessment of disease progression over one year in patients with Hereditary Sensory Neuropathy type 1. Complementary to the previous chapters, methods of longitudinal slice selection in this cohort are compared, in order to improve responsiveness of quantitative MRI as an outcome measure. It should be noted that part of the MRI results from this thesis have already been published by our group (Kugathasan et al., 2019) as part of a larger study. The author performed all work related to the MRI aspects of this study, whereas all other outcome measures were performed and analysed by Dr Kugathasan.

5.4 Results

Patients and controls were matched for age and gender. The severity of clinical disease as defined by CMTNS was heterogeneous (Table 5-2 and Figure 5-1).

5.4.1 Cross-sectional results

34 patients with HSN1 and 10 matched controls were analysed at baseline. As part of our Centre's wider HSN1 natural history study, each underwent a battery of clinical, neurophysiological and laboratory studies. This thesis is dedicated to the analysis of lower limb MRI data from that study. Other results are previously published by our group (Kugathasan et al., 2019).

5.4.1.1 Clinical and demographic details

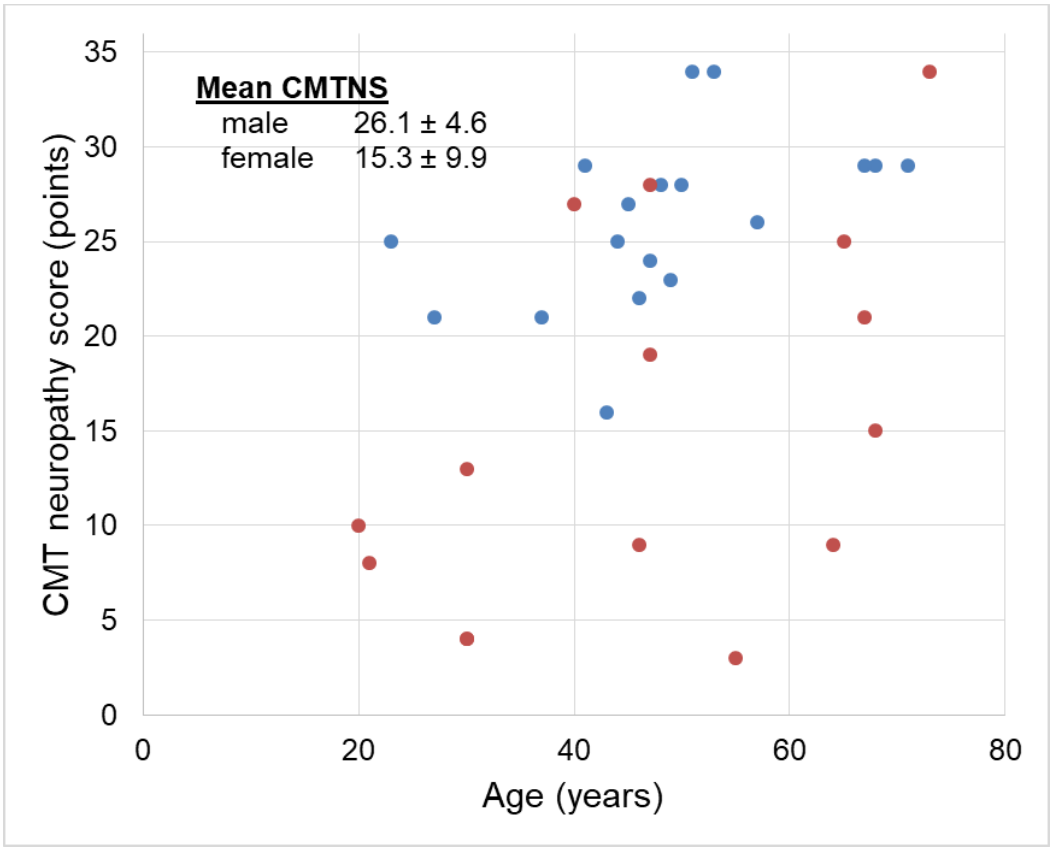
Demographic details are summarised in Table 5-1. There were no significant differences in age between the HSN1 and control groups (p=0.30). There were also no significant differences in age between males and females within the HSN1 group. A greater spread of disease severity was seen in females than males, with larger variation in CMTNS (range 3 to 34 in females vs 16 to 34 in males - Table 5-1). Males had more severe disease as demonstrated by CMTES (mean CMTES 19.1 ± 3.6 in males vs 12.2 ± 6.7, p=0.002 in females) and CMTNS (26.1 ± 4.6 in males vs 15.3 ± 9.9 in females, p=0.001). Age versus CMTNS divided for gender in the HSN1 cohort is plotted in Figure 5-1.

Table 5-1 – Demographic details in patients and controls

	HSN			Control					
	Overall	**Male**	**Female**	**Overall**	**Male**	**Female**	**p1**	**p2**	**p3**
Number	34	20	14	10	5	5			
Age (y) ± s.d	48.3 ± 14.9	49.3 ± 12.5	46.9 ± 18.0	43.0 ± 10.0	44.4 ± 7.7	41.6 ± 12.6	0.41	0.55	0.64
CMTES (0-28)	16.0 ± 6.6	19.1 ± 3.6	12.2 ± 6.7	ND	ND	ND	ND	ND	0.002
CMTNS (0-36)	21.1 ± 9.6	26.1 ± 4.6	15.3 ± 9.9	ND	ND	ND	ND	ND	0.001

y=years, s.d=standard deviation, HSN=hereditary sensory neuropathy, CMTNS=CMT neuropathy score, CMTES=CMT examination score, p value for χ^2 test for gender, and otherwise by independent t-test. p1=HSN1 v controls (female), p2=HSN1 v controls (male), p3=HSN1 males vs HSN1 females, ND=no data

Figure 5-1 – Age versus CMT neuropathy score in HSN1 cohort



Each point represents a single participant with HSN1. Blue=males, Red=females. CMTNS=Charcot-Marie-Tooth neuropathy score. CMTNS range is 0 to 36 points

5.4.1.2 Fat fraction

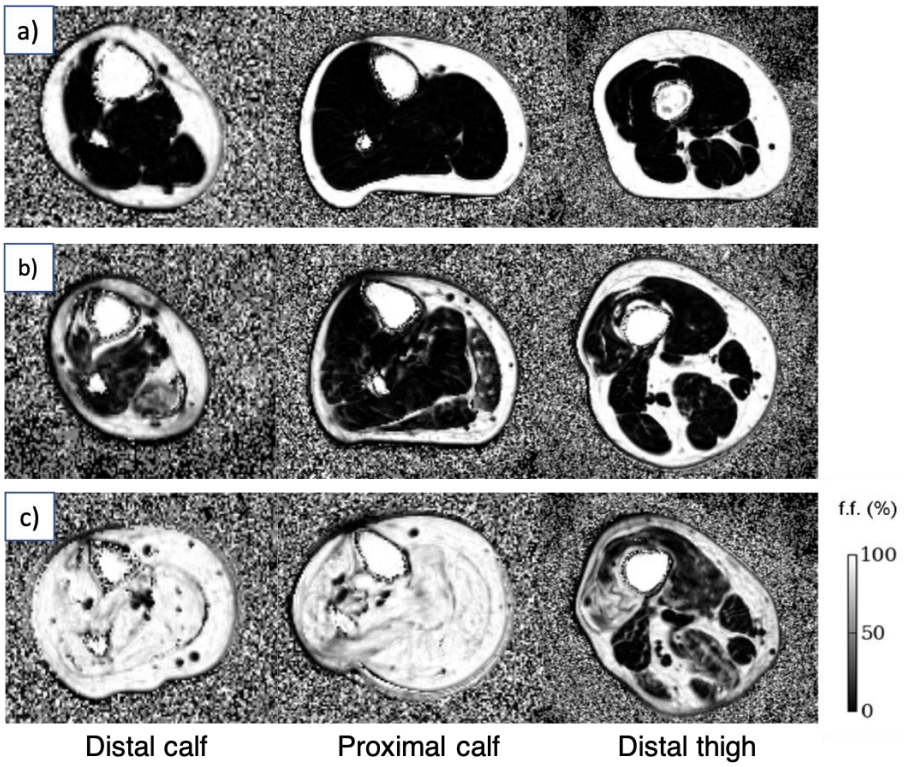
Example 3D 3-point Dixon fat fraction (FF) maps for are shown in Figure 5-2. Distal thigh slice was positioned at a distance 110mm rostral to the tibial plateau, whereas proximal calf slice is at a distance 110mm caudal and distal calf slice at a distance 330mm caudal to the tibial plateau (Figure 2-1).

5.4.1.2.1 Proximal calf

Combined bilateral proximal calf FF in the HSN1 cohort was significantly higher (33.8 ± 30.0%) than in matched controls (2.0 ± 1.5%, p<0.0001). Comparison of other individual and combined muscle mean FF revealed significantly higher values in all HSN1 muscles and combination of muscles compared with controls. FF was greater in males than females across the spectrum of individual and combined muscles, though this difference reached statistical significance only in left soleus (41.4 ± 31.2% in males vs 24.2 ± 30.9% in females, p=0.03) and left medial gastrocnemius (46.2 ± 33.7% in males vs 28.5 ± 30.4% in females, p=0.04). Combined bilateral proximal calf FF was 39.6 ± 29.5% in males and 27.3 ± 30.1% in females (p=0.17). Baseline proximal calf FF values for HSN1

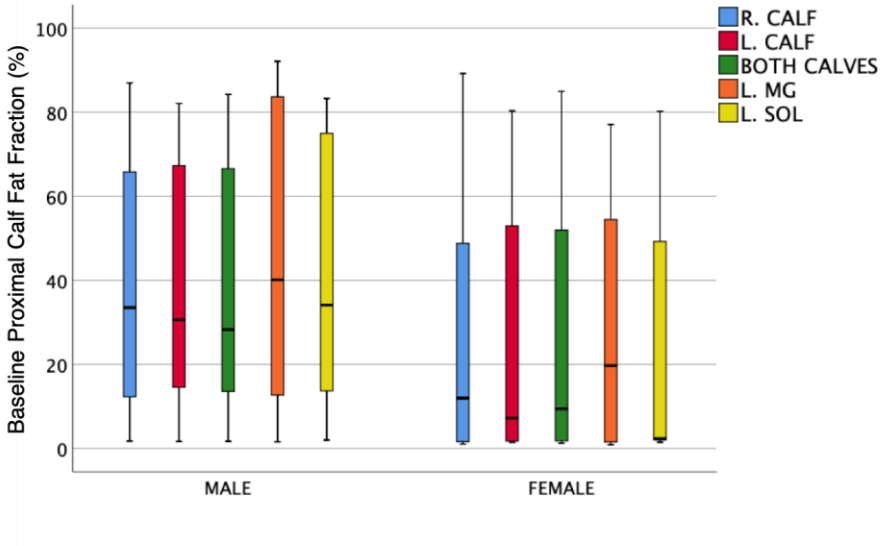
and control groups are given in Table 5-2, and median FF divided by gender in Figure 5-3. Scatterplot of age versus proximal calf FF divided for gender in the HSN1 cohort is shown in Figure 5-4.

Figure 5-2 – Example Dixon fat maps from control and HSN1 at three anatomical levels



Example fat fraction maps in a control (a) and two patients with HSN1. Participant in panel (b) has moderate disease, and in panel (c) severe disease. Each row of three images is from the same participant at the specified level

Figure 5-3 – Baseline proximal calf fat fraction divided by gender in HSN1



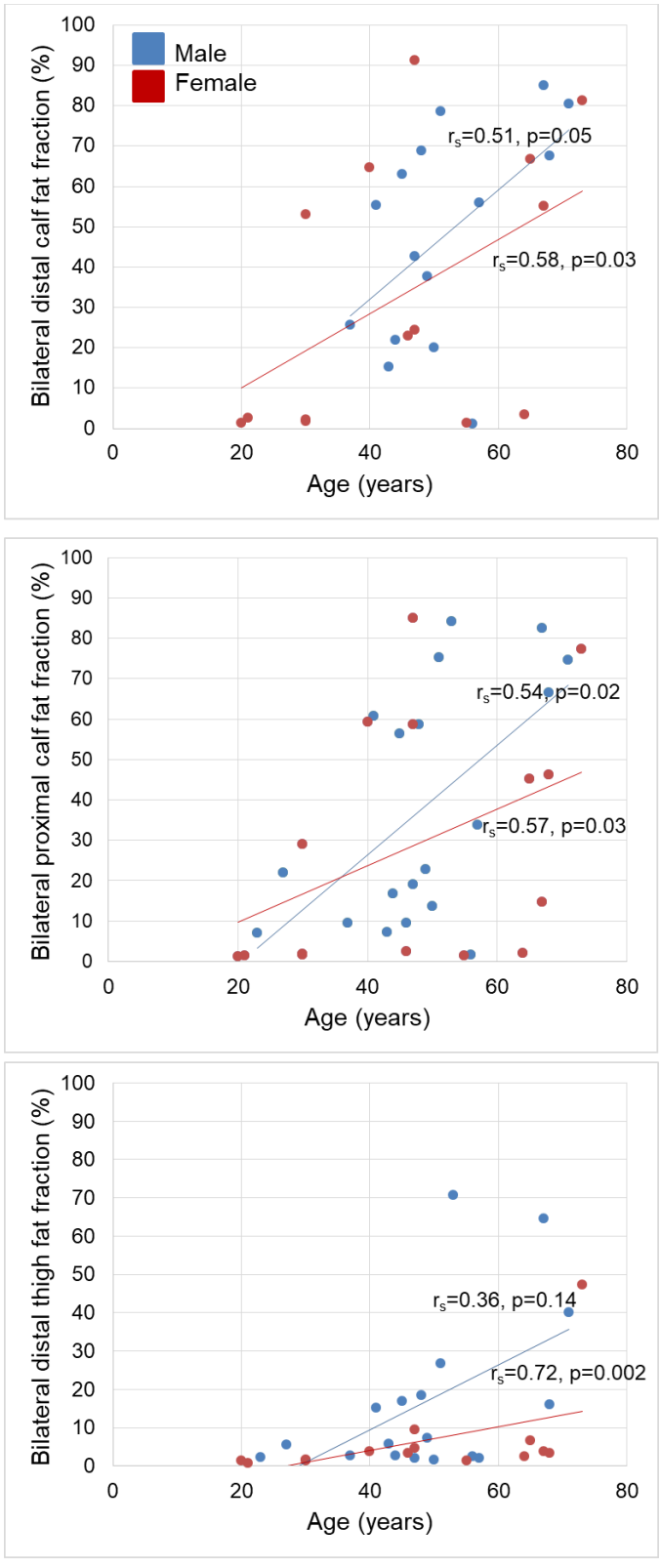
Boxes represent median and IQR, whiskers show range, and circles/asterisks are outliers. R.=right, L.=left, Sol=soleus, MG=medial gastrocnemius, LG=lateral gastrocnemius

Table 5-2 – Baseline proximal calf fat fraction in HSN1 and controls

	Fat fraction (%)					
	HSN			Control		
Muscle	Overall	Female	Male	Overall	p1	p2
R. TA	21.8 ± 22.2	21.3 ± 26.2	22.2 ± 18.6	1.0 ± 0.5	<0.0001	0.36
R. EHL	29.8 ± 28.0	23.8 ± 28.4	35.0 ± 27.4	1.4 ± 0.6	<0.0001	0.13
R. PL	38.1 ± 32.0	31.0 ± 34.1	34.4 ± 29.5	2.1 ± 1.2	<0.0001	0.13
R. LG	41.6 ± 37.9	37.0 ± 40.2	45.4 ± 36.5	2.0 ± 1.1	<0.0001	0.32
R. MG	41.3 ± 34.4	35.1 ± 36.4	46.3 ± 33.7	2.2 ± 1.9	<0.0001	0.32
R. SOL	35.8 ± 32.5	27.7 ± 31.7	43.1 ± 32.3	2.5 ± 2.3	<0.0001	0.07
R. PT	28.2 ± 27.3	22.6 ± 28.0	33.2 ± 26.5	1.4 ± 0.6	<0.0001	0.07
L. TA	22.5 ± 23.1	21.0 ± 27.0	23.8 ± 19.8	1.2 ± 0.7	<0.0001	0.22
L. EHL	31.3 ± 29.1	27.8 ± 30.4	34.5 ± 28.3	1.8 ± 1.4	<0.0001	0.33
L. PL	36.0 ± 31.9	28.5 ± 33.1	42.7 ± 30.3	2.1 ± 1.4	<0.0001	0.07
L. LG	41.4 ± 35.2	38.6 ± 39.1	43.8 ± 32.6	1.8 ± 1.3	<0.0001	0.58
L. MG	38.1 ± 33.0	28.5 ± 30.4	46.2 ± 33.7	2.3 ± 2.4	<0.0001	0.04
L. SOL	34.9 ± 32.3	24.2 ± 30.9	41.4 ± 31.2	2.5 ± 2.2	<0.0001	0.03
L. PT	27.9 ± 27.7	22.7 ± 28.7	32.1 ± 27.0	1.7 ± 1.2	<0.0001	0.09
R. CALF	34.3 ± 30.2	27.3 ± 29.9	40.5 ± 29.9	2.0 ± 1.5	<0.0001	0.11
L. CALF	34.0 ± 30.6	28.2 ± 31.5	39.2 ± 29.6	2.0 ± 1.6	<0.0001	0.13
BOTH CALVES	33.8 ± 30.0	27.3 ± 30.1	39.6 ± 29.5	2.0 ± 1.5	<0.0001	0.17

Values are mean baseline fat fraction ± standard deviation. P value for unpaired student t-test or Mann-Whitney U test as appropriate (p1=HSN1 versus control, p2=HSN1 males v females). HSN1=Hereditary Sensory Neuropathy, R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior. P value is highlighted in blue where difference is significant

Figure 5-4 – Combined bilateral fat fraction in males and females with HSN1 at each level



Age versus fat fraction at distal calf (a), proximal calf (b) and distal thigh (c) levels. Blue=male, red=female. r_s=spearman's rho correlation coefficient

5.4.1.2.2 Distal calf

Baseline combined bilateral distal calf fat FF in the HSN1 group was 40.1 ± 30.2% and 1.6 ± 1.2% in controls (p<0.0001). There were statistically significant differences between HSN1 and control groups in all individual and combined muscle FF (Table 5-3). Although FF was again greater in HSN1 males than females, this difference was not statistically significant at distal calf level. Baseline combined bilateral distal calf FF was 47.7 ± 26.6% in males, and 32.6 ± 32.6% in females (p=0.17). Mean difference between bilateral combined calf FF in the middle and distal blocks was 7.2 ± 13.2%, though this did not reach statistical significance (p=0.66). Scatterplot of age versus distal calf FF divided by gender in the HSN1 cohort is shown in Figure 5-4.

Table 5-3 – Baseline distal calf fat fraction in HSN1 and control

	Mean baseline fat fraction (%) ± s.d					
	HSN			Control		
Muscle	Overall	Female	Male	Overall	p1	p2
R. TA	35.5 ± 31.0	34.0 ± 33.4	37.3 ± 29.3	0.9 ± 0.8	<0.0001	0.71
R. EHL	30.7 ± 27.4	27.1 ± 29.6	34.9 ± 25.2	1.3 ± 1.3	<0.0001	0.27
R. PL	31.7 ± 27.4	27.3 ± 31.2	36.7 ± 22.5	1.9 ± 2.5	<0.0001	0.23
R. Sol	45.0 ± 35.7	38.9 ± 38.8	52.1 ± 31.8	2.1 ± 1.7	<0.0001	0.36
R. PT	38.1 ± 30.2	32.8 ± 32.6	44.2 ± 27.1	1.5 ± 0.8	<0.0001	0.25
L. TA	36.8 ± 31.3	33.2 ± 33.9	40.4 ± 29.2	2.1 ± 2.8	<0.0001	0.34
L. EHL	34.8 ± 29.1	28.5 ± 30.3	41.2 ± 27.4	1.4 ± 1.3	<0.0001	0.15
L. PL	31.6 ± 27.6	24.7 ± 30.8	38.4 ± 23.0	0.9 ± 1.0	<0.0001	0.11
L. Sol	44.0 ± 34.3	34.6 ± 36.7	53.4 ± 29.9	1.6 ± 2.4	<0.0001	0.21
L. PT	40.9 ± 30.3	32.9 ± 32.8	48.9 ± 26.2	1.5 ± 1.1	<0.0001	0.13
R. CALF	38.6 ± 30.2	33.3 ± 33.1	44.6 ± 26.4	1.7 ± 1.3	0.001	0.44
L. CALF	40.2 ± 30.6	32.0 ± 32.5	48.4 ± 27.1	1.4 ± 1.4	<0.0001	0.18
BOTH CALVES	40.2 ± 30.2	32.6 ± 32.6	47.7 ± 26.6	1.6 ± 1.2	<0.0001	0.17

Values are mean baseline fat fraction ± standard deviation. P value for unpaired student t-test or Mann-Whitney U test as appropriate (p1=HSN1 versus control, p2=HSN1 males v females). HSN1-Hereditary Sensory Neuropathy, R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus/brevis, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior. P value is highlighted in blue where difference is significant

5.4.1.2.3 Distal Thigh

In the HSN1 group, combined bilateral thigh FF was 9.7 ± 16.5% versus 1.8 ± 1.4% in controls (p=0.01). All combined FF values were statistically significantly higher in the HSN1 group compared with controls. All individual muscles in HSN1 had significantly higher baseline FF except rectus femoris, sartorius and gracilis bilaterally, left semitendinosus and right vastus medialis. Baseline combined bilateral distal thigh FF was 16.8 ± 21.2% in males and 6.2 ± 11.6% in females (p=0.05). FF was higher in males

than females across the spectrum of muscles and combined muscle measures, reaching statistical significance in a number of individual and combined muscles. Baseline distal thigh FF in HSN1 and controls is summarised in Table 5-4. Scatterplot of age versus distal thigh FF divided by gender in the HSN1 cohort is shown in Figure 5-4.

Table 5-4 – Baseline thigh fat fraction in HSN1 and controls

	Mean fat fraction (%) ± s.d					
	HSN			Control		
Muscle	Overall	Female	Male	Overall	p1	p2
R. VM	9.4 ± 18.2	4.2 ± 9.8	13.7 ± 22.4	1.6 ± 1.2	0.08	0.046
R. VI	9.8 ± 19.3	4.0 ± 9.1	14.6 ± 24.0	1.2 ± 0.9	0.005	0.06
R. VL	12.7 ± 21.8	6.4 ± 14.7	17.9 ± 25.5	1.2 ± 0.5	0.03	0.11
R. Grac	4.9 ± 7.2	2.4 ± 2.7	7.0 ± 9.0	2.0 ± 1.5	0.26	0.02
R. BF	10.3 ± 17.9	5.9 ± 10.9	14.0 ± 21.7	2.2 ± 1.8	0.02	0.06
R. SM	11.3 ± 18.8	4.4 ± 5.4	17.0 ± 23.8	2.7 ± 1.9	0.04	0.02
R. ST	10.1 ± 20.1	5.5 ± 13.0	13.8 ± 24.2	1.6 ± 1.0	0.03	0.04
R. AM	7.3 ± 11.4	2.7 ± 3.3	11.0 ± 14.2	3.0 ± 4.3	0.02	<0.0001
R. Sart	9.6 ± 13.5	5.5 ± 9.0	12.9 ± 15.8	3.4 ± 2.2	0.18	0.04
R. RF	11.2 ± 22.0	5.4 ± 15.8	16.1 ± 25.5	0.9 ±.0.8	0.05	0.19
L. VM	10.2 ± 17.7	5.2 ± 11.2	14.5 ± 21.0	1.6 ± 1.3	0.005	0.03
L. VI	9.9 ± 19.7	4.0 ± 8.9	14.7 ± 24.7	1.3 ± 1.0	0.003	0.03
L. VL	13.1 ± 20.5	6.8 ± 15.1	18.3 ± 23.3	1.5 ± 0.9	0.004	0.08
L. Grac	4.8 ± 8.8	3.1 ± 4.9	6.2 ± 11.0	1.6 ± 0.7	0.45	0.28
L. BF	9.0 ± 13.8	4.5 ± 6.6	12.7 ± 17.0	2.5 ± 2.1	0.02	0.046
L. SM	10.4 ± 15.5	5.0 ± 6.4	14.9 ± 19.3	2.0 ± 1.4	0.01	0.05
L. ST	10.2 ± 19.3	6.2 ± 14.2	13.6 ± 22.6	2.4 ± 2.0	0.15	0.14
L. AM	7.6 ± 13.2	2.8 ± 3.8	11.6 ± 16.7	2.8 ± 3.5	0.05	<0.0001
L. Sart	8.7 ± 12.7	5.8 ± 10.4	11.2 ± 14.1	3.3 ± 1.9	0.49	0.20
L. RF	13.3 ± 24.9	6.4 ± 19.8	19.0 ± 27.6	0.6 ± 0.4	0.15	0.06
R. THIGH	9.8 ± 17.0	5.9 ± 11.0	16.6 ± 21.7	1.8 ± 1.5	0.01	0.046
L. THIGH	9.7 ± 16.1	6.3 ± 12.1	16.9 ± 20.7	1.8 ± 1.4	0.01	0.046
BOTH THIGHS	9.7 ± 16.5	6.2 ± 11.6	16.8 ± 21.2	1.8 ± 1.4	0.01	0.05

Values are mean baseline fat fraction ± standard deviation. P value for unpaired student t-test or Mann-Whitney U test as appropriate (p1=HSN1 versus control, p2=HSN1 males v females). R.=right, L.=left, VM=vastus medialis, VI=vastus intermedius, VL=vastus lateralis, Grac=gracilis, BF=biceps femoris, SM-semimembranosus, ST=semitendinosus, AM=adductor magnus, Sart=Sartorius, RF=rectus femoris. P value is highlighted in blue where difference is significant

5.4.1.3 Muscle fat fraction distribution

Baseline proximal and distal calf median individual muscle FF in HSN1 and controls is shown in Figure 5-5 and Figure 5-6 respectively, and mean individual muscle FF in HSN1 and controls by gender at all three anatomical levels in Figure 5-10.

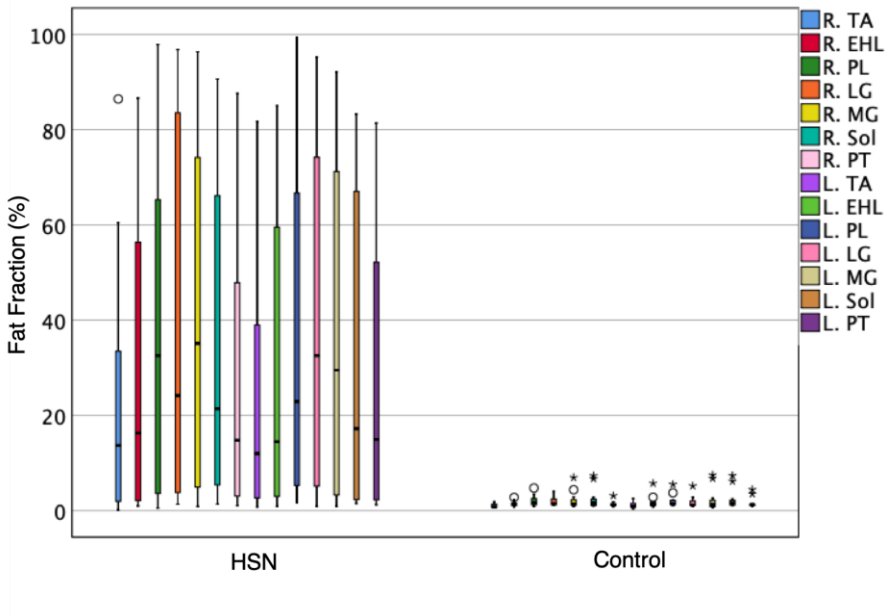
5.4.1.3.1 Proximal calf

In the HSN1 group at proximal calf level, the pattern of muscle involvement was the same in right and left calves, with posterior muscles affected more than anterior muscles, excepting tibialis posterior. The lateral gastrocnemius muscle was the most heavily fat infiltrated muscles in both calves: mean FF of 41.6 ± 37.9% (right), and 41.4 ± 35.2% (left). The tibialis anterior muscle was the least affected bilaterally: mean FF 21.8 ± 22.2% (right) and 22.5 ± 23.1% (left), though there was no statistically significant difference at baseline between HSN1 muscles, either when analysed as an entire cohort, or within gender groups.

In order to obtain a clearer picture of proximal calf FF pattern in HSN1, this analysis was repeated including only those HSN1 patients with bilateral combined proximal calf FF between 5% and 60% (n=20), thus excluding patients without a pattern due to ceiling or floor effect. In this cohort, there were significant inter-muscle FF differences: both right and left tibialis anterior were significantly less affected than peroneus longus, the gastrocnemii and right soleus. The gastrocnemii were also significantly more affected than tibialis posterior. Table 5-5 shows p values for inter-muscle comparison in this HSN1 subset at proximal calf level.

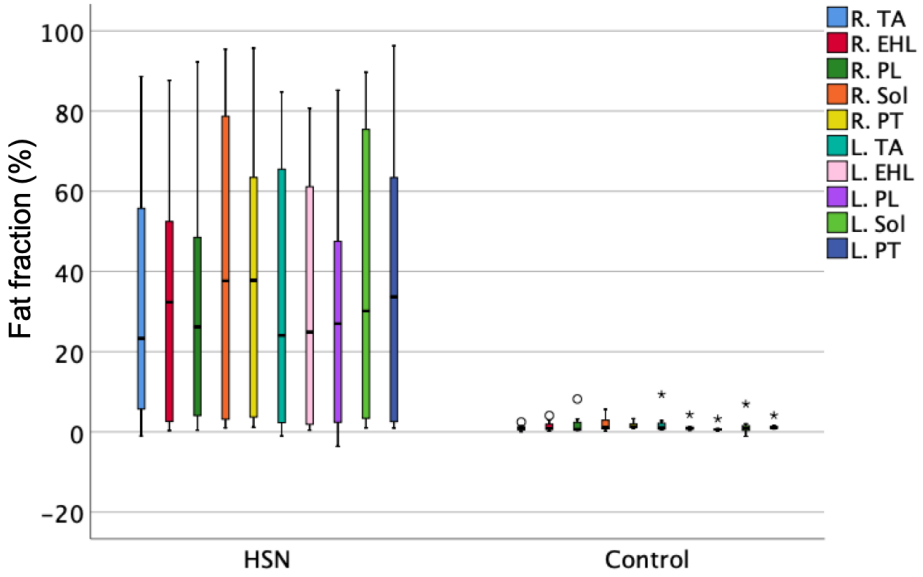
There were no significant differences between muscles in controls.

Figure 5-5 – Baseline proximal calf fat fraction in HSN1 and controls



Boxes represent median and IQR, whiskers show range, and circles/asterisks are outliers. R.=right, L.=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral head of gastrocnemius, MG=medial head of gastrocnemius, Sol=soleus, PT=tibialis posterior

Figure 5-6 – Distribution of fatty change in HSN1 and control at distal calf level



Boxes represent median and IQR, whiskers show range, and circles/asterisks are outliers. R.=right, L.=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, Sol=soleus, PT=tibialis posterior

Table 5-5 – p values for proximal calf inter-muscle FF differences in HSN1 cohort

Muscle	R.TA	R.EHL	R.PL	R.LG	R.MG	R.SOL	R.PT	L.TA	L.EHL	L.PL	L.LG	L.MG	L.SOL	L.PT
R.TA														
R.EHL	0.16													
R.PL	0.00	0.14												
R.LG	0.00	0.07	0.54											
R.MG	0.00	0.05	0.56	0.91										
R.SOL	0.03	0.41	0.54	0.26	0.24									
R.PT	0.35	0.57	0.05	0.02	0.01	0.16								
L.TA	0.94	0.19	0.01	0.01	0.00	0.04	0.40							
L.EHL	0.08	0.63	0.37	0.18	0.16	0.76	0.31	0.09						
L.PL	0.02	0.37	0.57	0.28	0.26	0.95	0.14	0.03	0.72					
L.LG	0.00	0.04	0.45	0.93	0.83	0.20	0.01	0.00	0.13	0.22				
L.MG	0.00	0.12	0.89	0.63	0.66	0.46	0.03	0.01	0.32	0.50	0.54			
L.SOL	0.07	0.58	0.41	0.20	0.18	0.82	0.27	0.08	0.95	0.77	0.15	0.35		
L.PT	0.43	0.56	0.05	0.03	0.01	0.17	0.96	0.48	0.32	0.15	0.01	0.04	0.28	

P values are for analysis of HSN1 patients with combined bilateral proximal calf FF between 5% and 60%. R.=right, L.=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral head of gastrocnemius, MG=medial head of gastrocnemius, Sol=soleus, PT=tibialis posterior. Cells are highlighted in blue if p<0.05. Comparison done by one way ANOVA with Bonferroni correction

5.4.1.3.2 Distal calf

At distal calf level, soleus was the most affected muscle and peroneus longus the least affected muscle bilaterally. The HSN1 group inter-muscle FF difference did not reach statistical significance at this level, in the whole group (right peroneus longus/brevis against right soleus: p=0.07), in intermediate FF only patients (bilateral combined distal calf FF between 5 and 60%, n=13) or when analysed based on gender. There were no significant differences between muscles in controls.

5.4.1.3.3 Distal Thigh

With the 95th centile in controls defining a 'normal' baseline thigh FF of <4.7%, 20/34 (58.8%) of the HSN1 group had a normal baseline combined bilateral thigh FF. Only four HSN1 patients (11.8%) had combined bilateral thigh FF >10%. In both thighs, vastus lateralis was the most affected muscle in the HSN1 group with FF of 12.7 ± 21.8% (R) and 13.1 ± 20.5% (L) in HSN1, and 1.2 ± 0.5% (R) and 1.5 ± 0.9% (L) in controls. The least affected muscles were gracilis and sartorius bilaterally. The differences in mean baseline bilateral thigh FF between muscles in the HSN1 group reached statistical significance between left vastus medialis (FF: 13.1 ± 20.5%) and left (FF 4.8 ± 8.8%, p=0.04) and right gracilis muscles (FF 4.9 ± 7.2%, p=0.03).

5.4.1.4 Cross-sectional area and remaining muscle area

5.4.1.4.1 Proximal calf

In the HSN1 cohort at baseline, combined bilateral proximal calf cross-sectional area (CSA) was significantly smaller (70.3 ± 25.8cm^2) compared with that in controls (108.7 ± 31.6 cm^2, p=0.0002). There were also significant differences in CSA between HSN1 and controls in almost all individual muscles, as well as in the left and right calf.

Table 5-6 – Baseline proximal calf cross-sectional area and remaining muscle area in HSN1 and control

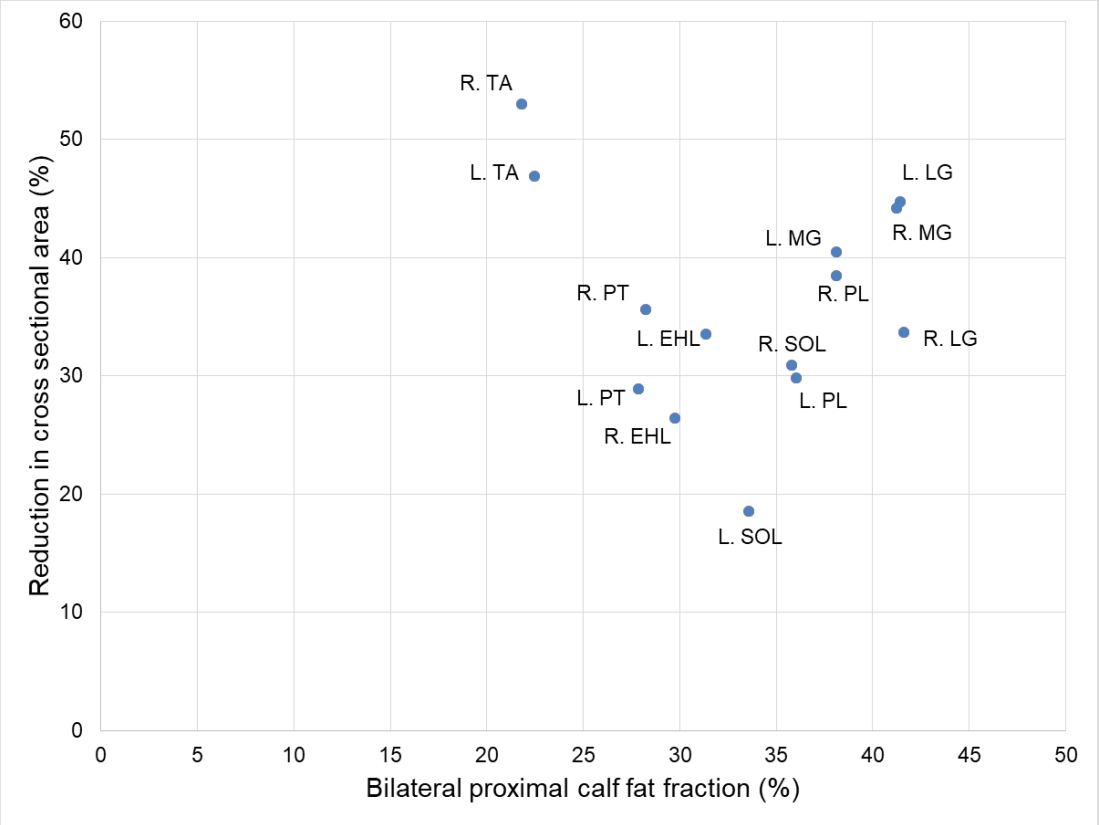
	CSA (cm^2)				RMA (cm^2)			
Muscle	**HSN**	**Control**	**HSN/Control (%)**	**p**	**HSN**	**Control**	**HSN/Control (%)**	**p**
R. TA	3.4 ± 1.5	7.3 ± 2.5	47.0	<0.0001	2.8 ± 1.6	7.2 ± 2.5	38.4	<0.0001
R. EHL	1.6 ± 0.8	2.2 ± 0.6	73.5	0.03	1.2 ± 0.9	2.2 ± 0.6	55.1	0.001
R. PL	3.0 ± 1.2	4.8 ± 1.6	61.5	0.0002	1.9 ± 1.4	4.7 ± 1.5	39.5	<0.0001
R. LG	2.3 ± 2.0	3.5 ± 2.5	66.3	0.12	1.5 ± 1.9	3.4 ± 2.4	45.1	0.01
R. MG	6.0 ± 3.3	10.8 ± 3.0	55.8	0.0001	3.8 ± 3.7	10.6 ± 2.9	35.5	<0.0001
R. SOL	12.5 ± 5.0	18.1 ± 5.6	69.1	0.003	7.1 ± 5.1	17.6 ± 5.2	40.2	<0.0001
R. PT	5.1 ± 1.9	8.0 ± 2.3	64.4	0.0002	3.7 ± 2.1	7.8 ± 2.2	46.6	<0.0001
L. TA	3.8 ± 2.0	7.1 ± 2.5	53.1	<0.0001	3.1 2.2	7.0 ± 2.4	43.9	<0.0001
L. EHL	1.6 ± 0.8	2.4 ± 0.9	66.5	0.005	1.1 ± 0.7	2.4 ± 0.9	47.6	<0.0001
L. PL	3.3 ± 1.4	4.7 ± 1.7	70.2	0.008	2.1 ± 1.6	4.6 ± 1.5	46.7	<0.0001
L. LG	2.5 ± 1.8	4.5 ± 2.8	55.3	0.007	1.6 ± 1.6	4.3 ± 2.6	37.4	0.0002
L. MG	6.3 ± 4.2	10.6 ± 3.2	59.5	0.004	4.1 ± 4.7	10.4 ± 3.0	40.1	0.0002
L. SOL	13.7 ± 5.9	16.8 ± 4.8	81.4	0.23	8.2 ± 5.6	16.3 ± 4.4	50.0	<0.0001
L. PT	5.7 ± 2.1	8.0 ± 2.2	71.1	0.004	4.1 ± 2.4	7.8 ± 2.1	52.5	<0.0001
R. CALF	33.9 ± 11.5	54.6 ± 15.9	62.1	<0.0001	21.7 ± 14.8	53.5 ± 15.0	40.7	<0.0001
L. CALF	36.4 ± 15.3	54.1 ± 15.8	67.3	0.002	24.2 ± 17.4	52.8 ± 14.7	45.8	<0.0001
BOTH CALVES	70.3 ± 25.8	108.7 ± 31.6	64.7	0.0002	46.0 ± 31.3	106.3 ± 29.6	43.2	<0.0001

Values are given in cm^2, HSN1=Hereditary Sensory Neuropathy type 1. CSA=cross-sectional area, RMA=remaining muscle area, p=significance value for Mann-Whitney U test or independent sample t-test where appropriate (HSN1 v controls), R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior. Rows highlighted in blue if difference is significant

Remaining muscle area (RMA) was significantly smaller in HSN1 than in controls in all muscles and groups of muscles. In overall bilateral calf, RMA was 46.0 ± 31.3cm^2 in HSN1 and 106.3 ± 29.6cm^2 in controls (p<0.0001). CSA and RMA values in HSN1 and controls at proximal calf level are summarised in Table 5-6.

Of particular note is the fact that despite significant increase in baseline proximal calf FF compared with controls, right lateral gastrocnemius and left soleus demonstrate non-significant muscle atrophy compared with controls. Mean bilateral proximal calf FF versus muscle atrophy (%) is shown in Figure 5-7.

Figure 5-7 – Proximal calf fat fraction versus reduction in cross-sectional area in HSN1 cohort



Each point represents a single muscle. Values are baseline fat fraction (%) plotted against percentage CSA reduction in HSN1 compared with controls. HSN1=hereditary sensory neuropathy. R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, PL=peroneus longus, Sol=soleus, PT=tibialis posterior

5.4.1.4.2 Distal calf

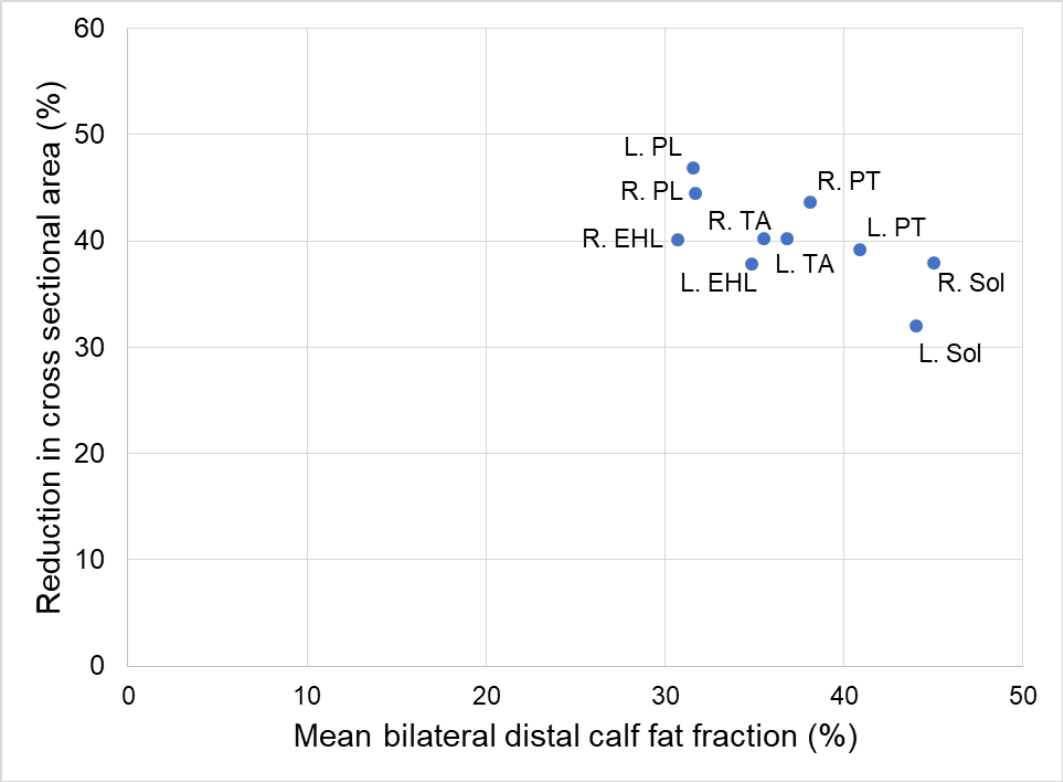
Combined bilateral distal calf CSA was significantly smaller in HSN1 (23.1 ± 10.2cm^2) than in controls (37.5 ± 20.2cm^2), p=0.006. RMA was also significantly smaller in all individual and groups of muscles except left soleus. Once again, it is notable that there was no significant muscle atrophy in right or left soleus despite this muscle being the most fatty infiltrated muscle in the HSN1 cohort, whereas other muscles with lesser FF, such as tibialis anterior and EHL demonstrated significant reduction in CSA. All RMA and CSA values at distal calf level are summarised in Table 5-7, and mean bilateral distal calf FF versus muscle atrophy (%) in Figure 5-8.

Table 5-7 – Baseline distal calf slice cross-sectional area and remaining muscle area in HSN1 and controls

	CSA (cm^2)				RMA (cm^2)			
Muscle	**HSN**	**Control**	**HSN/Control (%)**	**p**	**HSN**	**Control**	**HSN/Control (%)**	**p**
R. TA	0.7 ± 0.5	1.1 ± 0.8	59.8	0.04	0.4 ± 0.4	1.1 ± 0.7	39.7	0.03
R. EHL	1.5 ± 0.8	2.5 ± 0.7	59.9	0.001	1.1 ± 0.8	2.5 ± 0.7	43.7	<0.0001
R. PL	1.3 ± 0.9	2.4 ± 0.9	55.5	0.003	0.9 ± 0.8	2.3 ± 0.9	39.3	0.001
R. Sol	3.7 ± 1.9	6.0 ± 4.6	62.0	0.18	2.2 ± 2.0	5.8 ± 4.4	37.0	0.04
R. PT	4.0 ± 1.4	7.1 ± 3.6	56.3	0.04	2.6 ± 1.7	7.0 ± 3.5	36.7	0.005
L. TA	0.8 ± 0.5	1.3 ± 0.8	59.8	0.04	0.5 ± 0.5	1.2 ± 0.8	39.6	0.03
L. EHL	1.6 ± 0.8	2.6 ± 0.6	62.2	0.004	1.1 ± 0.9	2.6 ± 0.6	43.1	<0.0001
L. PL	1.4 ± 0.9	2.6 ± 0.9	53.1	0.001	1.0 ± 0.8	2.6 ± 0.9	37.7	<0.0001
L. Sol	4.4 ± 3.5	6.5 ± 5.0	68.0	0.47	2.4 ± 2.4	6.3 ± 4.7	38.2	0.09
L. PT	4.4 ± 1.6	7.3 ± 3.7	60.8	0.19	2.7 ± 1.8	7.1 ± 3.5	37.5	0.03
R. CALF	11.2 ± 12.6	19.1 ± 9.3	58.7	0.002	7.2 ± 5.3	18.7 ± 8.9	38.2	<0.0001
L. CALF	12.6 ± 6.8	18.4 ± 11.2	68.4	0.06	7.7 ± 5.8	18.0 ± 10.7	42.4	0.02
BOTH CALVES	23.1 ± 10.2	37.5 ± 20.2	61.5	0.006	14.4 ± 10.5	36.8 ± 19.3	39.0	0.008

p= Mann-Whitney U test or independent samples t-test (where appropriate) for fat fraction change (HSN1 versus controls). R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus/brevis, Sol=soleus, PT=tibialis posterior. Rows highlighted in blue if difference is significant

Figure 5-8 – Distal calf fat fraction versus reduction in cross-sectional area in HSN1 cohort



Each point represents a single muscle. Values are baseline fat fraction plotted against percentage cross-sectional area reduction in HSN1 compared with controls. HSN1=hereditary sensory neuropathy. R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus/brevis, Sol=soleus, PT=tibialis posterior

5.4.1.4.3 Distal thigh

At baseline, CSA was significantly smaller in the HSN1 cohort (162.1 ± 40.4cm^2) compared with controls (203.0 ± 45.2cm^2, p=0.007). There were also significant differences in CSA between HSN1 and controls in right and left thigh, as well as in certain individual muscles. Similarly, RMA was significantly smaller at baseline in the HSN1 cohort: combined bilateral thigh RMA 146.9 ± 47.4cm^2 in HSN1 and 198.8 ± 43.1cm^2 in controls (p=0.002). As per the calf, a number of individual muscles with significant difference between HSN1 and control in baseline FF had not yet developed significant muscle atrophy compared with controls (right BF, SM, and ST. Left VM, BF, SM and AM). Conversely, both left and right rectus femoris showed significantly reduction in CSA despite no change in FF.

All baseline thigh CSA and RMA values in HSN1 and control are given in Table 5-8 and mean bilateral distal thigh FF versus muscle atrophy (%) in Figure 5-9.

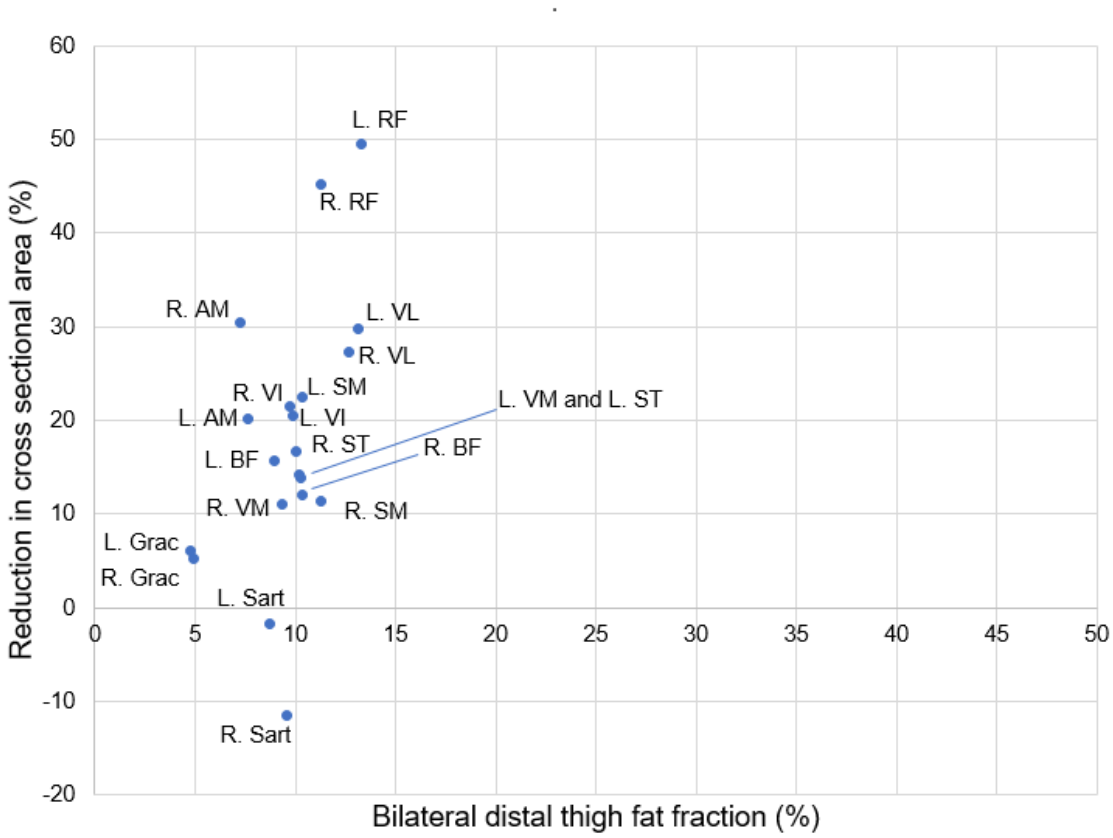
Table 5-8 – Baseline thigh cross-sectional area and remaining muscle area in HSN1 and controls

	CSA (cm^2)				RMA (cm^2)			
Muscle	**HSN**	**Control**	**HSN/Control (%)**	**p**	**HSN**	**Control**	**HSN/Control (%)**	**p**
R. VM	12.0 ± 3.4	13.5 ± 5.1	88.9	0.38	10.9 ± 4.0	13.3 ± 4.9	82.7	0.12
R. VI	13.1 ± 3.2	16.6 ± 3.4	78.5	0.003	11.7 ± 3.8	16.4 ± 3.4	71.2	0.001
R. VL	12.7 ± 3.8	17.5 ± 3.8	72.7	0.001	11.5 ± 4.9	17.3 ± 3.7	66.7	0.001
R. Grac	2.9 ± 0.8	3.1 ± 0.9	94.7	0.56	2.8 ± 0.7	3.0 ± 0.8	91.8	0.36
R. BF	12.3 ± 4.2	14.0 ± 4.8	87.9	0.27	11.0 ± 4.4	13.6 ± 4.5	81.0	0.10
R. SM	7.3 ± 3.0	8.2 ± 3.8	88.6	0.41	6.3 ± 2.8	7.9 ± 3.5	78.9	0.11
R. ST	5.7 ± 2.0	6.9 ± 2.6	83.2	0.14	5.2 ± 2.1	6.8 ± 2.6	76.2	0.04
R. AM	10.7 ± 5.4	15.5 ± 5.2	69.5	0.01	10.1 ± 5.2	15.0 ± 5.1	66.9	0.008
R. Sart	2.9 ± 0.9	2.6 ± 1.2	111.4	0.39	2.6 ± 0.9	2.5 ± 1.1	104.0	0.76
R. RF	2.3 ± 1.2	4.1 ± 1.5	54.9	<0.0001	2.1 ± 1.3	4.1 1.5	52.5	<0.0001
L. VM	11.7 ± 3.6	13.6 ± 4.9	86.1	0.18	10.6 ± 3.9	13.4 ± 4.7	79.0	0.06
L. VI	12.9 ± 3.8	16.2 ± 3.6	79.5	0.02	11.5 ± 4.3	16.0 ± 3.5	72.1	0.003
L. VL	13.0 ± 3.8	18.5 ± 3.5	70.2	<0.0001	11.7 ± 5.0	18.3 ± 3.5	64.0	<0.0001
L. Grac	2.7 ± 0.9	2.9 ± 1.1	93.9	0.60	2.6 ± 0.9	2.8 ± 1.1	90.3	0.40
L. BF	11.9 ± 4.4	14.1 ± 4.5	84.3	0.16	10.8 ± 4.3	13.7 ± 4.4	78.4	0.06
L. SM	6.5 ± 2.7	8.3 ± 3.8	77.5	0.09	5.7 ± 2.6	8.1 ± 3.7	70.3	0.02
L. ST	5.6 ± 2.1	6.5 ± 2.3	85.8	0.22	5.1 ± 2.2	6.4 ± 2.3	79.6	0.10
L. AM	11.1 ± 5.2	14.0 ± 4.6	79.8	0.12	10.4 ± 5.2	13.6 ± 4.8	76.8	0.08
L. Sart	2.6 ± 0.8	2.5 ± 1.1	101.7	0.89	2.4 ± 0.8	2.5 ± 1.1	96.5	0.78
L. RF	2.2 ± 1.2	4.3 ± 1.4	50.5	<0.0001	2.0 ± 1.3	4.3 ± 1.4	47.8	<0.0001
R. THIGH	81.9 ± 20.2	102.0 ± 2.3	80.3	0.01	74.2 ± 24.0	100.0 ± 21.8	74.3	0.003
L. THIGH	80.2 ± 20.6	101.0 ± 22.7	79.4	0.007	72.7 ± 23.7	99.0 ± 21.7	73.5	0.002
BOTH THIGHS	162.1 ± 40.4	203.0 ± 45.2	79.8	0.007	146.9 ± 47.4	198.9 ± 43.1	73.9	0.002

HSN1=Hereditary sensory neuropathy, RMA=remaining muscle area, CSA=cross-sectional area, p= Mann-Whitney U test or independent samples t-test (where appropriate). R=right, L=left, VM=vastus medialis, VI=vastus intermedius, VL=vastus lateralis, Grac=gracilis, BF=biceps femoris, SM=semimembranosus, ST=semitendinosus, AM=adductor magnus, Sart=Sartorius, RF=rectus femoris. Rows highlighted in blue if difference significant

The results of FF and CSA/RMA at the three anatomical locations are in keeping with variable degrees of fatty atrophy of the lower limbs in HSN1 patients.

Figure 5-9 – Distal thigh fat fraction versus reduction in cross-sectional area in HSN1 cohort



Each point represents a single muscle. Values are baseline fat fraction plotted against percentage CSA reduction in HSN1 compared with controls. HSN1=hereditary sensory neuropathy. R=right, L=left, VM=vastus medialis, VI=vastus intermedius, VL=vastus lateralis, Grac=gracilis, BF=biceps femoris, SM-semimembranosus, ST=semitendinosus, AM=adductor magnus, Sart=sartorius, RF=rectus femoris

5.4.1.5 Baseline correlations

5.4.1.5.1 Age and whole muscle FF

There were significant moderate-strong positive correlations between age and FF at all anatomical levels in males and females, except for thigh FF in males (Figure 5-4).

5.4.1.5.2 CMTNS/CMTES and whole muscle FF

There were significant, strong positive correlations between CMTNS and bilateral combined FF at all levels, and similarly between CMTES and bilateral combined FF at all levels. At all levels, strongest correlations were seen at calf levels, and in female HSN1 patients at all levels. In general, and as may be expected, correlations between FF and CMTES were stronger, albeit marginally than correlations between FF and CMTNS (Table 5-9).

Table 5-9 – Correlation matrix between combined bilateral lower limb fat fraction and CMTN/ES

	CMTNS			CMTES		
Cohort	**Distal calf FF**	**Proximal calf FF**	**Distal Thigh FF**	**Distal calf FF**	**Proximal calf FF**	**Distal Thigh FF**
Overall	0.866, p <0.001	0.853, p<0.001	0.758, p<0.001	0.877, p<0.001	0.855, p<0.001	0.764, p<0.001
Male	0.849, p<0.001	0.866, p<0.001	0.617, p=0.005	0.861, p<0.001	0.865, p<0.001	0.623, p=0.004
Female	0.871, p<0.001	0.859, p<0.001	0.782, p<0.001	0.909, p<0.001	0.876, p<0.001	0.804, p<0.001

Values are Spearman rho and associated p value for correlation between combined bilateral calf or thigh fat fraction and CMTNS or CMTES as indicated. CMTNS=CMT neuropathy score, CMTES=CMT examination score, FF=fat fraction

5.4.1.5.3 MRI measure correlations

5.4.1.5.3.1 Proximal and distal calf

At baseline in the HSN1 group, there were moderate negative correlations between CSA and proximal/distal calf FF in some individual muscles and combined muscle groups – these are summarised in Table 5-10. In controls, there were no correlations between baseline FF and other MRI measures.

Table 5-10 – Correlation between baseline proximal and distal fat fraction and cross-sectional area in the HSN1 group in the stated muscle

	CSA (r_s, p value)	
Muscle	**Proximal calf**	**Distal calf**
R. TA	-0.565, p<0.0001	-0.233, p=0.17
R. EHL	-0.409, p=0.005	-0.421, p=0.009
R. PL	-0.334, p=0.03	-0.364, p=0.03
R. LG	-0.244, p=0.11	
R. MG	-0.442, p=0.003	
R. SOL	-0.194, p=0.20	-0.197, p=0.243
R. PT	-0.406, p=0.006	-0.404, p=0.01
L. TA	-0.627, p<0.0001	-0.157, p=0.34
T. EHL	-0.261, p=0.08	-0.327, p=0.045
L. PL	-0.233, p=0.12	-0.411, p=0.009
L. LG	-0.276, p=0.07	
L. MG	-0.474, p=0.001	
L. SOL	-0.082, p=0.60	-0.017, p=0.92
L. PT	-0.299, p=0.049	-0.242, p=0.14
R. CALF	-0.378, p=0.01	-0.421, p=-0.01
L. CALF	-0.313, p=0.03	-0.159, p=0.33
BOTH CALVES	-0.339, p=0.02	-0.324, p=0.04

Correlation coefficient is between baseline fat fraction and cross-sectional area in the stated muscle/muscles. p=significance value for Spearman's rho. CSA=cross-sectional area, R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus (proximal calf), PL=peroneus longus/brevis (distal calf), LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior. r_s=rho Spearman correlation coefficient. Row highlighted in blue if p value <0.05

5.4.1.5.3.2 Distal thigh

There were weak negative correlations between combined bilateral thigh FF and RMA (rho=-0.309, p=0.04) but not with CSA (rho=-0.068, p=0.66). There were also weak negative correlations between left thigh FF and RMA (rho=-0.338, p=0.03).

5.4.2 Longitudinal results

Follow-up imaging was performed at a mean of 370 ± 37.4 days (range 307 to 525 days) in 25 patients with HSN1 and 10 controls. All longitudinal values were annualised for precise comparison and analysis.

5.4.2.1 Fat fraction change

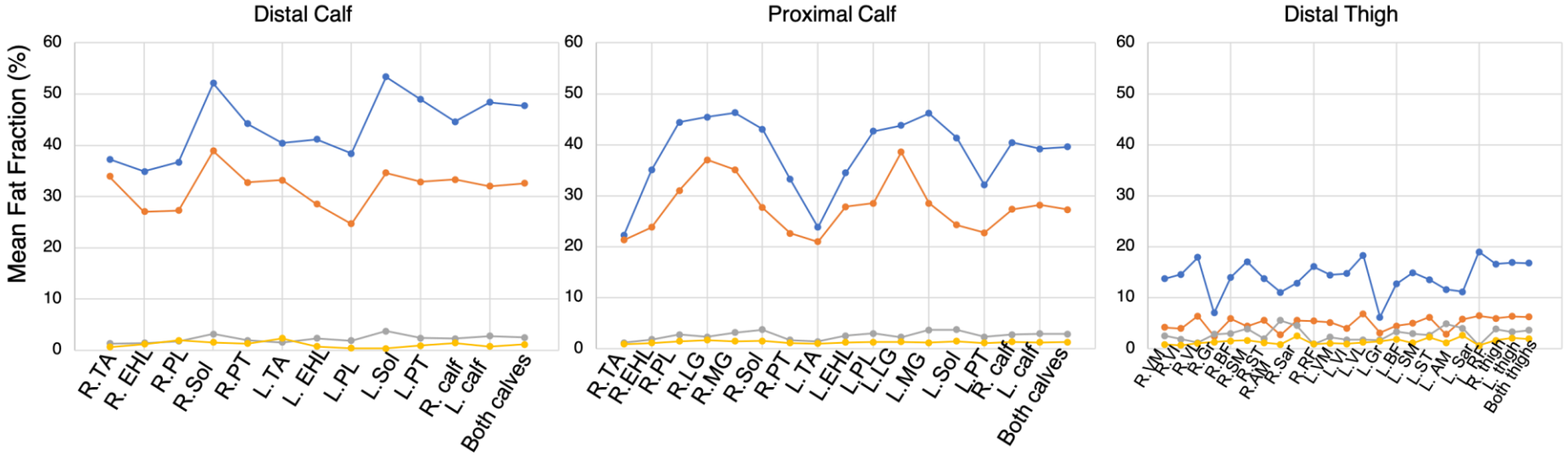
5.4.2.1.1 Proximal calf

In the HSN1 group at follow-up, bilateral combined calf FF was 32.5 ± 30.8% in the HSN1 group and 2.1 ± 1.6% in controls (p<0.0001 – Mann-Whitney U test). In the HSN1 group, there was a significant increase in combined bilateral calf FF of 2.6 ± 3.0% (95% CI 1.3 to 3.9, p=0.0001) above baseline, compared with the control group in which there was no significant change over 12 months (-0.0 ± 0.1%, p=0.42). All individual muscle groups in the HSN1 cohort showed statistically significant FF increase over 12 months compared with baseline, except left tibialis posterior in which there was no significant change. When compared with FF change in controls, there was significant FF change in all individual muscles and muscle groups except right peroneus longus and the left posterior tibial group.

Mean change in combined bilateral proximal calf FF was 3.3 ± 2.9% in males, and 2.0 ± 3.2% in females (p=0.18). The difference between FF change was significant in left tibialis anterior (p=0.04), left medial gastrocnemius (p=0.02), left soleus (p=0.01) and the combined measure left calf (p=0.02) – in all of which FF in the male group changed by significantly more than females.

Mean proximal calf FF change in HSN1 and controls is summarised in Table 5-11, and divided by gender in Table 5-12.

Figure 5-10 – Fat fraction in HSN1 and control by gender



Each point is mean fat fraction (%) for a single muscle in specified group. Lines present for visual aid. Lines colour coded for group: blue=HSN1 males, orange=HSN1 females, grey=male controls, yellow=female controls. R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus (proximal calf), PL=peroneus longus/brevis (distal calf), LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior, VM=vastus medialis, VI=vastus intermedius, VL=vastus lateralis, Grac=gracilis, BF=biceps femoris, SM=semimembranosus, ST=semitendinosus, AM=adductor magnus, Sart=Sartorius, RF=rectus femoris

Table 5-11 – Baseline proximal calf fat fraction and change in fat fraction over 12 months in HSN1 and controls

	HSN		Control				
Muscle	**Baseline FF (%)**	**Change (%)**	**Baseline FF (%)**	**Change (%)**	**p1**	**p2**	**SRM**
R. TA	18.2 ± 21.2	2.4 ± 4.0	1.1 ± 0.5	0.0 ± 0.3	0.002	0.007	0.59
R. EHL	25.2 ± 25.9	4.0 ± 5.5	1.4 ± 0.7	0.2 ± 0.5	<0.0001	0.002	0.72
R. PL	35.1 ± 31.7	2.9 ± 5.7	2.1 ± 1.2	0.1 ± 0.3	0.001	0.13	0.52
R. LG	37.3 ± 37.6	4.5 ± 8.3	2.0 ± 1.2	-0.2 ± 0.7	0.001	0.01	0.55
R. MG	36.8 ± 33.8	3.3 ± 6.1	2.3 ± 1.9	-0.0 ± 0.2	0.001	0.013	0.54
R. SOL	31.1 ± 31.7	1.6 ± 3.7	2.6 ± 2.4	-0.0 ± 0.2	0.01	0.03	0.45
R. PT	25.2 ± 26.7	3.6 ± 4.2	1.5 ± 0.7	0.1 ± 0.2	<0.0001	<0.0001	0.87
L. TA	20.3 ± 23.3	1.7 ± 3.3	1.2 ± 0.7	-0.1 ± 0.2	0.007	0.012	0.52
L. EHL	28.6 ± 28.0	2.0 ± 2.6	1.9 ± 1.5	0.0 ± 0.3	0.001	0.001	0.76
L. PL	32.4 ± 31.2	1.9 ± 3.7	2.2 ± 1.4	-0.0 ± 0.3	0.009	0.02	0.51
L. LG	37.2 ± 34.1	3.5 ± 5.0	1.9 ± 1.3	-0.2 ± 0.5	0.003	0.002	0.71
L. MG	32.7 ± 31.4	3.5 ± 5.0	2.5 ± 2.5	-0.1 ± 0.8	<0.0001	0.002	0.70
L. SOL	31.2 ± 32.0	2.0 ± 2.8	2.6 ± 2.3	-0.1 ± 0.3	0.001	0.001	0.72
L. PT	26.0 ± 27.6	1.2 ± 3.9	1.7 ± 1.2	0.0 ± 0.3	0.08	0.13	0.32
R. CALF	29.8 ± 29.2	3.2 ± 4.4	2.1 ± 1.5	0.0 ± 0.2	<0.0001	0.002	0.71
L. CALF	30.7 ± 30.3	1.8 ± 2.9	2.1 ± 1.7	-0.1 ± 0.2	0.002	0.004	0.61
BOTH CALVES	29.9 ± 29.8	2.6 ± 3.0	2.1 ± 1.6	-0.1 ± 0.1	<0.0001	<0.0001	0.84

Values are baseline fat fraction (%) and change in fat fraction (%) in those participants who had baseline and follow-up scan (n: HSN1=25, control=10). SRM=standardised response mean. p1=Wilcoxon signed ranks test or paired t-test where appropriate (baseline and follow-up in HSN1), p2=Mann-Whitney U test or independent samples t-test (where appropriate) for fat fraction change (HSN1 versus controls). R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL=peroneus longus, LG=lateral gastrocnemius, MG=medial gastrocnemius, Sol=soleus, PT=tibialis posterior. Rows highlighted in blue if both p1 and p2 significant

Table 5-12 – Proximal calf fat fraction change divided by gender

	Male					Female					
Muscle	**Baseline FF (%)**	**Change (%)**	**p1**	**p2**	**SRM**	**Baseline FF (%)**	**Change (%)**	**p1**	**p2**	**SRM**	**p3**
R. TA	19.3 ± 15.1	2.2 ± 4.6	0.14	0.32	0.48	17.3 ± 25.5	2.5 ± 3.8	0.004	0.02	0.67	0.61
R. EHL	33.9 ± 24.4	3.6 ± 2.7	0.001	0.02	1.40	18.4 ± 25.8	4.2 ± 7.1	0.004	0.03	0.59	0.27
R. PL	47.2 ± 27.0	2.5 ± 2.7	0.01	0.09	0.90	25.6 ± 32.7	3.3 ± 7.3	0.12	0.34	0.45	0.41
R. LG	42.6 ± 37.5	3.1 ± 3.5	0.02	0.07	0.88	33.1 ± 38.5	5.8 ± 10.9	0.01	0.04	0.53	0.85
R. MG	43.8 ± 33.1	3.4 ± 3.3	0.01	0.04	1.01	31.3 ± 34.6	3.2 ± 7.7	0.15	0.38	0.41	0.27
R. SOL	40.9 ± 30.4	2.3 ± 3.3	0.01	0.07	0.90	23.4 ± 31.6	0.6 ± 3.7	0.56	0.69	0.16	0.58
R. PT	32.3 ± 23.3	4.3 ± 3.9	0.01	0.04	1.10	19.6 ± 28.7	3.1 ± 4.5	0.005	0.02	0.70	0.47
L. TA	22.2 ± 16.9	2.9 ± 4.5	0.06	0.18	0.64	18.9 ± 28.0	0.8 ± 1.5	0.14	0.16	0.54	0.04
L. EHL	36.6 ± 25.5	2.1 ± 2.7	0.03	0.015	0.76	23.1 ± 29.5	1.8 ± 2.5	0.01	0.07	0.72	0.65
L. PL	44.2 ± 27.9	2.6 ± 3.3	0.03	0.11	0.78	23.2 ± 31.4	1.4 ± 4.1	0.22	0.44	0.34	0.38
L. LG	44.7 ± 31.7	4.1 ± 3.7	0.003	0.02	1.10	30.9 ± 36.0	3.1 ± 6.0	0.14	0.26	0.52	0.11
L. MG	44.9 ± 32.3	4.7 ± 4.2	0.004	0.02	1.12	22.4 ± 27.8	2.4 ± 5.5	0.02	0.09	0.44	0.02
L. SOL	41.9 ± 29.8	3.4 ± 3.2	0.01	0.04	1.06	18.5 ± 29.0	0.8 ± 1.5	0.14	0.19	0.50	0.02
L. PT	34.3 ± 24.6	2.2 ± 5.4	0.20	0.40	0.41	19.0 ± 29.0	0.4 ± 1.6	0.38	0.46	0.25	0.01
R. CALF	38.3 ± 27.7	3.4 ± 3.0	0.004	0.03	1.10	23.0 ± 29.5	3.0 ± 5.4	0.001	0.002	0.56	0.29
L. CALF	39.7 ± 28.2	3.3 ± 2.9	0.003	0.02	1.15	23.6 ± 31.0	0.5 ± 2.3	0.06	0.03	0.24	0.02
BOTH CALVES	39.0 ± 27.8	3.3 ± 2.9	0.003	0.02	1.16	22.8 ± 29.5	2.0 ± 3.2	0.001	<0.0001	0.62	0.18

Baseline fat fraction, mean FF change in HSN1 divided by gender. p1=paired t-test or Wilcoxon sun rank test as appropriate (HSN1), p2=independent sample t-test or Mann Whitney U test (HSN1 v controls), p3=independent sample t-test or Mann Whitney U test (male v female HSN1). Rows highlighted in blue if change significant versus baseline and controls. Cell highlighted in orange if significant difference in change between genders. NB control FF change not shown

5.4.2.1.2 Distal calf

In the HSN1 group, there was significant change in MRI determined FF (Table 5-13). Overall bilateral distal calf FF change was 2.2 ± 2.7% in HSN1 and 0.0 ± 0.5% in controls (p=0.01). Right calf and left calf FF also showed significant change over 12 months: for right calf FF change was 2.5 ± 3.4% in HSN1 and 1.9 ± 1.3% in controls (p=0.004), and for left calf 2.2 ± 2.5% in HSN1 and 0.3 ± 0.6% in controls (p=0.006). Mean change in combined bilateral distal calf FF was 2.3 ± 2.7% in males, and 2.1 ± 2.7% in females (p=0.81). Compared with controls, there was significant FF change only in right soleus in females (p=0.04). The difference between gender groups in FF change was significant only in right soleus in which the male group FF changed by 0.2 ± 3.2% and females by 1.9 ± 2.5% (p=0.046).

Mean distal calf FF change in HSN1 and controls is summarised in Table 5-13 and divided by gender in Table 5-14.

Table 5-13 – Distal calf fat fraction change in HSN1 and controls

	HSN		Control				
Muscle	**Baseline FF (%)**	**Change (%)**	**Baseline FF (%)**	**Change (%)**	**p1**	**p2**	**SRM**
R. TA	32.6 ± 33.3	3.6 ± 7.3	0.9 ± 0.8	0.2 ± 0.6	0.09	0.16	0.49
R. EHL	28.2 ± 28.8	2.2 ± 6.3	1.4 ± 1.4	-0.3 ± 1.6	0.26	0.10.	0.36
R. PL	29.4 ± 28.9	3.3 ± 7.4	2.1 ± 2.6	-1.0 ± 3.0	0.06	0.69	0.44
R. SOL	42.0 ± 37.9	1.2 ± 2.9	2.2 ± 1.8	-0.1 ± 0.8	0.09	0.04	0.4
R. PT	35.7 ± 32.1	3.4 ± 4.9	1.6 ± 0.8	-0.0. ± 0.5	0.03	0.48	0.69
L. TA	33.5 ± 31.4	1.2 ± 5.1	2.1 ± 3.0	-0.0. ± 3.6	0.46	0.57	0.24
L. EHL	29.2 ± 28.9	2.3 ± 3.3	1.5 ± 1.3	0.0 ± 0.6	0.005	0.002	0.71
L. PL	28.9 ± 28.2	0.6 ± 4.9	1.0 ± 1.1	0.6 ± 1.6	0.67	0.47	0.12
L. SOL	38.9 ± 35.8	2.8 ± 4.4	1.7 ± 2.6	1.0 ± 1.3	0.01	0.05	0.63
L. PT	37.1 ± 32.8	2.2 ± 3.9	1.6 ± 1.2	0.1 ± 0.5	0.01	0.04	0.58
R. CALF	35.7 ± 32.8	2.5 ± 3.4	1.9 ± 1.3	-0.3 ± 1.0	0.004	0.004	0.74
L. CALF	35.5 ± 32.1	2.2 ± 2.5	1.5 ± 1.5	0.3 ± 0.6	0.002	0.006	0.87
BOTH CALVES	35.6 ± 31.9	2.2 ± 2.7	1.7 ± 1.2	0.0 ± 0.5	0.003	0.01	0.83

Values are baseline fat fraction (%) and change in fat fraction (%) in those participants who had baseline and follow-up MRI scans (n: HSN1=25, control=10). SRM=standardised response mean. p1=Wilcoxon test or paired t-test where appropriate (baseline and follow-up in HSN1), p2=independent samples t-test or Mann -Whitney U test for fat fraction change (HSN1 versus controls). R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL= peroneus longus/brevis, Sol=soleus, PT=tibialis posterior. Rows highlighted in blue if both p1 and p2 significant

Table 5-14 – Distal calf fat fraction change divided by gender

	Male					Female					
Muscle	**Baseline FF (%)**	**Change (%)**	**p1**	**p2**	**SRM**	**Baseline FF (%)**	**Change (%)**	**p1**	**p2**	**SRM**	**p3**
R. TA	32.0 ± 31.7	2.2 ± 6.7	0.38	0.60	0.33	33.0 ± 31.7	4.7 ± 8.0	0.03	0.24	0.58	0.70
R. EHL	35.5 ± 28.1	1.5 ± 6.4	0.59	0.86	0.23	23.1 ± 29.3	2.9 ± 6.4	0.33	0.66	0.44	0.71
R. PL	32.0 ± 26.4	1.3 ± 7.1	0.60	0.82	0.18	27.6 ± 31.4	4.9 ± 7.6	0.06	0.11	0.64	0.30
R. SOL	49.3 ± 34.3	0.2 ± 3.2	0.86	0.95	0.06	36.9 ± 40.7	1.9 ± 2.5	0.01	0.04	0.78	0.046
R. PT	41.9 ± 30.7	4.2 ± 4.7	0.03	0.20	0.89	31.4 ± 33.6	2.8 ± 5.3	0.33	0.18	0.53	0.60
L. TA	35.1 ± 26.8	2.5 ± 6.4	0.28	0.50	0.39	32.3 ± 35.3	0.3 ± 4.0	0.79	0.94	0.08	0.35
L. EHL	38.0 ± 28.2	2.2 ± 3.6	0.10	0.28	0.62	23.2 28.9	2.3 ± 3.1	0.02	0.17	0.75	0.65
L. PL	33.7 ± 25.4	2.3 ± 5.7	0.25	0.57	0.41	25.6 ± 30.5	-0.7 ± 4.1	0.58	0.47	-0.17	0.17
L. SOL	50.3 ± 34.6	2.5 ± 4.3	0.12	0.36	0.58	31.0 ± 35.8	3.0 ± 4.6	0.06	0.68	0.65	1.00
L. PT	45.0 ± 31.1	3.0 ± 5.3	0.12	0.32	0.57	31.7 ± 33.9	1.6 ± 2.5	0.04	0.35	0.67	0.43
R. CALF	41.5 ± 29.5	2.8 ± 3.2	0.04	0.20	0.89	31.7 ± 34.5	2.3 ± 3.8	0.05	0.08	0.62	0.49
L. CALF	44.2 ± 30.6	2.6 ± 3.0	0.03	0.18	0.87	29.6 ± 33.0	1.9 ± 2.2	0.01	0.27	0.86	0.54
BOTH CALVES	42.9 ± 29.8	2.3 ± 2.7	0.03	0.20	0.85	30.6 ± 33.5	2.1 ± 2.7	0.02	0.17	0.77	0.81

Baseline fat fraction, mean FF change in HSN1 divided by gender. R=right, L=left, TA=tibialis anterior, EHL=extensor hallucis longus, PL= peroneus longus/brevis, Sol=soleus, PT=tibialis posterior. p1=paired t-test or Wilcoxon rank sum test as appropriate (HSN1), p2=independent sample t-test or Mann Whitney U test (HSN1 v controls), p3=independent sample t-test or Mann Whitney U test (male v female HSN1). Rows highlighted in blue if change significant both versus baseline and controls. Cell highlighted in orange if significant difference in change between genders. NB control FF change not shown

5.4.2.1.3 Distal thigh

There was significant change in combined FF measures at 12 months. In combined bilateral thigh FF, change was 0.6 ± 1.2% in HSN1 and -0.1 ± 0.2% in controls (p=0.01). There was significant FF change in several muscles in the HSN1 group compared with baseline, though this change was not significant when compared with matched controls. Mean change in combined bilateral distal thigh FF was 1.0 ± 1.6% in males, and 0.2 ± 0.7% in females (p=0.13). Compared with controls, there was no significant FF change in females and in only three muscles in males. The difference in FF change between gender groups was significant in left vastus lateralis (p=0.02), left adductor magnus (p=0.01), and the combined measure left thigh (p=0.049), in all of which FF change was greater in males.

Mean distal thigh FF change in HSN1 and controls are given in Table 5-15 and divided by gender in Table 5-16.

Table 5-15 – Baseline and change in fat fraction at thigh level in HSN1 and control

	HSN		Control				
Muscle	Baseline FF (%)	Change (%)	Baseline FF (%)	Change (%)	p1	p2	SRM
R. VM	6.2 ± 10.4	0.3 ± 1.8	1.5 ± 1.2	-0.2 ± 0.7	0.55	0.46	0.14
R. VI	6.5 ± 10.0	0.3 ± 1.8	1.3 ± 1.0	0.1 ± 0.4	0.33	0.90	0.17
R. VL	10.0 ± 15.2	0.9 ± 2.1	1.3 ± 0.6	0.2 ± 0.3	0.01	0.06	0.20
R. Grac	4.2 ± 3.7	1.2 ± 1.8	2.1 ± 1.5	0.0 ± 0.6	0.003	0.01	0.50
R. BF	8.2 ± 11.0	0.8 ± 1.1	2.4 ± 2.0	0.0 ± 0.3	0.002	0.03	0.74
R. SM	8.2 ± 10.6	1.1 ± 1.9	2.8 ± 1.9	-0.1 ± 0.4	<0.0001	0.001	0.59
R. ST	8.1 ± 16.3	0.9 ± 2.4	1.7 ± 1.1	0.1 ± 0.2	0.03	0.20	0.39
R. AM	4.7 ± 4.4	0.3 ± 1.5	3.0 ± 4.4	-0.1 ± 0.4	0.35	0.44	0.20
R. Sart	7.7 ± 9.3	0.6 ± 2.5	3.7 ± 2.0	-0.0 ± 0.6	0.09	0.20	0.24
R. RF	9.8 ± 17.6	1.4 ± 3.2	1.1 ± 1.1	0.3 ± 0.8	0.08	0.53	0.40
L. VM	8.4 ± 12.8	1.0 ± 2.0	1.5 ± 1.4	-0.2 ± 0.3	0.02	0.03	0.50
L. VI	6.5 ± 9.7	0.5 ± 1.5	1.2 ± 1.0	-0.1 ± 0.3	0.045	0.045	0.35
L. VL	11.3 ± 16.5	0.9 ± 1.9	1.5 ± 0.8	-0.1 ± 0.2	0.01	0.05	0.49
L. Grac	3.0 ± 4.5	0.3 ± 1.5	1.7 ± 1.1	0.2 ± 0.8	0.11	0.20	0.21
L. BF	6.9 ± 9.6	0.5 ± 0.9	2.3 ± 2.3	-0.4 ± 0.8	0.005	0.002	0.58
L. SM	7.3 ± 10.9	0.3 ± 3.8	2.2 ± 1.5	0.0 ± 0.6	0.17	0.30	0.09
L. ST	8.0 ± 15.2	0.5 ± 1.9	1.6 ± 1.3	-0.5 ± 0.9	0.53	0.24	0.25
L. AM	5.2 ± 5.8	0.7 ± 1.5	2.7 ± 3.4	-0.3 ± 0.4	0.048	0.02	0.50
L. Sart	7.8 ± 9.8	1.5 ± 2.0	3.1 ± 2.2	-0.4 ± 1.3	<0.0001	0.004	0.75
L. RF	12.2 ± 21.3	1.0 ± 2.5	0.6 ± 0.4	-0.0 ± 0.2	0.04	0.14	0.41
R. THIGH	7.0 ± 9.5	0.5 ± 1.2	1.9 ± 1.5	0.0 ± 0.2	0.02	0.09	0.43
L. THIGH	7.2 ± 10.0	0.6 ± 1.2	1.8 ± 1.4	-0.2 ± 0.3	0.03	0.01	0.49
BOTH THIGHS	7.1 ± 9.7	0.6 ± 1.2	1.8 ± 1.4	-0.1 ± 0.2	0.01	0.01	0.47

Values are baseline mean fat fraction (%) and change in fat fraction (%/year) in those participants who had baseline and follow-up MRI scans at thigh level (n: HSN1=27, control=10). SRM=standardised response mean. p1=paired t-test (baseline and follow-up in HSN1), p2= independent samples t-test or Mann-Whitney U test for fat fraction change (HSN1 versus controls). R=right, L=left, VM=vastus medialis, VI=vastus intermedius, VL=vastus lateralis, Grac=gracilis, BF=biceps femoris, SM-semimembranosus, ST=semitendinosus, AM=adductor magnus, Sart=sartorius, RF=rectus femoris Rows highlighted in blue if both p1 and p2 significant. SRM=standardised response mean

Table 5-16 – Distal thigh fat fraction change divided by gender

	Male					Female					
Muscle	Baseline FF (%)	Change (%)	p1	p2	SRM	Baseline FF (%)	Change (%)	p1	p2	SRM	p3
R. VM	7.8 ± 11.0	0.4 ± 2.6	0.64	0.44	0.15	4.5 ± 10.1	0.1 ± 0.9	0.73	0.56	0.17	0.81
R. VI	8.6 ± 10.2	0.6 ± 2.6	0.47	0.61	0.23	4.3 ± 9.4	0.1 ± 0.8	0.78	0.39	0.09	0.34
R. VL	12.6 ± 13.8	1.4 ± 3.0	0.11	0.44	0.48	6.9 ± 15.2	0.5 ± 0.5	0.003	0.10	0.96	0.78
R. Grac	4.2 ± 3.5	1.6 ± 2.5	0.03	0.03	0.65	2.5 ± 2.8	0.8 ± 0.9	0.01	0.22	0.90	0.49
R. BF	8.4 ± 10.6	0.8 ± 0.9	0.02	0.10	0.84	5.5 ± 11.2	0.8 ± 1.3	0.03	0.19	0.66	0.91
R. SM	10.2 ± 12.7	1.6 ± 2.6	0.01	0.01	0.64	4.7 ± 5.5	0.7 ± 1.2	0.01	0.06	0.64	0.43
R. ST	8.8 ± 15.6	1.4 ± 2.7	0.13	0.18	0.53	5.8 ± 13.5	0.5 ± 2.2	0.12	0.89	0.25	0.27
R. AM	6.3 ± 4.2	0.5 ± 2.0	0.39	0.36	0.27	3.0 ± 3.3	0.1 ± 0.9	0.98	0.50	0.11	0.27
R. Sart	8.9 ± 9.0	0.6 ± 3.5	0.16	0.27	0.18	5.8 ± 9.2	0.6 ± 1.6	0.21	0.61	0.35	0.57
R. RF	12.0 ± 15.2	2.5 ± 3.9	0.11	0.38	0.64	5.7 ± 16.4	0.5 ± 2.3	0.46	0.99	0.20	0.40
L. VM	10 ± 12.5	1.7 ± 2.7	0.06	0.19	0.64	5.3 ± 11.6	0.4 ± 0.9	0.11	0.10	0.46	0.11
L. VI	8.2 ± 9.3	0.9 ± 2.1	0.18	0.36	0.44	4.2 ± 9.2	0.2 ± 0.7	1.00	0.18	0.31	0.25
L. VL	14.6 ± 15.1	2.0 ± 2.5	0.01	0.05	0.80	7.1 ± 15.6	0.1 ± 0.6	0.54	0.56	0.17	0.02
L. Grac	2.8 ± 2.4	1.0 ± 1.0	0.01	0.32	0.91	2.6 ± 4.8	-0.2 ± 1.7	0.73	0.90	-0.10	0.07
L. BF	8.7 ± 11.6	0.7 ± 0.7	0.01	0.04	0.98	4.5 ± 6.8	0.3 ± 1.0	0.22	0.07	0.35	0.27
L. SM	9.5 ± 13.0	1.6 ± 3.1	0.13	0.32	0.50	4.9 ± 6.5	-0.6 ± 4.1	0.57	0.74	-0.16	0.16
L. ST	9.5 ± 14.1	0.9 ± 2.5	0.28	0.16	0.34	6.0 ± 14.8	0.2 ± 1.4	0.59	0.66	0.15	0.41
L. AM	6.3 ± 5.8	1.4 ± 1.9	0.03	0.05	0.77	3.1 ± 3.9	0.1 ± 0.9	0.88	0.62	0.11	0.01
L. Sart	6.9 ± 5.7	1.5 ± 1.9	0.02	0.14	0.81	5.8 ± 10.8	1.5 ± 2.5	0.1	0.02	0.69	0.69
L. RF	16.8 ± 21.2	1.5 ± 3.5	0.19	0.35	0.43	6.9 ± 20.5	0.6 ± 1.3	0.05	0.31	0.50	0.40
R. THIGH	8.4 ± 8.8	0.9 ± 1.7	0.19	0.19	0.53	4.9 ± 9.3	0.3 ± 0.7	0.05	0.75	0.37	0.54
L. THIGH	8.8 ± 9.5	1.1 ± 1.5	0.03	0.09	0.75	4.9 ± 9.6	0.2 ± 0.7	0.38	0.27	0.24	0.05
BOTH THIGHS	8.6 ± 9.1	1.0 ± 1.6	0.06	0.12	0.65	4.9 ± 9.4	0.2 ± 0.7	0.10.	0.19	0.32	0.13

Baseline fat fraction, mean FF change in HSN1 divided by gender. p1= paired t-test or Wilcoxon sun rank test as appropriate (HSN1), p2=independent sample t-test or Mann Whitney U test (HSN1 v controls), p3= independent sample t-test or Mann Whitney U test (male v female HSN1). Rows highlighted in blue if change significant versus baseline and controls. Cell highlighted in orange if significant difference in change between genders. NB control FF change not shown

5.4.2.2 Change in muscle cross-sectional area and remaining muscle area

5.4.2.2.1 All levels

There were no significant changes in CSA or RMA in HSN1 or control groups at any level over the course of the study. Combined bilateral CSA and RMA results at each anatomical level for HSN1 and controls are given in Table 5-17.

Table 5-17 – Cross-sectional area and remaining muscle area in HSN1 and Controls

	CSA (cm^2)					
	HSN			Control		
	Baseline	Follow up	p1	Baseline	Follow up	p2
Distal calf	22.3 ± 8.6	21.2 ± 10.0	0.14	41.8 ± 20.5	43.5 ± 23.4	0.28
Proximal calf	67.6 ± 24.6	67.9 ± 25.2	0.83	110.8 ± 32.5	113.6 ± 33.1	0.28
Distal thigh	154.9 ± 34.9	161.0 ± 33.6	0.51	197.4 ± 54.5	186.3 ± 62.1	0.61

	RMA (cm^2)					
	HSN			Control		
	Baseline	Follow up	p1	Baseline	Follow up	p2
Distal calf	14.8 ± 10.9	13.6 ± 11.7	0.06	40.9 ± 19.5	42.5 ± 22.2	0.29
Proximal calf	47.0 ± 30.3	45.6 ± 31.6	0.14	108.3 ± 30.4	111.0 ± 30.7	0.27
Distal thigh	143.4 ± 40.0	146.5 ± 39.2	0.69	196.9 ± 42.0	185.1 ± 52.8	0.61

Values are area ± standard deviation. HSN1=hereditary sensory neuropathy, CSA=cross-sectional area, RMA=remaining muscle area. p1 and p2 =paired t-test in HSN1 and control respectively

5.4.2.3 Standardised Response Mean

5.4.2.3.1 Proximal calf

In the HSN1 group, significant FF change was seen at follow up in quantitative MRI (qMRI) determined FF. Maximum standardised response mean (SRM) across all muscles was 0.87 in the right posterior tibialis group. SRM for combined bilateral calf FF was 0.84. Individual and combined muscle SRM are given in Table 5-11.

5.4.2.3.2 Distal calf

In the HSN1 group, significant FF change was seen in left and right calf FF, combined bilateral distal calf FF as well as in left EHL and right PT. Maximum SRM was 0.87 for left calf FF. Combined bilateral distal calf FF had SRM of 0.83. Individual and combined muscle SRM are given in Table 5-13.

5.4.2.3.3 Distal thigh

In the HSN1 group, significant FF change was seen in left thigh and combined bilateral thigh FF, as well as in several individual muscles. Maximum SRM was 0.75

for left sartorius. Combined thigh FF had SRM of 0.47. Individual and combined muscle SRM are given in Table 5-15.

5.4.2.3.4 SRM for gender groups

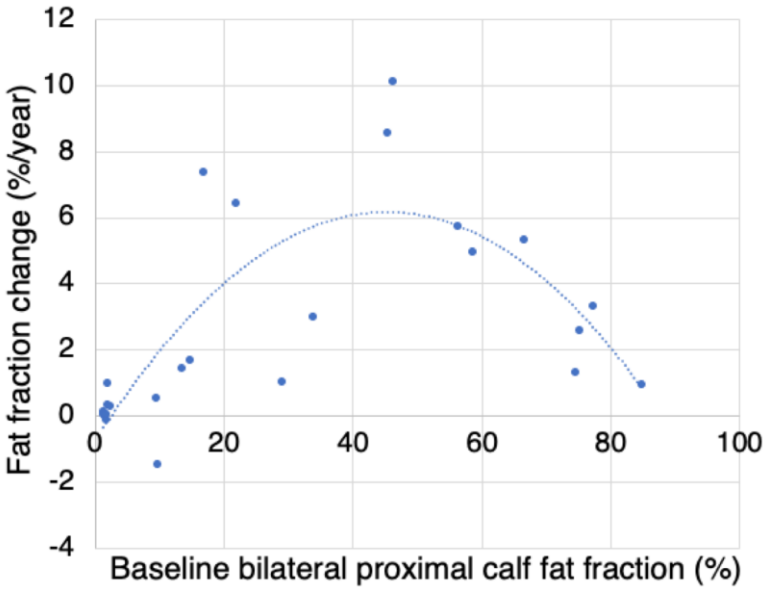
At distal calf level, the only significant FF change was in right soleus in females (FF change 1.9 ± 2.5%, SRM 0.78). At proximal calf level, qMRI measured significant change in both male and female groups in all combined measures, with maximum SRM of 1.16 in males and 0.62 in females (combined bilateral calf). The largest SRM in individual muscles for males was in right EHL (FF change 3.6 ± 2.7 with SRM 1.40) and in females in right tibialis posterior with FF 3.1 ± 4.5 and SRM 0.70. At thigh level, there was no significant change in female FF, and in only three muscles in males, with largest SRM of 0.98 in left biceps femoris.

5.4.2.4 Pattern of change in HSN1 cohort

At all three anatomical locations in HSN1 patients, a similar pattern of fat change was seen with respect to baseline FF, with largest FF change seen in those patients with an intermediate combined baseline FF (Figure 5-11, Figure 5-12 and Figure 5-13) with line of best fit best shown broadly to be a two point polynomial trend line.

5.4.2.4.1 Proximal calf

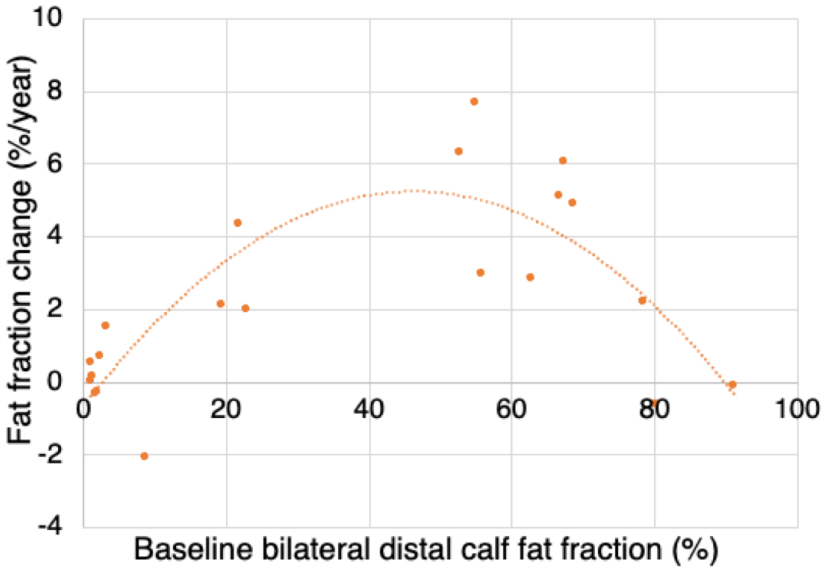
Figure 5-11 – Fat fraction change based on baseline proximal calf fat fraction in HSN1



Each point represents a single patient with HSN1. NB: 2 order polynomial trend line superimposed for visual aid, with formula of best fit: $y = -0.0035x^2 + 0.3158x - 0.7737$

5.4.2.4.2 Distal calf

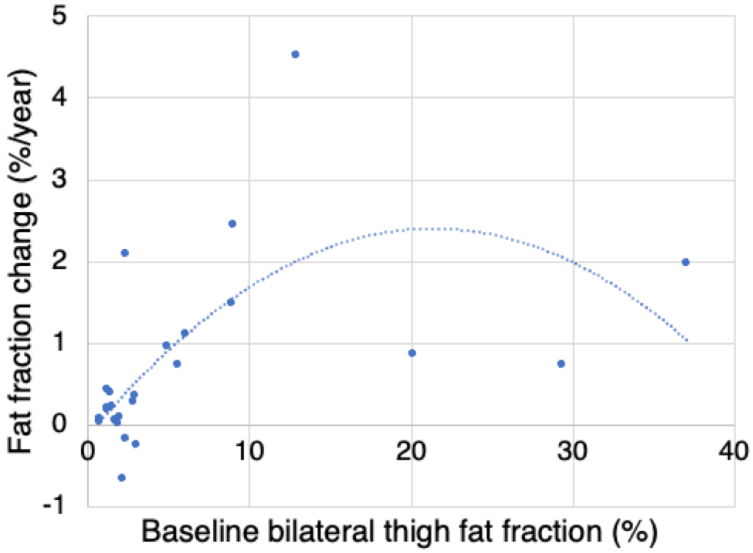
Figure 5-12 – Fat fraction change based on baseline distal calf fat fraction in HSN1



Each point represents a single patient with HSN1. NB: 2 order polynomial trend line superimposed for visual aid, with formula of best fit: $y = 0.0028x^2 + 0.2559x – 0.6532$

5.4.2.4.3 Distal thigh

Figure 5-13 – Fat fraction change based on baseline distal thigh fat fraction in HSN1



Each point represents a single patient with HSN1. NB: 2 order polynomial trend line superimposed for visual aid, with formula of best fit: $y = -0.0056x^2 + 0.2374x – 0.1332$

5.4.2.5 Subgroup analysis

On the basis of the demonstrated association between baseline FF and FF change, as was done in the CMT1 longitudinal natural history study, HSN1 patients were

stratified at each level based on baseline bilateral combined calf or thigh FF thus: group 1 with FF<20%; group 2 with FF between 20% and 70%, group 3 with FF>70%.

5.4.2.5.1 Proximal Calf

Mean combined bilateral calf FF change (%/year) over 12 months in these groups were 0.9 ± 2.1% (p=0.16); 5.8 ± 2.7% (p=0.0005) and 2.1 ± 1.1% (p=0.03) for group 1, 2 and 3 respectively. In group 2 and 3 in which FF change over 12 months was significant, SRM was 1.96 and 1.85, a marked improvement in responsiveness compared with the SRM derived from analysing the HSN1 group as a whole. Subgroup results for proximal calf are summarised in Table 5-18, and FF change against baseline FF is shown in Figure 5-11.

Table 5-18 – Subgroup analysis in HSN1 cohort based on baseline proximal calf fat fraction

Group	Mean baseline FF (%) ± s.d	Mean FF change ± s.d (95% CI)	p	SRM
<20%	6.0 ± 5.9	0.9 ± 2.1 (-0.4 to 2.1)	0.16	0.41
20-70%	44.7 ± 15.6	5.8 ± 2.7 (3.2 to 8.1)	0.0008	1.96
>70%	78.0 ± 4.7	2.1 ± 1.1 (0.3 to 3.8)	0.03	1.85

Group refers to the division based on baseline combined calf fat fraction (FF): group 1 FF<20%, group 2 FF between 20 and 70%, group 3 FF>70%. CI=confidence interval for the change. SRM=standardised response mean, p value is for paired t-test for mean FF change in HSN1 patients (baseline versus follow-up) – values normally distributed in all groups

5.4.2.5.2 Distal Calf

Results for distal calf are show in Table 5-19. As was seen with analysis of the proximal calf slice, the patient group with intermediate baseline FF showed maximal FF change (5.1 ± 1.9%, p=0.001) with a marked improvement in SRM up to 2.65.

Table 5-19 – Mean fat fraction change based on baseline distal calf fat fraction in HSN1

Group	Mean baseline FF (%) ± s.d	Mean FF change ± s.d (95% CI)	p	SRM
<20%	6.0 ± 7.5	0.3 ± 1.2 (-0.7 to 1.2)	0.52	0.22
20-70%	60.3 ± 6.8	5.1 ± 1.9 (3.1 to 7.2)	0.001	2.65
>70%	83.2 ± 6.9	0.5 ± 1.5 (-3.2 to 4.2)	0.62	0.33

FF=fat fraction, s.d=standard deviation, HSN1=Hereditary Sensory Neuropathy, p value is for paired t test in HSN1, SRM=standardised response mean

5.4.2.5.3 Distal Thigh

In patients with baseline FF <20%, there was a significant increase in FF over 12 months (0.6 ± 1.0%, p=0.01) compared with baseline, and an improvement in SRM from 0.47 (whole HSN1 group analysis at thigh level) to 0.58. There was no significant change in FF in other subgroups (Table 5-20).

Table 5-20 - Mean fat fraction change based on baseline distal thigh fat fraction in HSN1

Group	Mean baseline FF (%) ± s.d	Mean FF change ± s.d (95% CI)	p	SRM
<20%	3.5 ± 3.2	0.6 ± 1.0 (0.2 to 1.2)	0.01	0.58
20-70%	28.9 ± 8.5	5.8 ± 2.7 (3.2 to 8.1)	0.1	1.75
>70%	NP			

FF=fat fraction, s.d=standard deviation, p value is for paired t test in HSN1 patients; SRM=standardised response mean. NP=no patients. Row highlighted in blue if significant change

5.4.2.6 Stratification based on gender

As for the overall HSN1 cohort, male and female groups were stratified according to baseline FF <20%, 20-70% and >70%. Results are shown in Table 5-21. As seen in analysis of the whole HSN1 cohort, greatest change at both calf levels was seen in the cohort with baseline intermediate FF. There was marked improvement in responsiveness using this method, with maximum SRM seen in intermediate muscles in both groups. In females, greatest SRM was for distal calf FF change: 6.4 ± 1.3% (SRM=4.96) and in males, in proximal calf FF change: 5.1 ± 1.3% (SRM=3.93).

Table 5-21 – Baseline and fat fraction change according to gender and baseline FF

Slice	Group	Gender	Mean baseline FF (%) ± sd	Mean FF change ± sd	p	SRM
Distal Calf	**<20%**	**Male**	4.8 ± 5.3	-1.0 ± 1.5	0.51	-0.68
		Female	1.9 ± 0.8	0.4 ± 0.7	0.23	0.56
	20-70%	**Male**	59.4 ± 20.2	3.4 ± 2.3	0.01	1.47
		Female	58.1 ± 7.4	6.4 ± 1.3	0.01	4.96
	>70%	**Male**	79.3 ± 1.3	0.8 ± 2.0	0.68	0.39
		Female	ID			
Proximal calf	**<20%**	**Male**	10.4 ± 6.5	1.8 ± 3.9	0.43	0.46
		Female	4.0 ± 4.7	0.5 ± 0.6	0.04	0.83
	20-70%	**Male**	47.4 ± 18.8	5.1 ± 1.3	0.001	3.93
		Female	40.2 ± 9.7	6.6 ± 4.9	0.14	1.35
	>70%	**Male**	75.0 ± 0.5	2.0 ± 0.9	0.20	2.15
		Female	81.1 ± 5.4	2.1 ± 1.7	0.32	1.28
Distal Thigh	**<20%**	**Male**	5.0 ± 4.3	1.3 ± 1.6	0.04	0.79
		Female	2.4 ± 1.7	0.1 ± 0.4	0.56	1.42
	20-70%	**Male**	24.7 ± 6.5	-0.1 ± 1.1	0.95	-0.10
		Female	ID			
	>70%	**Male**	ND			
		Female	ND			

ND=no data, ID=inadequate data, p=p value for paired t-test (HSN1), SRM=standardised response mean

5.4.2.6.1 Conclusion of subgroup analysis

Across both calf sites, highest SRM was seen in the intermediate group with baseline FF between 20% and 70%. In the thigh group, maximum significant change was seen

in the group of patients with FF<20%. There were further improvements in SRM when dividing the HSN1 group by gender. The effect of stratification is to homogenise the mean change and standard deviation of change, thus improving responsiveness of the outcome measure.

5.4.2.7 Refinement in responsiveness of quantitative MRI

Based on these subgroup analyses at each anatomical level, and findings from the CMT1A natural history study, data was analysed using further methods of longitudinal slice selection.

5.4.2.7.1 Severity specific slice selection

5.4.2.7.1.1 Analysis of distal calf slice in HSN1 cohort with proximal combined bilateral calf fat fraction <20%

Subgroup analysis of distal calf slice in those thirteen HSN1 patients in whom baseline combined bilateral proximal calf FF was <20%, revealed a significant FF change of 1.7 ± 2.4%. Results are summarised in Table 5-22 and spread of FF change for the entire HSN1 cohort at distal calf level shown in Figure 5-12.

Table 5-22 – Baseline and fat fraction change for combined bilateral proximal and distal calf in HSN1 patients

	Proximal calf		Distal calf	
HSN patient	**Baseline FF (%)**	**Change (%)**	**Baseline FF (%)**	**Change (%)**
1	1.7	-0.1	1.1	0.0
2	2.0	0.4	3.2	1.5
3	1.5	0.0	1.2	0.1
4	1.4	0.1	2.3	0.7
5	9.4	0.5	19.3	2.1
6	14.6	1.7	54.8	7.7
7	1.3	0.1	1.1	0.5
8	13.6	1.4	x	x
9	1.6	0.1	1.6	-0.3
10	1.9	1.0	1.9	-0.3
11	16.9	7.4	21.7	4.4
12	2.4	0.3	22.8	2.0
13	9.6	-1.5	x	x

Each row represents fat fraction values (baseline FF and FF change) from a single HSN1 patient at proximal and distal calf level. FF=fat fraction

Using distal calf slice FF for analysis in these patients in place of proximal calf slice FF, results in a mean FF change of 3.1 ± 3.0%, and SRM of 1.04, a considerable improvement in responsiveness compared with the responsiveness of using proximal slice FF as the outcome measure across the whole cohort (FF change 2.6 ± 3.0%, SRM 0.84), despite the fact that some patients with proximal calf level FF<20% also had distal calf FF <20%. This suggests that the 'leading edge' of length-dependent

pathology is yet further distally in a number of these patients, raising the question of imaging even further distally. Analysis of foot muscle FF is ongoing in our centre. Herein lies one of the problems with severity dependent slice selection – the difficulty in predicting when and where a rapid phase of fat fraction change will occur.

5.4.2.7.1.2 Analysis of distal thigh slice in subgroup with baseline combined bilateral proximal calf fat fraction >70%

In the four patients with baseline proximal calf FF >70%, mean distal thigh FF ± s.d was 23.3 ± 13.3%. There was a significant change in FF over the study period in this subgroup (FF change 1.2 ± 0.6%, p=0.03)

5.4.2.7.2 Combining anatomical sites – a composite outcome measure

5.4.2.7.2.1 Combining all three anatomical sites in each participant

Calculating a mean FF across all three anatomical sites in each participant revealed a baseline mean FF of 23.2 ± 20.9% in HSN1 versus 2.0 ± 1.5% in controls, and follow up mean FF of 25.5 ± 21.6% in HSN1 and 1.9 ± 1.4% in controls. FF change was 2.3 ± 3.3% in HSN1 (paired t-test 0.002) and -0.1 ± 0.2% in controls (independent samples t-test p=0.03). SRM of 0.69 by this method of slice selection.

If the thigh slice is omitted, thus negating some of the muscle with no change over the study period, FF changes from 32.2 ± 20.9% at baseline to 35.5 ± 30.0% at 12 months with FF change of 3.1 ± 4.1% (p=0.0007) in HSN1, and -0.1 ± 0.3% in controls (independent samples t-test p=0.02). SRM for this method of slice selection is 0.77.

5.4.2.7.2.2 Slice selection based on baseline proximal calf fat fraction

Of the 25 HSN1 patients who underwent serial imaging:

- eight had overall bilateral proximal calf FF between 20 and 70%
- thirteen had overall bilateral proximal calf FF <20%, of which eleven had distal imaging which could be analysed.
- four had overall bilateral proximal calf FF >70%.

For patients with overall bilateral proximal calf FF <20% or >70%, distal calf or distal thigh slice were used for serial assessment respectively, combining these with the proximal calf slice from the eight patients with baseline FF between 20 and 70%. This resulted in a mean baseline FF of 25.3 ± 21.4% with highly significant mean FF change of 3.00 ± 3.06% (p=0.0001) and SRM of 0.97.

This compares with FF change of 2.2 ± 1.7%/year, 2.1 ± 1.4%/year and 0.6 ± 1.2%/year with SRM of 0.83, 0.84 and 0.47 when solely distal calf, proximal calf or distal thigh slice respectively are used across the cohort.

This method of longitudinal slice selection has increased the mean change, though the standard deviation has also increased, albeit proportionally less. It should be noted that in one patient in whom the distal thigh slice was used, FF was 6.1%, and in seven patients in whom the distal calf slice was used, FF was <3.2%. This suggests that for the former, the leading pathological edge of disease lies a little more distally in the thigh, and for the latter, more distally in the foot. It is clear that if these patients were also captured at the anatomical location of maximum change, SRM would be again vastly improved. As an example, removal of the seven patients in whom combined bilateral distal calf FF <3.2% results in further improvement in SRM from 0.97 to 1.39. Baseline FF and FF change for HSN1 patients by this method are listed in Table 5-23.

Table 5-23 – Baseline fat fraction and fat fraction change in HSN1 group, using differing anatomical location dependent on baseline proximal calf slice fat fraction

Slice location	Baseline FF (%)	Change (%)
Distal thigh	6.1	1.1
	29.3	0.7
	37.1	2.0
	20.1	0.9
Proximal calf	45.3	8.6
	33.8	3.0
	66.6	5.3
	21.9	6.4
	29.0	1.0
	56.4	5.7
	58.6	5.0
	46.3	10.1
Distal calf	1.1	0.0
	3.2	1.5
	1.2	0.1
	2.3	0.7
	19.3	2.1
	54.8	7.7
	1.1	0.5
	x	x
	1.6	-0.3
	1.9	-0.3
	21.7	4.4
	22.8	2.0
	x	x

Baseline FF ± s.d	25.3 ± 21.4
Mean FF change	3.00
s.d of change	3.06
SRM	0.97

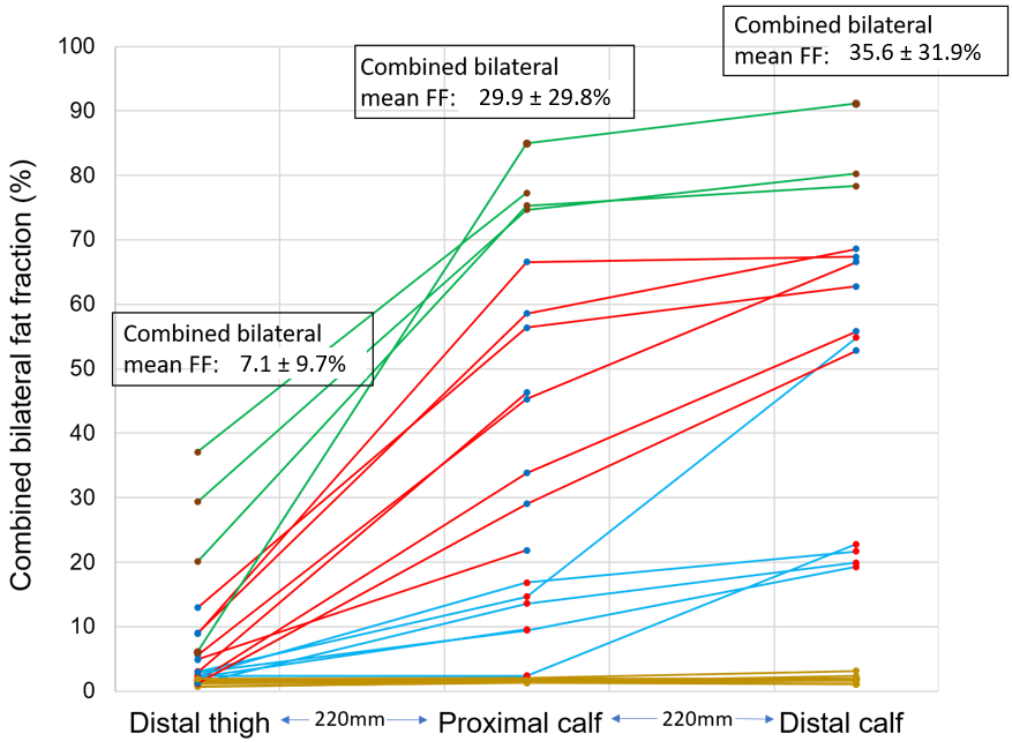
Results are baseline fat fraction (%) and fat fraction change (%/year). Cells are colour coded for anatomical location: blue=distal thigh, red=proximal calf, green=distal calf. s.d=standard deviation, FF=fat fraction, SRM=standardised response mean

5.4.3 Gradient assessment in HSN1 cohort

Figure 5-14 summarises combined bilateral limb FF at distal thigh, proximal calf and distal calf levels. Several points are noted. There are clear length-dependent proximal-distal muscle fat gradients in this cohort of HSN1 patients, with distal calf most heavily affected, followed by proximal calf and then distal thigh. Generally, distal thigh slice starts to become affected once distal calf slice reaches a FF of between 50-60%, or proximal slice FF of ~40%. The mean FF is skewed by many more patients with low/normal baseline FF. There is not a proportional increase in FF across the lower limb in keeping with intermediate FF predicting more rapid FF change.

Although lines between points in Figure 5-14 suggest a linear increase in FF between levels, this is unlikely to be the case based on results in CMT1A presented in Chapter 4. Lines are for visual aid only.

Figure 5-14 – Combined bilateral fat fraction (calf or thigh as indicated) at three anatomical locations in HSN1, (n=25)



Each line represents a single patient with HSN1. Lines are colour coded for baseline proximal calf fat fraction: <20%=blue, 20-70%=red, >70%=green. All levels normal=yellow. FF=fat fraction. Distance between slices is indicated in millimetres. NB – two patients' distal axial MRI slice was affected by artefact and not analysed

5.5 Discussion

Hereditary Sensory Neuropathy type 1 is an inherited peripheral neuropathy with the potential to cause significant morbidity, and for which a possible treatment is available. The rarity and phenotypic heterogeneity of HSN1 translates to the need for a highly responsive outcome measure for future clinical trials.

This prospective natural history study is the first to analyse intramuscular fat accumulation in patients with HSN1, and demonstrates the 3-point Dixon method of FF calculation to be a highly responsive outcome measure over 12 months in a cohort of patients with clinically heterogeneous HSN1. In line with the central strand of this thesis, methods are assessed by which outcome measure responsiveness can be further improved in order to allow adequately powered clinical trials of feasible duration.

5.5.1 Cross-sectional assessment

5.5.1.1 Lower limb fat infiltration

We have demonstrated fat infiltration and muscle atrophy in the lower limbs of a cohort of patients with clinically heterogeneous HSN1. This is an important finding, not least because even though HSN1 is a predominantly sensory neuropathy, qMRI provides a measure of disease activity which can be used for cross-sectional and longitudinal assessment.

5.5.1.2 Baseline qMRI measures and pattern of muscle fat infiltration

Assessment of the qMRI pattern of lower limb involvement demonstrates a number of novel findings relating to chronic fatty atrophy in this disease.

At proximal calf level, there is an interesting pattern of muscle involvement. In both legs, tibialis anterior is the least fat infiltrated muscle, with peroneus longus, medial and lateral heads of gastrocnemius and soleus the most affected. This finding contrasts with the distribution reported in CMT1A (Chapter 3 and 4) in which soleus and lateral gastrocnemius are amongst the least affected calf muscles. In both inherited neuropathies, the tibialis posterior group is relatively preserved.

At distal calf level, the difference between muscles in the HSN1 cohort is less marked, due to the more severe muscle fat involvement at that level, however soleus remains the most fat infiltrated muscle with peroneus longus least affected. At distal thigh level, the pattern is more difficult to assess due to fewness of participants with proximal fat infiltration. Vasti are more affected than posterior thigh muscles, and vastus lateralis significantly more fatty infiltrated than gracilis – the least affected muscle.

We have found consistency between the lower limbs, with no significant right/left FF difference at any anatomical level examined.

Of particular note, and mirroring results reported in the CMT1A cohort (Chapter 3) is the finding that several individual muscles with significant fat infiltration at baseline compared with controls, have non-significant reduction in CSA when compared with controls. This finding reflects the differing response of individual muscles to chronic denervation with varying degree of temporal disconnect between fatty infiltration and muscle atrophy. As may be expected, this finding is most notable at thigh level, where it occurs in seven individual muscles. As one proceeds further caudally, to proximal and then distal calf level, this dichotomy occurs much less often, presumably due to the fact that calf muscle has had more time to atrophy than thigh muscle. At distal and proximal calf levels, where all muscles have significant fat infiltration compared with controls, only soleus, and left tibialis posterior (at distal calf level) and left soleus and right lateral head of gastrocnemius (at proximal calf level), have a non-significant difference in CSA compared with controls, despite these being amongst the muscles with greatest FF at both levels. This point is further emphasised by the lack of change in CSA or RMA over the study period despite significant changes in FF.

5.5.1.3 Gender

CMTNS and CMTES results reveal significantly more severe disease in HSN1 males than females. This is reflected in the lower limb FF findings: although bilateral combined FF was increased in males compared with females at all anatomical levels, this difference did not reach statistical significance at distal calf level, and only did so in two individual muscles at proximal calf level. At thigh level however, there was significantly higher FF in HSN1 males than females in combined measures and several individual muscles.

The pattern of muscle fat involvement is the same in males and females, paralleling that which is reported in the entire cohort.

5.5.1.4 Muscle fat gradients

As clinically expected, presence of muscle fat gradients are confirmed in this study (Figure 5-14). This once again affirms, as discussed in detail in Chapter 4, the critical importance of precise longitudinal slice selection. Even with axial slices of 5mm without a gap, there is a risk of measurement error in excess of the measurable real change – an emerging problem for longitudinal imaging studies.

Accurate computerised muscle segmentation will be an important development allowing multiple axial slices to be segmented for analysis, though this technique cannot be implemented until shown to be highly correlated with the current gold standard of manual muscle segmentation. This is further discussed in Chapter 6.

5.5.1.5 Cross-sectional correlations

Cross-sectional results confirm high criterion validity of qMRI FF, by strong positive correlations between CMTES/CMTNS and qMRI determined FF at all anatomical levels. At all levels, there was stronger correlation between FF and RMA, than between FF and CSA, reflecting the former being a more accurate assessment of fatty atrophy.

5.5.2 Longitudinal assessment

5.5.2.1 Change in fat fraction and other MRI measures

Of central importance to this thesis is the demonstration that qMRI determined FF by the 3-point Dixon method measures highly significant FF change over a period of 12 months at all three anatomical levels examined: distal calf, proximal calf and distal thigh, in a cohort of patients with clinically heterogeneous HSN1. Overall bilateral lower limb slice FF change in proximal calf, distal calf and distal thigh was 2.6 ± 3.0%/year (paired t-test $p<0.0001$, two tailed t-test $p<0.0001$), 2.2 ± 2.7%/year (paired t-test $p= 0.003$, two-tailed t-test 0.001) and 0.6 ± 1.2%/year (paired t-test $p=0.01$, two-tailed t-test $p=0.001$) respectively. Similarly, significant FF change over the study period is measured by qMRI in right and left calf at both levels, and in left, but not right distal thigh. At proximal calf level, FF in almost all muscles changed significantly over the study period, whereas there were fewer such muscles at both thigh and distal calf level, showing that overall in this HSN1 cohort, the active pathological process is most prominent at proximal calf level.

There were no significant changes in overall CSA and RMA in HSN1 or controls over the study period which indicates a slowly progressive atrophic process undetectable over short trial durations. It is unlikely that muscle CSA will be useful as an outcome measure in future clinical trials.

5.5.3 Quantitative MRI outcome measure responsiveness

Of critical importance is the finding that alongside a statistically significant change in FF measured by MRI, qMRI determined FF is shown to be a highly responsive outcome measure in a second inherited peripheral neuropathy. Table 5-24 summarises baseline FF, FF change and SRM for all methods of longitudinal slice selection examined in this chapter. Each is discussed in turn below.

5.5.3.1 Overall bilateral lower limb fat fraction

Using overall bilateral lower limb FF at a single level as an outcome measure (at any three of the analysed anatomical levels), qMRI is shown to have large responsiveness at distal calf and proximal calf, and minimal responsiveness at distal thigh level with SRM of 0.84, 0.83 and 0.47 respectively. The calculated SRM values at both calf levels are comparable to the results of data from the CMT1A natural history study published by this

centre, in which calf FF change was 1.2 ± 1.5%/year with SRM of 0.83 (Morrow et al., 2016).

At all imaged levels, FF change has been shown to be related to baseline FF in the HSN1 group. Based on this observation, we have demonstrated the ability to markedly improve qMRI FF responsiveness by patient group homogenisation.

5.5.3.2 Combination of anatomical levels

When all three anatomical levels are combined with an overall unweighted mean value for each participant, as may be expected there is improvement in the SRM to 0.69 compared with assessment at thigh level alone, however a reduction in SRM compared with proximal or distal calf analysis. The latter is due to the large number of 'normal' thigh scans in the HSN1 group (20/34=58.8%) which changed little over the study duration, thus diluting the mean FF change and increasing standard deviation. When excluding the distal thigh slice from analysis, there is improvement in SRM to 0.77, though this remains less responsive when proximal or distal calf are analysed alone (SRM=0.84 and 0.83 respectively).

5.5.3.3 Single limb versus both lower limbs

There was significant FF change in all single limb slice FF except in right distal thigh. One may expect using the combination of both limbs to result in an SRM somewhere between the SRM of the two individual limbs, as was the case when combining the three anatomical sites, and although this is the case at distal calf level: right calf FF change 2.5 ± 3.4%/year (SRM=0.74), left calf FF change 2.2 ± 2.5%/year (SRM=0.87), overall FF change 2.2 ± 2.7 (SRM=0.83) it is not so at the other two levels, given that combined calf or thigh FF is weighted for CSA. In the proximal calf, right calf FF change was 3.2 ± 4.4%/year (SRM=0.71) and in the left calf 1.8 ± 2.9%/year (SRM=0.61), compared with overall distal calf FF change of 2.6 ± 3.0%/year (SRM=0.84). At distal thigh level, left thigh FF changed by 0.6 ± 1.2 (SRM=0.49) compared with overall bilateral thigh FF change of 0.6 ± 1.2 (SRM=0.47). There were no significant differences between the change in right versus left leg.

These results show that using both limbs is not necessarily superior in outcome measure responsiveness to a single limb. As per findings in the CMT1A cohort (Chapter 3), we therefore suggest analysis of a single limb for longitudinal studies, unless marked clinical asymmetry is apparent. In addition to halving analysis time, this affords several advantages including the possibility of higher resolution acquisition in a similar duration imaging slot (if only one leg is imaged), and inclusion of patients with below knee amputations, an unfortunately frequent complication of HSN1.

5.5.3.4 Single muscle analysis

Interestingly, measurement of FF in some individual HSN1 muscles was highly responsive to change, and in the thigh, greatly exceeding the SRM of bilateral thigh FF. At distal calf level, no muscle exceeded the overall bilateral calf FF SRM of 0.84: left extensor hallucis longus showed greatest responsiveness with FF change of 2.3 ± 3.3%/year (SRM=0.71). At thigh level, FF in left sartorius changed by 1.5 ± 2.0%/year (SRM=0.75).

An interesting finding was that at proximal calf level, FF in right tibialis posterior changed by 3.6 ± 4.2%/year (SRM=0.87) whereas FF in left tibialis posterior showed no FF change at the same level (p=0.04 for difference between right and left posterior tibialis change) despite almost identical baseline FF. The reason for this disparity is not clear – through factors driving a more rapid FF change in any individual muscle are clearly complex.

Certain muscles such as soleus, showed no significant change or little change over 12 months. We hypothesise that in certain muscles, change occurs quite quickly – over several years, whereas in others it occurs over many more years and subsequently less responsive as outcome measures, though this needs to be further examined in longer natural history studies.

5.5.3.5 Stratification of patients based on fat fraction at baseline

At both proximal and distal calf level, the largest SRM was seen in that subgroup of patients with intermediate baseline combined calf FF (between 20% and 70%). In that subgroup at distal calf level, there was significant change in FF of 5.1 ± 1.9%/year, paired t-test p=0.001 (SRM=2.65), and at proximal calf level, FF change of 5.8 ± 2.7%/year, paired t-test p=0.0008 (SRM=1.96). Similarly, patients with baseline proximal calf FF >70% had FF change of 2.1 ± 1.1%/year (SRM=1.85). The effect through homogenising the group is to increase the mean change over time by capturing those intermediate FF patients (at proximal calf level, mean change increased from 2.6 to 5.8%), and to reduce the standard deviation of the change (at proximal calf level, standard deviation decreased from 3.0 to 2.7%). The effect is further enhanced when combining gender and intermediate baseline FF, leading to the greatest study SRM of 4.96 (for combined bilateral distal calf FF in females with intermediate baseline distal calf FF).

Although this concept is of great interest to trial design in rare diseases, and could certainly be capitalised on in rare disorders, as discussed in Chapters 1 and 3 it may not be practical to limit clinical trial participation to a certain subset of patients in a disease such as HSN1, in which numbers are few, recruitment difficult and attrition rates not negligible. In addition, treatments which may have greatest effect on those with early disease will be missed, and a study's external validity affected.

5.5.3.6 Severity specific slice selection

A compromise therefore is to include all participants with longitudinal slice for analysis at that anatomical location at which FF is intermediate, and where change may be expected to be greatest and homogenous across the group, thus excluding the 'ceiling' or 'floor' effect seen with other outcome measures.

In our analysis, such an approach led to an overall FF change across the HSN1 group of 3.0 ± 3.1%/year (paired t-test p=0.0001) with SRM of 0.97, an improvement on all other qMRI determined outcome measures which included all HSN1 patients, and only bettered by subgroup analysis.

Of importance in interpreting this result is that in this HSN1 cohort, a majority of patients (7/13 – 54%) with baseline overall bilateral proximal calf level FF <20%, also had baseline overall distal calf level FF <20% - with subsequent little change over 12 months at both calf levels (floor effect). Conversely, one patient with baseline overall bilateral proximal calf level FF >70%, had baseline overall bilateral distal thigh FF of <10% (ceiling effect). This is in keeping with the length-dependent pathology being in the foot in the former, and further distally in the thigh or more proximal calf in the latter. Further optimisation of qMRI responsiveness would be attained if the anatomical location with intermediate FF were selected as baseline in each participant – this would necessitate careful trial design.

With this in mind, development of qMRI of the foot is underway in our Centre, as a critical next step. This is likely to take some time and must needs include demonstration of scan-rescan reliability in healthy controls and determination of normative data for individual foot muscles across the spectrum of age, gender, size and other possible confounders, and then to show validity in patients, as has been done for calf and thigh imaging. This is further discussed in Chapter 6.

Table 5-24 – Baseline fat fraction, fat fraction change and standardised response mean by different methods of longitudinal slice selection in HSN1 group

	Slice selection	Details	n	Mean baseline FF ± sd	FF change ± sd (95% CI)	p1	p2	SRM
Entire group	Bilateral combined	Distal Calf	22	35.6 ± 31.9	2.2 ± 2.7 (1.0 to 3.4)	0.003	0.01	0.83
		Proximal Calf	25	29.2 ± 29.8	2.6 ± 3.0 (1.3 to 3.9)	<0.0001	<0.0001	0.84
		Distal Thigh	25	7.1 ± 9.7	0.6 ± 1.2 (0.1 to 1.1)	0.01	0.01	0.47
		Three anatomical sites	22	23.3 ± 20.9	2.3 ± 3.3 (0.9 to 3.7)	0.002	0.03	0.69
		Combined distal and proximal	22	32.2 ± 20.9	3.1 ± 4.1 (1.5 to 4.8)	0.001	0.02	0.77
		Severity specific slice	23	25.3 ± 21.4	3.0 ± 3.1 (1.6 to 4.3)	0.0001	na	0.97
	Partial slice - proximal calf	Right Calf	25	29.8 ± 29.2	3.2 ± 4.4 (1.3 to 5.0)	<0.0001	0.002	0.71
		Left Calf	25	30.7 ± 30.3	1.8 ± 2.9 (0.6 to 3.0)	0.002	0.004	0.61
		Right Tibialis Posterior	25	25.2 ± 26.7	3.6 ± 4.2 (1.9 to 5.4)	<0.0001	<0.0001	0.87
		Right EHL	25	25.2 ± 25.9	4.0 ± 5.5 (1.7 to 6.2)	<0.0001	0.002	0.72
	Partial slice - distal calf	Right Calf	22	35.7 ± 32.8	2.5 ± 3.4 (0.9 to 4.2)	0.004	0.004	0.74
		Left Calf	22	35.5 ± 32.1	2.2 ± 2.5 (1.0 to 3.4)	0.002	0.006	0.87
		Left EHL	22	29.2 ± 28.9	2.3 ± 3.3 (0.8 to 3.8)	0.005	0.002	0.71
		Left Tibialis Posterior	22	37.1 ± 32.8	2.2 ± 3.9 (0.4 to 4.2)	0.01	0.04	0.58
	Partial slice - distal thigh	Right Thigh	25	7.0 ± 9.5	0.5 ± 1.2 (0.02 to 1.1)	0.02	0.09	0.43
		Left Thigh	25	7.2 ± 10.0	0.6 ± 1.2 (0.1 to 1.1)	0.03	0.01	0.49
		Left Sartorius	25	7.8 ± 9.8	1.5 ± 2.0 (0.7 to 2.4)	<0.0001	0.0004	0.75
		Right Biceps Femoris	25	8.2 ± 11.0	0.8 ± 1.1 (0.4 to 1.3)	0.002	0.03	0.74
Partial group	Proximal calf	FF<20%	13	6.0 ± 5.9	0.9 ± 2.1 (-0.4 to 2.1)	0.16	0.19	0.41
		FF 20% to 70%	8	44.7 ± 15.6	5.8 ± 2.7 (3.2 to 8.1)	0.0008	na	1.96
		FF>70%	4	78.0 ± 4.7	2.1 ± 1.1 (0.3 to 3.8)	0.03	na	1.85
		FF 20-70% (males)	5	47.4 ± 18.8	5.1 ± 1.3 (3.5 to 6.7)	0.001	na	3.93
		FF<20% (females)	9	4.0 ± 4.7	0.5 ± 0.6 (0.03 to 0.9)	0.04	0.049	0.83
	Distal calf	FF<20%	12	6.0 ± 7.5	0.3 ± 1.2 (-0.7 to 1.2)	0.52	0.98	0.22
		FF 20% to 70%	7	60.3 ± 6.8	5.1 ± 1.9 (3.1 to 7.2)	0.001	na	2.65
		FF>70%	3	83.2 ± 6.9	0.5 ± 1.5 (-3.2 to 4.2)	0.62	na	0.33
		FF 20-70% (males)	6	59.4 ± 20.2	3.4 ± 2.3 (1.0 to 5.9)	0.01	na	1.47
		FF 20-70% (females)	3	58.1 ± 7.4	6.4 ± 1.3 (3.2 to 9.6)	0.01	na	4.96
	Distal thigh	FF<20%	22	3.5 ± 3.2	0.6 ± 1.2 (0.2 to 1.2)	0.01	0.049	0.58
		FF 20% to 70%	3	28.9 ± 8.5	2.4 ± 0.6 (-0.5 to 2.9)	0.10	na	1.75
		FF>70%	0	No patients				
		FF<20% (males)	9	5.0 ± 4.3	1.3 ± 1.6 (0.04 to 2.5)	0.04	0.08	0.79

FF=fat fraction, sd=standard deviation, CI=confidence interval, SRM=standardised response mean, EHL=extensor halluces longus, n=number, na=not available due to no control data, p1=paired t-test with baseline, p2=two tailed t-test (HSN1 v controls). Rows are highlighted in blue if both p values (if available) are <0.05

5.6 Conclusion

Quantitative MRI measures significant FF change over 12 months in a second inherited neuropathy – Hereditary Sensory Neuropathy type 1. Significant change was detected at all anatomical levels: proximal calf, distal calf and distal thigh, with qMRI determined FF change at proximal calf level most responsive to change.

In keeping with results from previous chapters, qMRI outcome measure responsiveness can be increased by methods aimed at patient group homogenisation, including partial slice analysis and intermediate baseline FF subpopulation analysis. This study has also demonstrated for this first time the application of severity specific slice analysis, allowing outcome measure responsiveness to be maximised without sacrificing external validity of results, and the benefits of universal participation.

This study demonstrates that qMRI provides a highly responsive outcome measure for future trials in HSN1, thus strongly supporting its application as a primary outcome measure for clinical trials in HSN1.

6 Conclusion

Over the past decade, the search for effective outcome measures for use in clinical trials of rare neuromuscular diseases (NMD) has intensified, alongside the development of novel potential treatments. There is major interest in quantitative MRI (qMRI) for this purpose, given its recently proven reliability, validity and responsiveness in a range of NMD.

This thesis provides evidence which strongly supports MRI as a suitable outcome measure for trials in NMD, demonstrating qMRI determined fat fraction (FF) by the 3-point Dixon method to be a highly responsive outcome measure in two inherited peripheral neuropathies. In Charcot-Marie-Tooth disease type 1A (CMT1A), qMRI measures significant FF progression over up to five year follow-up and is shown for the first time to have longitudinal validity, with strong positive correlation with progression in CMT examination score (CMTES) and remaining muscle area (RMA). In Hereditary Sensory Neuropathy type 1 (HSN1), qMRI measures significant FF change with large responsiveness at both calf levels examined.

In both CMT1A and HSN1, significant improvement of outcome measure responsiveness can be achieved by optimisations related to baseline FF, specific muscle and slice selection and muscle fat gradients, each directed at maximising and homogenising mean FF change and reducing measurement error.

With this demonstration of continued large internal responsiveness and longitudinal validity in these two inherited neuropathies, with careful thought to outcome measure design, qMRI is trial ready for even the rarest NMD.

6.1 Novel findings presented in this thesis – in brief

6.1.1 CMT1A natural history study

This thesis presents results from the longest ongoing single centre study to date which uses qMRI as an outcome measure in CMT1A, and reveals significant ongoing fat accumulation in calf muscles of patients with CMT1A at up to five years: combined bilateral calf FF change (%/year from baseline) of 1.2 ± 1.1%, 0.9 ± 0.9%, 0.7 ± 0.6% for timepoints two, three and four. Quantitative MRI determined FF is shown to be highly responsive over extended follow up (standardised response mean for combined bilateral calf: 1.04, 1.01 and 1.07 for timepoints two, three and four), with further improvements in standardised response mean (SRM) afforded by extending trial duration and

stratification of patients based on baseline FF (maximum SRM for combined bilateral calf: 2.05). Baseline STIR is shown to predict FF change, and T2 is shown to be significantly higher in CMT1A muscles with normal FF indicating a pre-fatty phase of pathology which may be a potential target for treatments. Alongside these findings, qMRI measures significant change in RMA at extended follow up, again with large responsiveness (combined bilateral calf RMA change (%/year from baseline) of -2.3 ± 2.6cm^2 and -2.0 ± 2.2cm^2 at timepoints three and four) though is less responsive than matched FF change (SRM for combined bilateral calf RMA: -0.87 and -0.89 at timepoints three and four). Surprisingly, change in CMTES and CMTES-lower limb (CMTES-LL) was significant at extended follow-up (0.6 ± 0.6 and 0.5 ± 0.5 points/year at timepoints three and four, with SRM of 1.00 and 1.12), which is at odds with previously published data. There were significant moderate positive correlations between combined bilateral calf FF progression and change in CMTES/CMTES-LL (r_s =0.723 and 0.706 for CMTES, and 0.669 and 0.638 for CTMES-LL at timepoint three and four respectively) and moderate negative correlations with RMA (r_s= -0.512 at timepoint four) demonstrating for the first time, longitudinal validity of qMRI FF in CMT1A.

6.1.2 CMT1A gradient study

Significant quantifiable calf muscle fat gradients have been demonstrated in all calf muscles of CMT1A patients and controls, with potential to introduce marked measurement error if not taken into account in trial design. Fat gradient magnitude is up to 12.6%/cm in CMT1A and 6.7%/cm in controls, with muscles with intermediate FF most affected by gradients (74.4% of intermediate muscles). Peroneal innervated muscles are shown to have more intermediate fat infiltrated muscles (59.0%) than tibial innervated muscles (30.8%), which the author hypothesises may relate to anatomical location of motor endplates. The method of longitudinal slice selection affords improvement in outcome measure responsiveness, with greatest advantage seen when analysing maximum 'volume' of muscle (maximum SRM 1.46). The benefit to responsiveness from this method is greatest in muscles with gradients, and less important in those without.

6.1.3 HSN1 natural history study

In this rare, predominantly sensory neuropathy, qMRI determined FF by 3D 3-point Dixon is shown to be a highly responsive outcome measure at distal calf (combined bilateral distal calf FF change 2.2 ± 2.7, SRM of 0.83) and proximal calf level (combined bilateral proximal calf FF change 2.6 ± 3.0, SRM of 0.84) in a cohort of patients with variable clinical phenotype. Muscle fat gradients are demonstrated for the first time in this disease – albeit based on assessment at three levels rather than within muscles themselves – with a length-dependent pattern in the lower limb: combined bilateral mean FF 35.6 ± 31.9% → 29.9 ± 29.8% → 7.1 ± 9.7% at distal calf, proximal calf and distal

thigh level. As seen in CMT1A, intermediate baseline FF predicts greatest FF change at all levels: 5.8 ± 2.7% and 5.1 ± 1.9%/year for distal and proximal calf, with marked improvement in SRM (1.96 and 2.65 respectively). Slice selection based on baseline proximal calf FF led to an improvement in SRM to 0.97 (bilateral lower limb FF change 3.0 ± 3.1%/year), but this can be further improved by more comprehensive imaging including of foot muscles. We have also demonstrated a length-dependent disconnect between the dual processes of fat infiltration and atrophy seen at all levels analysed, but most significantly at thigh level, less often at proximal calf level and in only two individual muscles at distal calf level. The pattern of calf muscle fatty atrophy in HSN1 is shown to be different to that seen in CMT1A, with more marked involvement of soleus and lateral gastrocnemius muscles. Tibialis posterior remains least affected however, as seen in CMT1A.

6.2 Quantitative MRI responsiveness

Outcome measure responsive is critical to trial design, and even more so in NMD, which are universally rare. The work in this thesis gives an evidence base for the refinement of qMRI for clinical trials. Several suggestions and observations are noted below, based on this work.

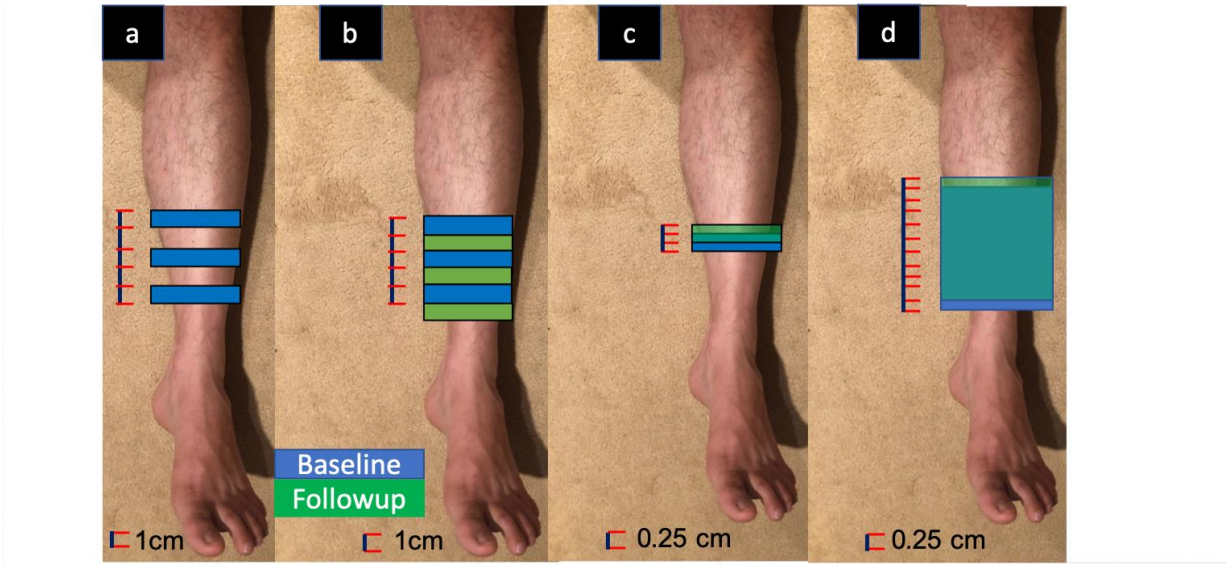
6.2.1 Muscle fat gradients and accurate patient positioning

It is clear from the work in Chapter 4 of this thesis, that patient and MRI block positioning are critical to longitudinal reliability of qMRI. With axial imaging slices positioned 10mm or 20mm apart, as is currently the case in many neuromuscular centres (our CMT1A natural history study used 10mm thick slices with 10mm gap between slices), there is risk of FF measurement error of up to 12%, solely due to block placement (if slices are perfectly interleaved), which in turn renders longitudinal analysis futile when looking for real FF changes in the region of 1-3%. In the worst case scenario with these slice parameters, there is actually no overlapping muscle at all (Figure 6-1b) longitudinally and the effects of muscle fat gradients becomes highly unpredictable. One can appreciate in this scenario that increasing the number of analysed slices does not change the proportion of overlapping muscle, and so one would not expect an improvement in responsiveness by simply increasing slice number, unless this were to randomly occur.

Each neuromuscular centre must carefully consider how best to tackle this real problem, be it with more detailed imaging sequences with reduced space between slices, as is being pursued in our Centre with the addition of 3D 3-point Dixon (5mm slice thickness with no gap between slices), with more strict positioning measurements based on scout imaging, or ideally with both. When analysing 3D 3-point Dixon imaging, one can

appreciate that if a single slice is analysed longitudinally, block mispositioning would mean that in the worst case scenario (i.e. longitudinal slice out of position by 2.5mm), only 50% of the muscle analysed would overlap (Figure 6-1c). If the number of slices analysed are increased, then a similar longitudinal mispositioning of 2.5mm would result in a reduction of the error proportional to overlapping muscle, as is demonstrated in Chapter 4 and Figure 6-1d. This would not lead to a difference in the measurement error compared with using fewer slices if the muscle fat gradient is constant, however as shown in Chapter 4, a constant FF gradient is rarely the case, in which case there would be an averaging out of the effect of muscle fat gradients in addition to reduction in 'noise' by averaging FF over a larger number of voxels. Of course, one must be mindful that the latter would only be the case if the 'noise' were to be evenly distributed throughout the imaging field. If for example, proximal or distal imaging encompasses tendons, artefact from these structures may result in a paradoxical increase in measurement error if included.

Figure 6-1 – Baseline and follow-up mispositioning comparing different interslice distance

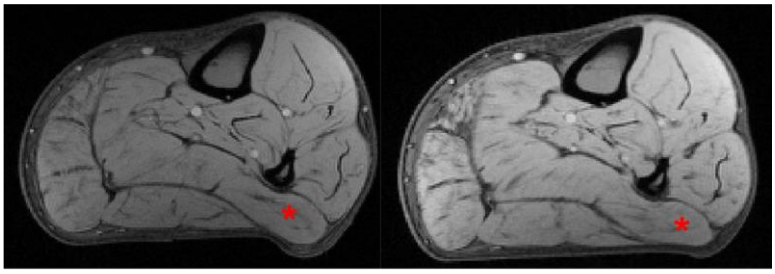


Example of baseline (blue) and follow up (green) imaging. In panels a and b, each axial slice is 1cm thick with a 1cm gap between slices. If baseline and follow up scans are perfectly interwoven (b), there is no overlapping muscle (turquoise) between baseline and follow up scan, and the effects of gradients become unpredictable. In panels c and d (representative of 3D 3-point Dixon imaging) slices are 5mm thick with no gap. If a single perfectly interwoven slice is analysed, 50% overlap of muscle analysed results (c), and if the analysed block is larger, this same error of placement becomes proportionally less important if muscle fat gradient is not constant. Superimposed MRI slices not to scale

Finally, in terms of positioning, whether the leg is relaxed firmly against the MRI gurney, or semi-suspended by placing the ankle on a support, affects the shape of the posterior leg muscles on imaging (Figure 6-2). This should not be a confounding factor for FF measurement as long as the lower limb is parallel with the long axis of the MRI bore on

each occasion, and the same leg positioning is recorded and maintained longitudinally. The distortion should also not affect cross-sectional area (CSA), given that muscle is not compressible at these pressures, though this needs to be examined in more detail longitudinally.

Figure 6-2 – Muscle distortion due to limb positioning

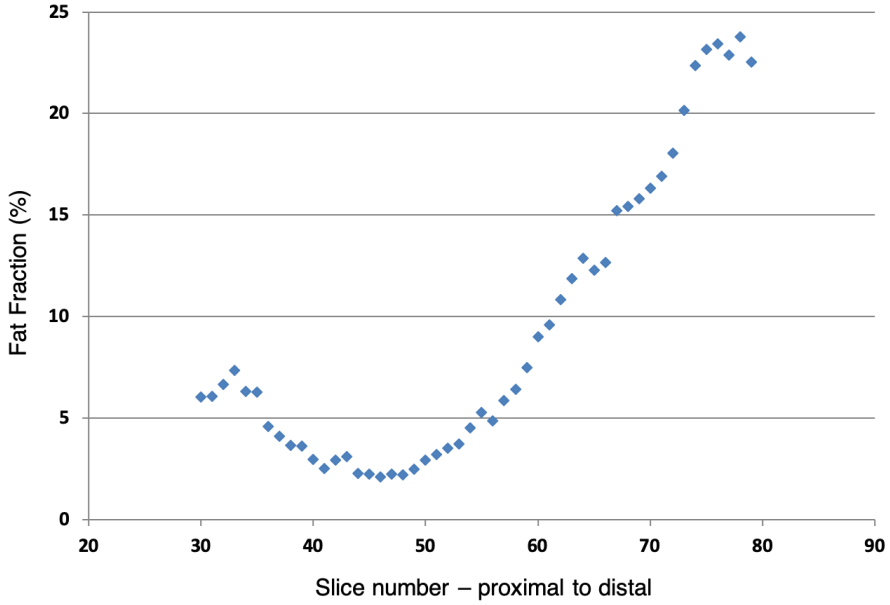


Distortion can be seen in the posterior aspect of soleus (*) between scans, due to lower limb positioning against the scanning table

6.2.2 Muscle fat gradient analysis in other neuromuscular diseases

We have demonstrated the presence of muscle fat gradients in two inherited neuropathies. As discussed, quantitative assessments of muscle fat gradients have been reported in several muscle diseases (Kan et al., 2009; Lareau-Trudel et al., 2015; Hooijmans et al., 2017). It is likely that all NMD will have at least some proximal-distal gradients which need to be accounted for in analysis. It is important however to establish the presence and precise distribution of muscle fat gradients in other NMD, as this is likely to significantly impact exact choice of qMRI outcome measure and method of analysis used. Our centre has started assessing muscle fat gradients on 3D 3-point Dixon imaging, with excellent initial results (Figure 6-3), though as discussed below, this will not be realistic until automated segmentation becomes widely available.

Figure 6-3 – Fat fraction (%) in right anterior compartment of a single patient with HSN1



Fat fraction derived from 3D 3-point Dixon imaging. Anterior compartment fat fraction is a combination of tibialis anterior and extensor hallucis longus. Each slice is 5mm thick with no gap

6.2.3 Prognostic enrichment

In a recent paper on clinical trials in Duchenne muscular dystrophy (DMD), the Federal Drug Agency (FDA) suggested that 'for drugs that may slow clinical decline but are not expected to improve or reverse pre-existing muscle dysfunction, it may be useful to consider prognostic enrichment (i.e. the use of inclusion criteria to select patients with characteristics that predict more rapid clinical decline during the planned study). Such criteria might include a history of rapid deterioration before study entry or more severe functional deficit at enrolment.'

Two forms of 'prognostic enrichment' are demonstrated in this thesis

6.2.3.1 Bespoke slice selection

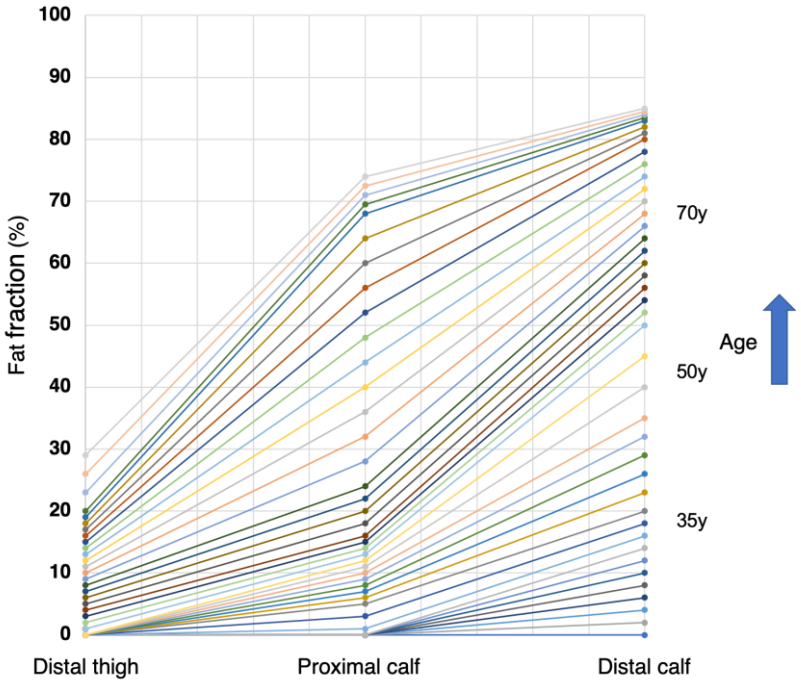
We have suggested severity specific slice selection as one method by which mean FF change can be maximised and homogenised across a cohort of patients with clinically heterogenous disease. In this thesis, this concept is clearly demonstrated in the HSN1 cohort with improvement in SRM of qMRI determined combined bilateral FF from 0.84 (proximal calf) to 0.97 (severity specific slice selection) with this method – Table 5-23. As noted in Chapter 5, even with a severity specific slice selection model, many of the distal calf slices had low/normal baseline FF with resultant little change over 12 months, which reduces the SRM both by reducing mean change and increasing standard deviation of mean change. It is in these patients that foot imaging is suggested for further enhancement of responsiveness.

6.2.3.1.1 Model of HSN fat accumulation

Based broadly on values from Table 5-18, Table 5-19 and Table 5-20, a model of fat accumulation in HSN1 is presented in Figure 6-4. The rapid phase of fat accumulation at each level does not coincide temporally, thus supporting a severity specific slice approach to outcome measure selection. This model is readily applicable to other length-dependent neuropathies such as CMT1A.

Of course, based on results from Chapter 4, the line joining each point (particularly at intermediate fat fractions), is not linear as shown in this model, as positive and negative gradients are present between each level.

Figure 6-4 – Model of lower limb fat accumulation in a single HSN1 patient



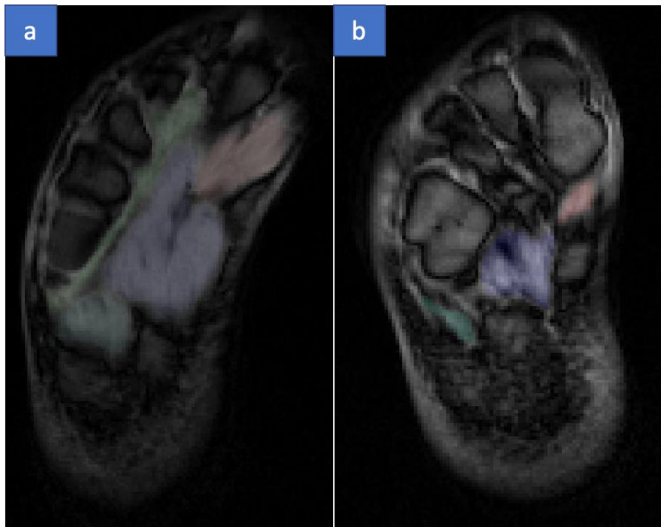
Each horizontal line represents a single timepoint. As each anatomical level reaches an intermediate fat fraction, it enters a more rapid phase of fat accumulation relative to the other levels, making that level optimal for outcome measure selection

6.2.3.1.2 Foot imaging

Quantitative MR imaging of the feet has been suggested in this thesis for analysis in those patients with normal distal calf FF (Chapter 5). At our Centre, we have added foot imaging to our 3D 3-point Dixon protocol. Although initial results are promising, imaging is not straightforward given the variable anatomical distortions (e.g. – differing ankle and toe angles, variable foot anatomy/achilles tightening/toe deformities) seen in patients with NMD. Two representative axial foot MR images are shown in Figure 6-5. It is clear that reproducible positioning is complicated for reasons mentioned, and segmentation is challenging due to muscle size and distorted anatomy. Before foot MR imaging can be used in clinical trials, demonstration of reliability, validity and responsiveness must be

demonstrated, as with any new outcome measure. For these reasons, it is unlikely that it will be applicable in the near future, though it is a potentially very powerful addition to the neuromuscular qMRI repertoire.

Figure 6-5 – Axial foot imaging in HSN1



Axial foot MRI in two patients with HSN1: a) clinically mild, b) clinically moderate. Regions of interest are superimposed on raw Dixon sequence (TE=3.45s). Green= interosseous, light blue=lateral group, blue=central group, red=medial group

6.2.3.2 Bespoke patient selection

A second form of prognostic enrichment demonstrated in this thesis is patient enrolment based on baseline FF. In our hands, maximum change is seen in intermediate fat infiltrated muscles in all groups studied (Chapters 3, 4 and 5) with marked improvement in responsiveness in CMT1A (Table 3-23) and HSN (Table 5-18) and further improvements by increased muscle volume analysed (Table 4-7). The drawback with this method is the exclusion from trials of patients with low or high baseline FF with potential to affect external validity or generalisability of study results. It is important to know the proposed mechanism of action of the medication to be trialled, in order to determine which subgroup may be best suited to the trial. For example, a medication may be effective in early disease without benefit in patients who have accumulated a certain amount of fat, in which case it would be futile to look at patients with intermediate fat infiltration at baseline in hopes of larger responsiveness.

With both of these methods of prognostic enrichment, careful consideration would have to be given to trial design, and the precise details of the primary (or secondary) outcome measure used. For both methods, one option may be a pilot study prior to the trial, to determine the exact distribution of fat infiltration in each patient with subsequent decision on bespoke slice and/or patient selection as appropriate. It is clear that stratification based on clinical or functional measures would be insufficiently sensitive. It is worth mentioning at this point however that in our CMT1A natural history study, all patients

aged <35y had no significant calf muscle fat infiltration, and this too may be a complimentary way in which to stratify participants.

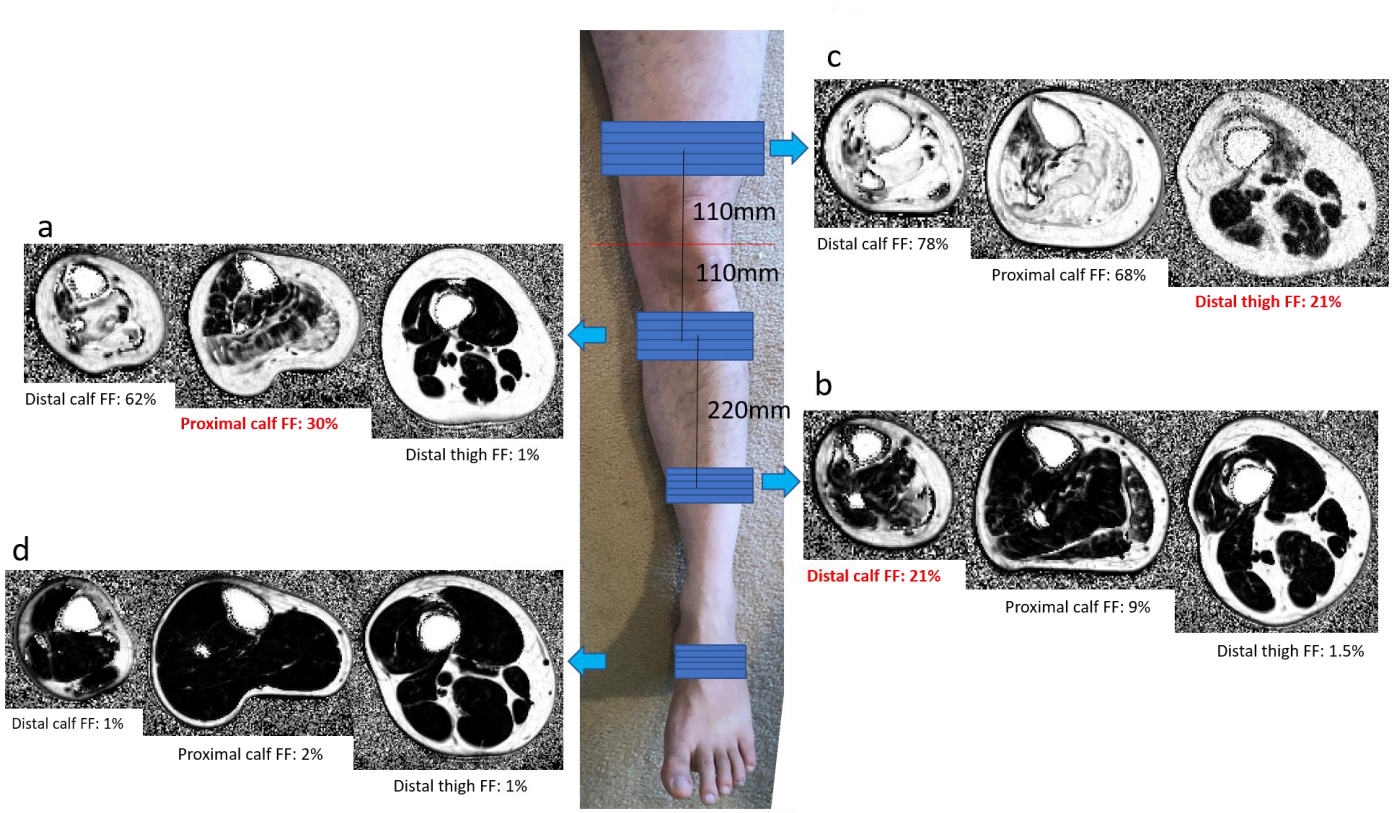
6.2.4 Optimal method of slice selection

In summary therefore, based on the findings in this thesis: the author makes the following overall suggestions (Figure 6-6) to maximise qMRI determined FF outcome measure responsiveness in a cohort of clinically heterogenous patients with a length-dependent peripheral neuropathy.

Firstly, a single limb is adequate and roughly comparable in responsiveness to both limbs combined. Secondly, and perhaps most critically is accurate slice positioning, based on scout imaging rather than surface anatomy. Thirdly is using 3D 3-point Dixon imaging if possible, to improve slice selection capabilities, but if not available, point two becomes ever more important. Fourthly is examination of whole slice FF in place of partial slice. Fifthly is examining at a site at which baseline FF is intermediate in each patient (foot imaging needs further development), and sixthly examining larger volume of muscle (within reason pending autosegmentation). Finally, a trial of ~24 months duration is preferred to one of shorter duration.

These suggestions do not mention the method by which the largest SRM were seen in this thesis: which is examining a homogenous cohort of patients with intermediate baseline FF, with exclusion of other patients (high and normal baseline FF) from analysis, though as discussed this has impact on external validity of results.

Figure 6-6 – Author's suggested method of longitudinal slice selection for qMRI determined fat fraction



The author suggests longitudinal analysis of multiple adjacent whole slices at one anatomical location in each patient, with site chosen based on baseline location of intermediate whole slice fat fraction: a=proximal calf, b=distal calf, c=distal thigh and d=foot. Foot imaging needs further development, but would be invaluable if found to be reliable and valid. These patients would currently have distal calf imaging, though unlikely to show great FF change. NB – in this example, the three more proximal imaging blocks are separated by 220mm (as per our HSN1 protocol)

6.3 Place of qMRI outcome measures in clinical trials

The exact place and nature of qMRI outcome measures in clinical trials will depend on the disease being studied and the trial phase. In slowly progressive inherited neuropathies in which other outcome measures have been shown to be poorly responsive, and in which there is otherwise no evidence of longitudinal validity, there are few reasons why qMRI derived FF should not be considered as the primary outcome measure going forward – this of course requires detailed discussion with FDA, European Medicines Agency (EMA) as appropriate. Other measures, including other qMRI derived measures, clinical and functional must also be included giving a more rounded body of evidence.

In the early phases of clinical trials (phase II), it may be appropriate to stratify patients based on baseline FF, using only those with intermediate FF.

6.4 Implications for trial design from this thesis

The findings in this thesis have important implications for outcome measure choice for future trial design. The specific qMRI outcome measure will depend on the disease to be studied, and it is worth knowing the distribution of fat and/or oedema in each disease before making this decision. Of course, this is already known from qualitative studies for many NMD.

In this thesis, qMRI determined FF by 3-point Dixon measured significant change over trial duration in two representative inherited neuropathies. In CMT1A, significant change was measured at extended follow up in almost all muscles/muscle groups analysed. In HSN1, significant change was measured in all combined calf muscle groups and many individual muscles at all levels at 12 months. Various methods have been discussed which further refine qMRI responsiveness. This therefore gives the researcher a wide array of choices for potential primary or secondary outcome measure for an upcoming clinical trial.

As an example, based on responsiveness measured by SRM in Chapters 3 and 4 of this thesis, a future study in patients with CMT1A powered to detect a 50% reduction in disease progression over 12 months with 80% power at $p<0.05$ significance would require (using Lehr's formula – Equation 1-3) $\sim 16/(0.91 \times 0.5)^2 = 77$ participants in each arm if right calf FF were to be used as the primary outcome measure, reducing to ~59 participants if combined bilateral calf FF were used. Marked further reduction in participants needed for equivalent power is achieved if a cohort of patients stratified for baseline FF are assessed. Based on results from the CMT1A gradient analysis, the number of patients would be further reduced to ~28 with whole right calf muscle MRI-

determined FF by method e (FF based on weighted average of six slices centred at the point 14cm distal to the right tibial plateau) and to ~20 if the peroneal group – moderate FF is used (method e).

Rather than highlighting the precise numbers, which may vary slightly, this gives concrete demonstration of the general trend, that increasing outcome measure responsiveness two-fold, results in a four-fold reduction in participants needed in each arm for the same statistical power (Jacobson et al., 1999). This is of vital importance in the rare diseases with which we are dealing. Examples of SRM and n. for different qMRI determined fat measures from CMT1A chapters of this thesis are shown in Table 6-1. A similar table for HSN is shown in Chapter 5 (Table 5-24).

Table 6-1 – Number of participants needed in each arm of a hypothetical study in CMT1A, depending on qMRI outcome measure selection

Study	Outcome measure	SRM	n
Natural history study	Right TA muscle - all patients	0.73	120
	Right calf - all patients	0.91	77
	Left calf - all patients	0.99	65
	Bilateral calf - all patients	1.04	59
	Bilateral calf - intermediate FF	1.94	17
Gradient study - all by method e	MG muscle - all patients	0.88	83
	PL muscle - all patients	1.26	40
	Peroneus group - all patients	1.46	30
	Right calf - intermediate FF	1.50	28
	Peroneal group - intermediate FF	1.78	20
	PT muscle - intermediate FF	3.51	5

SRM=standardised response mean, n=number, TA=tibialis anterior, FF=fat fraction, MG=medial gastrocnemius, PL=peroneus longus, PT=tibialis posterior, method e=FF based on weighted average of six slices centred at the point 14cm distal to the right tibial plateau

The demonstrated responsiveness of qMRI improves further on the published 12 month data from our CMT1A cohort (FF change 1.2 ± 1.5%/year, SRM=0.83) and provides the foundation for a potentially major advancement in trial design for NMD.

Finally, in this thesis, the author has concentrated on methods to improve responsiveness of qMRI as an outcome measure, though it must be assured that the significant change measured by the outcome measure is relevant to the patient. The outcome measure must reflect the specific NMD studied and its natural history. Inclusion of overall functional measures in any clinical trial is critical in this regard, with longitudinal correlation of significant importance.

6.5 Considerations and future developments

The number of clinical trials in NMD is likely to rapidly increase in the near future. A high degree of precision is needed in trial design and execution. It is incumbent on the researcher to have a firm grasp of the potential pitfalls in order to address these as best as possible prior to trial commencement.

6.5.1 Imaging

6.5.1.1 MRI systems

Even the most precise longitudinal patient positioning technique is rendered meaningless if the MRI protocols used are not standardised, reproducible and validated. A good example of this is demonstrated in the first year of our published CMT1A/IBM natural history study data, when a routine software upgrade resulted in a systematic bias between pre and post upgrade T2 values in the order of ~10%, which was fortunately identified and rectified by the authors (Morrow et al., 2016). As mentioned, looking for longitudinal change in the region of 1-3% would be impossible if such problems were to go unrecognised.

6.5.1.2 Multi-site imaging

Quantitative neuromuscular MRI has not yet been used as an outcome measure in drugs trials in NMD, including in the recent Spinal muscular atrophy (SMA) and DMD trials. This is in part due to concerns regarding the potentially high degree of variability in platforms (including scanner and field strength) and protocols/acquisition parameters across sites, which would limit the reliability and validity across sites. Although inter-site validity is important, and protocols must be harmonised across sites (Hollingsworth et al., 2012), the main concern in longitudinal studies is that each participant's longitudinal MRI scan is done on the same machine to ensure reliability for each participant.

6.5.1.3 Increasing speed of acquisition and signal to noise ratio

All imaging trials in NMD which have been reviewed in this thesis, have been performed using either 1.5 or 3 Tesla MRI systems, which are the standard for use in clinical practice in the UK today. 3 Tesla systems are prone to B1 field inhomogeneity in some sequences (T1 and MTR), though the Dixon method is essentially immune to this.

MRI units with higher magnetic field strength are beginning to be approved for clinical practice (FDA approved for clinical use Oct. 2017) and research. The increase in magnetic field strength between 3 and 7 Tesla is by a magnitude of 2.3, with an increase in signal to noise ratio of ~2 – allowing shortened scan time and improved spatial resolution (Juras et al., 2019). Due to the fact that chemical shift is increased at 7 Tesla, there is improvement in spectral resolution, and this may be beneficial for studies using MR spectroscopy, though is yet to be applied to patients with NMD (Alizai, Chang and

Regatte, 2018). Apart from the number of technical challenges higher field strengths would present, given the cost it is unlikely that 7 Tesla will be widely available in the near future, and thus one of the practical disadvantages would be cross-site validation.

Even with the current preferred use of 3 Tesla MRI systems for research, there can be a discordance in exact parameters between scanners from different manufacturers, and it is important that follow-up scans are always on the same machine, to avoid any errors related to subtle machine differences.

Since 2-point 3D Dixon was first described, there have been many further developments to improve its application, including methods to increase speed of acquisition (Schlaeger. et al., 2018) and multi-echo techniques, though it is not yet clear whether the latter improves FF quantification. Reeder et al. combined gradient-echo imaging with IDEAL (3-point IDEAL) to optimise signal to noise ratio and allow better modelling of the lipid spectrum (Reeder et al., 2007; Hu et al., 2012) but the basic process remains unchanged.

6.5.1.4 Whole body imaging and upper limb imaging

To date, there have been no studies in patients with CMT using qMRI of the upper limbs. As discussed in Chapter 3, qMRI determined FF at mid forearm level is shown to be a responsive outcome measure over 12 months in non-ambulant patients with DMD (Ricotti et al., 2016) thus giving support to include these patients in future clinical trials. In CMT1A however, patients are rarely non ambulant, and thus imaging of upper limbs whilst adding time to an MRI protocol, may not add very much further information to that gained from targeted lower limb imaging. It could however be considered in more rapidly progressive neuropathies such as HSN, in which patients may have had lower limb amputation, and in non-length dependent neuropathies (Gaeta et al., 2012).

The technology is available for whole body imaging to be done in a relatively short period of time (Karlsson et al., 2015), and it is clearly very useful in certain clinical circumstances, for diagnosis of neuromuscular conditions (Tomas et al., 2019). Given the large amount of data, Hankiewicz et al. have suggested using heat maps for a visual representation of disease activity (Hankiewicz et al., 2015). In terms of fat and water quantification however, apart from the obvious obstacles discussed below related to muscle segmentation, there are issues relating to optimal positioning for the upper limb (currently the arms are most often imaged separately to allow each limb to be in the middle of the magnetic field) and artefact produced by respiration in the chest/abdomen. Over and above these issues, whole body imaging is currently unlikely to add further to qMRI responsiveness in a length-dependent neuropathy such as CMT1 or HSN1, in which most abnormal muscle is in distal lower limbs. Indeed in the extended follow up in

our cohort we have not analysed thigh imaging due to fewness of patients with involvement at that level.

6.5.2 Analysis

6.5.2.1 Statistical methods

Although perhaps currently the most commonly used in medical research, the SRM is not the only method by which responsiveness of an outcome measure can be quantified. Other methods include calculating the 'standardised effect size' (ES) which is similar to the SRM, though compares size of change with the standard deviation at baseline rather than the standard deviation of the change. Both of these methods are potentially limited by the fact that they do not relate changes in the outcome measure to changes on a clinical or health status measure in the individual patient, risking a statistically significant yet clinically meaningless change in the outcome measure. To address this point, Guyatt and colleagues developed a responsiveness index (Guyatt, Walter and Norman, 1987) which includes in its calculation the minimally clinically important change in that measure. This would seem a superior measure of responsiveness, but its application to trials in NMD is currently limited by the fact that the minimally clinically important FF change is not yet known. Herein lies the importance of establishing longitudinal validity as has been demonstrated in this thesis.

6.5.2.2 Automated muscle segmentation

The field is in desperate need of a method of lower limb muscle auto-segmentation. Over the last several years, there has been progress with techniques that can quantify total fat and volume on Dixon or T1 weighted MRI. Although these techniques may be reliable in normal muscle, as FF increases the reliability reduces (Positano et al., 2009; Makrogiannis et al., 2012). Gadermayr and colleagues reported an average error of between 2 and 6% in their study (Gadermayr et al., 2018), which is unacceptable when looking for much smaller changes over short trial duration. It is likely that the first step forward will be to semi-automated segmentation (Kemnitz et al., 2017) with the checking process itself still highly time consuming. There is also interesting novel work on muscle segmentation based on muscle texture (Rodrigues et al., 2017), though this is still in its infancy. Further refinement will be need if auto-segmentation is to replace the nuanced art of manual segmentation.

With respect to manual segmentation, the author estimates that if each individual lower limb muscle were to be segmented on a single axial slice, this would take (including time needed to check each region of interest) ~30 minutes per individual calf slice, and perhaps ~40 minutes for each thigh slice. Based on this figure, a longitudinal study with 25 participants in each arm (single slice bilateral thigh and calf imaging) would take ~25 days of 9am-5pm non-stop muscle segmentation. This time increases exponentially and

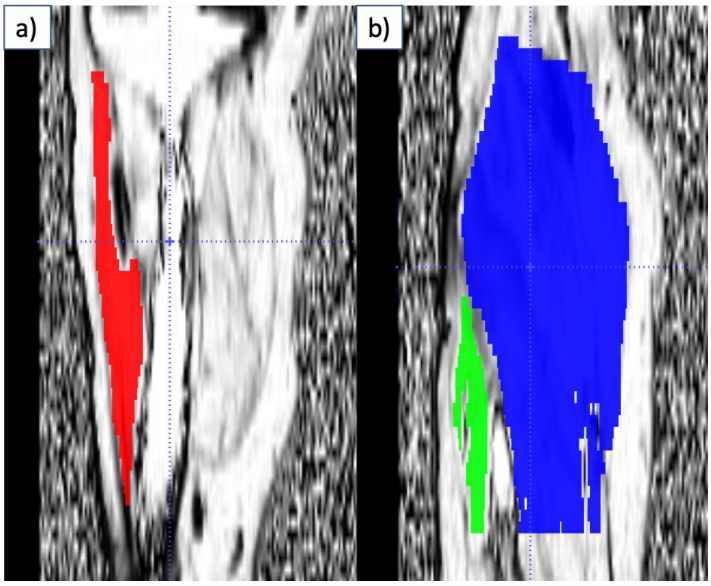
becomes untenable for a single examiner when considering emerging 3D Dixon protocols which may include up to 100 axial slices per lower limb. An equivalent longitudinal study segmenting all muscles on all slices in this case would take ~100 hours/scan → ~3.5 years of muscle segmentation! This of course cannot be done by a single individual, and so teams would be needed with the obvious problems that would introduce, including inter-operator variability.

This time would be reduced significantly if single muscle or a single segmentation for entire calf or thigh is used whilst waiting for much needed accurate automatic segmentation to become applicable, though this may lead to reduction in responsiveness as detailed above.

6.5.2.2.1 Coronal versus axial regions of interest

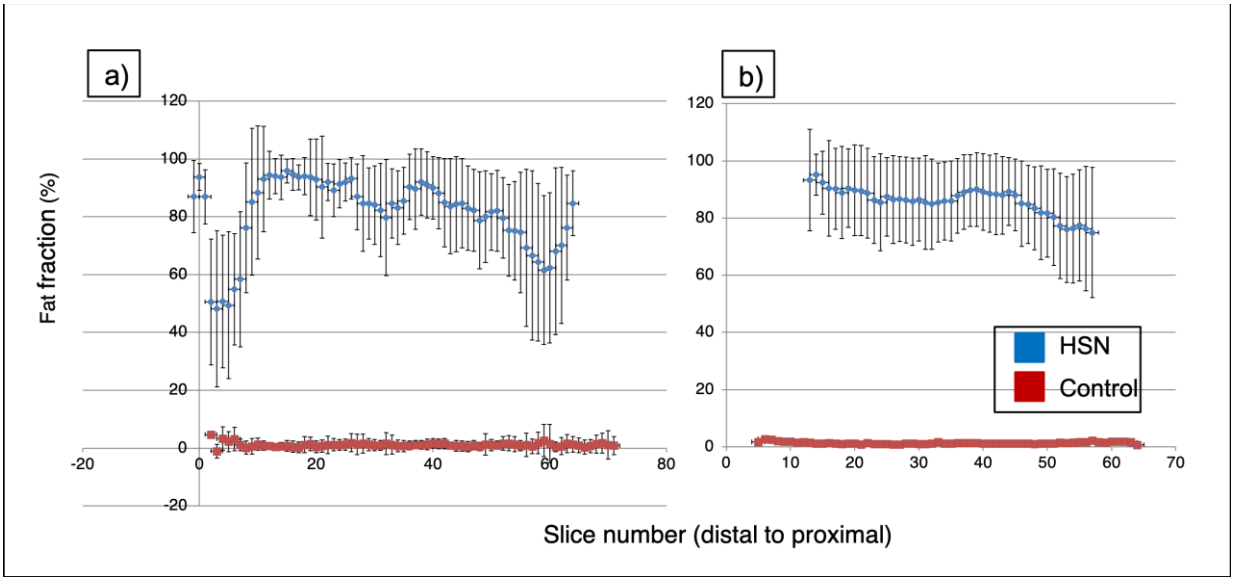
Whilst automated segmentation is in development, a relevant question is whether there are other methods by which lower limb muscles can be segmented in order to reduce segmentation time. We have demonstrated that there is large variability in positioning of small regions of interest leading to increased standard deviation of change, and thus poor reliability and low responsiveness as an outcome measure (Table 3-18). Another approach trialled by our team has been to segment muscles on coronal (Figure 6-7), rather than axial section. Preliminary cross-sectional results in a small number of our HSN cohort however has revealed large variation in FF in both controls and patients (Figure 6-8), akin to that seen with small regions of interest, and it is unlikely therefore that his technique would be adequately sensitive to change over time for implementation in clinical trials.

Figure 6-7 – Coronal muscle segmentation in right lower limb (3D Dixon)



Coronal sections through the right lower limb. Anterior compartment in red (a), Peroneus longus/brevis in green, posterior compartment in blue (b)

Figure 6-8 – Coronal versus axial segmentation or peroneal muscles in a patient with HSN



Fat fraction derived from coronal (a) and whole muscle axial (b) regions of interest in the right lower limb of a single patient with HSN (blue), and control (red). Error bars are one standard deviation

6.5.2.3 Other quantitative MRI methods

Although the Dixon method is the most common fat quantification method for trials in NMD, there are other qMRI sequences which merit further consideration in future studies.

6.5.2.3.1 MR spectroscopy

In essence, given that most MRI scanners rely on detecting proton signals, scanning of ^{1}H (proton MR spectroscopy (MRS)) is the most commonly found type of MRS in practice. Indeed it has been used extensively to quantify and monitor intramuscular fat accumulation in various NMD. In DMD, proton MRS quantifies intramuscular fat, and was shown to monitor corticosteroid treatment (Arpan et al., 2014; Willcocks et al., 2016). Spectroscopy with different nuclei, have also shown some promise in fat quantification. ^{23}Na MRS has a lower signal-to-noise ratio than ^{1}H MRS and thus requires stronger magnetic field strength (Madelin and Regatte, 2013). Some of the skeletal muscle channelopathies are due to mutations in sodium channel genes, resulting in fluctuation in muscle. As such, there may be a role for ^{23}Na MRS in monitoring these diseases and their response to treatment.

6.5.2.3.2 Diffusion imaging

The basis for the use of diffusion tensor imaging (DTI) is that water molecule diffusivity increases when there is inflammation in skeletal muscle, and will then reduce with progressive fat infiltration (Qi et al., 2008). Fractional anisotropy (FA) refers to the degree that the movement of water molecules is in one direction. Structures with high FA include

white matter tracts, peripheral nerves and muscle fibres. Inflammation or ultrastructual abnormalities in muscle disrupts the natural lateral barriers, allowing horizontal movement of water molecules which is represented by reduced FA (Budzik et al., 2014). The apparent diffusivity coefficient (ADC) increases as diffusivity increases, and can be used to quantify diffusivity. In their study in patients with DMD, Ponrartana found that muscle strength negatively correlated with FA and positively with ADC, however more advanced disease correlated positively with FA and negatively with ADC, which the authors felt may have been due to an artificial increase of ADC and decrease of FA in muscles with >45% FF (Ponrartana et al., 2015). This does throw the reliability and validity of this method into question. Other diffusivity based imaging such as tractography (Damon, Heemskerk and Ding, 2012) applied to muscle cells may add further quantitative information, but are not appropriate for use as outcome measures.

6.5.3 Emerging methods to assess muscle longitudinally

6.5.3.1 Muscle texture analysis

Computerised texture analysis of skeletal muscle has been found to differentiate normal from dystrophic muscle with good sensitivity and specificity (Herlidou et al., 1999). With respect to the possibility of using texture analysis for longitudinal assessment, Arpan et al. built histograms of lower leg muscles in thirty patients with DMD based on each muscle pixel's T2 time, and compared these to histograms from matched controls. The 95th centile of all control histograms was used as the threshold value and applied to DMD and control histograms. The authors then measured the percentage of T2 pixels above the threshold, and the heterogeneity of muscle involvement (by histogram width). They found that this method could provide additional measures with which to monitor disease involvement, and that this alternate approach correlated strongly with function (Arpan et al., 2013).

There are certain hurdles to the use of texture analysis in multi-centre trials. As per any MRI system, parameters must be uniform across centres. It has been shown that voxel size greatly impacts results – if too large, texture information is lost, and if too small, artefact seems to obscure the subtleties of texture. This is a field in ongoing development, which may come to complement other qMRI methods.

6.5.3.2 MR elastography

This technique assesses the propagation of a standardised mechanical wave in soft tissue, with its velocity dependent on tissue qualities. Elasticity is then quantified differentiating normal and abnormal muscle (Ringleb et al., 2007): fat infiltration reduces, whilst inflammation increases elasticity. Although this may prove useful in early muscle

disease (Pichiecchio et al., 2018), the sensitivity of this technique does not compare with other qMRI methods at present.

6.6 Final conclusions

It is an exciting time to be involved in neuromuscular research. Basic scientific research into novel treatments for inherited NMD is burgeoning with international collaboration facilitating ground-breaking discoveries. In tandem, is equally important research to find the most suited outcome measure to anchor meaningful clinical trials. Quantitative MRI is now firmly positioned at the front of the queue.

The rarity and slow progression of many NMD makes measurement of change challenging, hindering clinical trials. This thesis has provided a swathe of novel evidence-based suggestions for refinement of qMRI as an outcome measure for clinical trials in NMD.

Based on findings in this thesis, the author's opinion is that maximum qMRI responsiveness is attained by accurate positioning alongside a carefully considered combination of analysis of maximum muscle volume, and some derivative of patient or severity based stratification. Once foot imaging is available, severity specific slice selection may prove an even more responsive choice. The final decision will be dependent on characteristics of the particular neuropathy and patient cohort, in addition to the projected effect of the trial medication.

Quantitative MRI is ready to be applied as the primary outcome measure in upcoming trials in inherited neuropathies.

7 Publications and abstracts relating to this thesis

7.1 Publications

Evans, M., Laura, M., Hoskote, C., Reilly, M.M. Cervical spinal cord compression complicating the clinical course of Charcot-Marie-Tooth type 1. *BMJ Case Reports* 2015:bcr2015213486. doi: 10.1136/bcr-2015-213486

Kugathasan, U., **Evans, M. R. B**., Morrow, J. M., Sinclair, C. D. J., Thornton, J. S., Yousry, T. A., … Reilly, M. M. (2019). Development of MRC Centre MRI calf muscle fat fraction protocol as a sensitive outcome measure in Hereditary Sensory Neuropathy Type 1. *Journal of Neurology, Neurosurgery and Psychiatry*. doi: 10.1136/jnnp-2018-320198

Ricotti, V., **Evans, M. R. B**., Sinclair, C. D. J., Butler, J. W., Ridout, D. A., Hogrel, J. Y., … Thornton, J. S. (2016). Upper limb evaluation in duchenne muscular dystrophy: Fat-water quantification by MRI, muscle force and function define endpoints for clinical trials. *PLoS ONE*. doi: 10.1371/journal.pone.0162542

Morrow, J. M., **Evans, M. R. B**., Grider, T., Sinclair, C. D. J., Thedens, D., Shah, S., … Reilly, M. M. (2018). Validation of MRC Centre MRI calf muscle fat fraction protocol as an outcome measure in CMT1A. *Neurology*. doi: 10.1212/WNL.0000000000006214

Rossor, A.M., **Evans, M.R.B**., Reilly, M.M. A practical approach to the genetic neuropathies. April 2015. *Practical Neurology* 15(3). doi: 10.1136/practneurol-2015-001095

U. Kugathasan, M. Laurá, P.J. Tomaselli, **M.R.B. Evans.**, ... Reilly, M.M. Hereditary Sensory Neuropathy Type 1 (*SPTLC1*): phenotypic variation in patients with the English founder mutation. doi: 10.1016/S0960-8966(17)30296-1

Bishop, C. A., Ricotti, V., Sinclair, C. D. J., Butler, J., **Evans, R. B. M**., Morrow, J. M., … Janiczek, R. L. (2014). Semi-automated analysis of diaphragmatic motion during deep breathing using dynamic MRI in both healthy controls and non-ambulant Duchenne muscular dystrophy. *Neuromuscular Disorders*. doi: 10.1016/j.nmd.2014.06.200

7.2 Abstracts

Evans, M., Morrow, J., Sinclair, C., Shah, S., Hanna, M., Reilly, M., … Yousry, T. (2014). MRI QUANTIFICATION OF FAT GRADIENTS IN CALF MUSCLES IN CMT1A. *Journal of Neurology, Neurosurgery & Psychiatry*. doi: 10.1136/jnnp-2014-309236.173

Evans, MRB., Morrow, JM., Sinclair, CDJ., … Hanna MG. Accurate slice selection improves responsiveness of quantitative lower limb muscle MRI in CMT1A patients. Volume 52. September 2015. *Muscle & Nerve.*

Morrow, JM., **Evans, MRB**., Grider T, … Reilly MM, Quantification of intramuscular fat accumulation in CMT1A using MRI: an international longitudinal study. March 2017. *Neuromuscular Disorders.* 27(1):S26. doi: 10.1016/S0960-8966(17)30295-X

Kugathasan, U., Laura, M., **Evans, MRB**., … Reilly, MM. Natural history study in hereditary sensory neuropathy type 1. December 2017. *Journal of Neurology Neurosurgery & Psychiatry* 88(Suppl 1):A65.1-A65. doi: 10.1136/jnnp-2017-ABN.223

Morrow, JM., Rawah, E., Sinclair, CDJ., **Evans, MRB**., … Yousry, TA. MRI quantification of fixed myopathy in hypokalaemic periodic paralysis identifies potential therapeutic window. March 2014. *Neuromuscular Disorders* 24(1):S17. doi: 10.1016/S0960-8966(14)70054-9

Kugathasan, U., Laura, M., Tomaselli P., **Evans, MRB**, … Reilly, M. M. Hereditary Sensory Neuropathy Type 1 (SPTLC1): phenotypic variation in patients with the English founder mutation. March 2017. *Neuromuscular Disorders* 27(1):S26-S27. doi: 10.1016/S0960-8966(17)30296-1

Sinclair, Christopher D.J., Morrow, J. M., Janiczek, R. L., **Evans, M. R. B.,** Rawah, E., Shah, S., … Thornton, J. S. (2016). Stability and sensitivity of water T2 obtained with IDEAL-CPMG in healthy and fat-infiltrated skeletal muscle. *NMR in Biomedicine*. doi: 10.1002/nbm.3654

8 References

Alizai, H., Nardo, L., Karampinos, D. C., Joseph, G. B., Yap, S. P., Baum, T., Krug, R., Majumdar, S., and Link, T. M. (2012) 'Comparison of clinical semi-quantitative assessment of muscle fat infiltration with quantitative assessment using chemical shift-based water/fat separation in MR studies of the calf of post-menopausal women', *European Radiology*, 22(7), 1592–1600.

Alizai, H., Chang, G., and Regatte, R. R. (2018) 'MR Imaging of the Musculoskeletal System Using Ultrahigh Field (7T) MR Imaging', *PET Clinics*, 551–565.

Amato, A. A., Sivakumar, K., Goyal, N., David, W. S., Salajegheh, M., Praestgaard, J., Lach-Trifilieff, E., Trendelenburg, A. U., Laurent, D., Glass, D. J., Roubenoff, R., Tseng, B. S., and Greenberg, S. A. (2014) 'Treatment of sporadic inclusion body myositis with bimagrumab', *Neurology*, 83(24), 2239–2246.

Andersen, G., Dahlqvist, J. R., Vissing, C. R., Heje, K., Thomsen, C., and Vissing, J. (2017) 'MRI as outcome measure in facioscapulohumeral muscular dystrophy: 1-year follow-up of 45 patients', *Journal of Neurology*, 264(3), 438–447.

Andreassen, C. S., Jakobsen, J., Ringgaard, S., Ejskjaer, N., and Andersen, H. (2009) 'Accelerated atrophy of lower leg and foot muscles-a follow-up study of long-term diabetic polyneuropathy using magnetic resonance imaging', *Diabetologia*, 52(7), 1182-1191.

Aquilonius, S. M, Askmark, H., Gillberg, P. G, Nandedkar, S., Olsson, Y., and Stålrberg, E. (1984) 'Topographical localization of motor endplates in cryosections of whole human muscles', *Muscle & Nerve*, 7(4), 287–293.

Arechavala-Gomeza, V., Anthony, K., Morgan, J., and Muntoni, F. (2012) 'Antisense Oligonucleotide-Mediated Exon Skipping for Duchenne Muscular Dystrophy: Progress and Challenges', *Current Gene Therapy*, 12(3), 152–160.

Arpan, I., Forbes, S. C., Lott, D. J., Senesac, C. R., Daniels, M. J., Triplett, W. T., Deol, J. K., Sweeney, H. L., Walter, G. A., and Vandenborne, K. (2013) '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 in Biomedicine*, 26(3), 320–328.

Arpan, I., Willcocks, R. J., Forbes, S. C., Finkel, R. S., Lott, D. J., Rooney, W. D., Triplett, W. T., Senesac, C. R., Daniels, M. J., Byrne, B. J., Finanger, E. L., Russman, B. S., Wang, D. J., Tennekoon, G. I., Walter, G. A., Sweeney, H. L., and Vandenborne, K. (2014) 'Examination of effects of corticosteroids on skeletal muscles of boys with DMD using MRI and MRS', *Neurology*, 83(11), 974–980.

Attarian, S., Vallat, J. M., Magy, L., Funalot, B., Gonnaud, P. M., Lacour, A., Péréon, Y., Dubourg, O., Pouget, J., Micallef, J., Franques, J., Lefebvre, M. N., Ghorab, K., Al-Moussawi, M., Tiffreau, V., Preudhomme, M., Magot, A., Leclair-Visonneau, L., Stojkovic, T., *et al.* (2014) 'An exploratory randomised double-blind and placebo-controlled phase 2 study of a combination of baclofen, naltrexone and sorbitol (PXT3003) in patients with Charcot-Marie-Tooth disease type 1A', *Orphanet Journal of Rare Diseases*, 9(1), 1-15.

Beaton, D. E., Hogg-Johnson, S., and Bombardier, C. (1997) 'Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders', *Journal of Clinical Epidemiology*, 50(1), 79–93.

Berciano, J., García, A., Calleja, J., and Combarros, O. (2000) 'Clinico-electrophysiological correlation of extensor digitorum brevis muscle atrophy in children with Charcot-Marie-Tooth disease 1A duplication', *Neuromuscular Disorders*, 10(6), 419–424.

Berciano, J., Gallardo, E., García, A., Infante, J., Mateo, I., and Combarros, O. (2006)

‘Charcot-Marie-Tooth disease type 1A duplication with severe paresis of the proximal lower limb muscles: A long-term follow-up study’, *Journal of Neurology, Neurosurgery and Psychiatry*, 77(10), 1169–1176.

Blat, Y. and Blat, S. (2015) ‘Drug discovery of therapies for duchenne muscular dystrophy’, *Journal of Biomolecular Screening*, 1189–1203.

Bonati, U., Schmid, M., Hafner, P., Haas, T., Bieri, O., Gloor, M., Fischmann, A., and Fischer, D. (2015) ‘Longitudinal 2-point dixon muscle magnetic resonance imaging in becker muscular dystrophy’, *Muscle and Nerve*, 51(6), 918–921.

Bonati, U., Hafner, P., Schädelin, S., Schmid, M., Naduvilekoot Devasia, A., Schroeder, J., Zuesli, S., Pohlman, U., Neuhaus, C., Klein, A., Sinnreich, M., Haas, T., Gloor, M., Bieri, O., Fischmann, A., and Fischer, D. (2015) ‘Quantitative muscle MRI: A powerful surrogate outcome measure in Duchenne muscular dystrophy’, *Neuromuscular Disorders*, 25(9), 679–685.

Bonati, U., Holiga, Š., Hellbach, N., Risterucci, C., Bergauer, T., Tang, W., Hafner, P., Thoeni, A., Bieri, O., Gerlach, I., Marquet, A., Khwaja, O., Sambataro, F., Bertolino, A., Dukart, J., Fischmann, A., Fischer, D., and Czech, C. (2017) ‘Longitudinal characterization of biomarkers for spinal muscular atrophy’, *Annals of Clinical and Translational Neurology*, 4(5), 292–304.

Bönnemann, C. G., Wang, C. H., Quijano-Roy, S., Deconinck, N., Bertini, E., Ferreiro, A., Muntoni, F., Sewry, C., Béroud, C., Mathews, K. D., Moore, S. A., Bellini, J., Rutkowski, A., and North, K. N. (2014) ‘Diagnostic approach to the congenital muscular dystrophies’, *Neuromuscular Disorders*, 24(4), 289–311.

Braathen, G. J. (2012) ‘Genetic epidemiology of Charcot-Marie-Tooth disease’, *Acta Neurologica Scandinavica*, 126(S193), 1–22.

Bryan, W. W., Reisch, J. S., McDonald, G., Herbelin, L. L., Barohn, R. J., and Fleckenstein, J. L. (1998) ‘Magnetic resonance imaging of muscle in amyotrophic lateral sclerosis’, *Neurology*, 51(1), 110–113.

Budzik, J. F., Balbi, V., Verclytte, S., Pansini, V., Le Thuc, V., and Cotten, A. (2014) ‘Diffusion tensor imaging in musculoskeletal disorders’, *Radiographics*, 34(3), E56-72.

Bugiardini, E., Morrow, J. M., Shah, S., Wood, C. L., Lynch, D. S., Pitmann, A. M., Reilly, M. M., Houlden, H., Matthews, E., Parton, M., Hanna, M. G., Straub, V., and Yousry, T. A. (2018) ‘The diagnostic value of MRI pattern recognition in distal myopathies’, *Frontiers in Neurology*, 2018 Jun 26;9:456.

Burns, J., Ouvrier, R. A., Yiu, E. M., Joseph, P. D., Kornberg, A. J., Fahey, M. C., and Ryan, M. M. (2009) ‘Ascorbic acid for Charcot-Marie-Tooth disease type 1A in children: a randomised, double-blind, placebo-controlled, safety and efficacy trial’, *The Lancet Neurology*, 8(6), 537–544.

Bushby, K., Finkel, R., Wong, B., Barohn, R., Campbell, C., Comi, G. P., Connolly, A. M., Day, J. W., Flanigan, K. M., Goemans, N., Jones, K. J., Mercuri, E., Quinlivan, R., Renfroe, J. B., Russman, B., Ryan, M. M., Tulinius, M., Voit, T., Moore, S. A., *et al.* (2014) ‘Ataluren treatment of patients with nonsense mutation dystrophinopathy’, *Muscle and Nerve*, 50(4), 477–487.

Buyse, G. M., Voit, T., Schara, U., Straathof, C. S. M., D’Angelo, M. G., Bernert, G., Cuisset, J. M., Finkel, R. S., Goemans, N., McDonald, C. M., Rummey, C., and Meier, T. (2015) ‘Efficacy of idebenone on respiratory function in patients with Duchenne muscular dystrophy not using glucocorticoids (DELOS): A double-blind randomised placebo-controlled phase 3 trial’, *The Lancet*, 385(9979), 1748–1757.

Carlier, P. G. (2014) ‘Global T2 versus water T2 in NMR imaging of fatty infiltrated muscles: Different methodology, different information and different implications’, *Neuromuscular Disorders*, 390–392.

Carlier, P. G., Azzabou, N., de Sousa, P. L., Hicks, A., Boisserie, J. M., Amadon, A.,

Carlier, R. Y., Wary, C., Orlikowski, D., and Laforêt, P. (2015) 'Skeletal muscle quantitative nuclear magnetic resonance imaging follow-up of adult Pompe patients', *Journal of Inherited Metabolic Disease*, 38(3), 565–572.

Chabanon, A., Seferian, A. M., Daron, A., Péréon, Y., Cances, C., Vuillerot, C., De Waele, L., Cuisset, J. M., Laugel, V., Schara, U., Gidaro, T., Gilabert, S., Hogrel, J. Y., Baudin, P. Y., Carlier, P., Fournier, E., Lowes, L. P., Hellbach, N., Seabrook, T., *et al.* (2018) 'Prospective and longitudinal natural history study of patients with Type 2 and 3 spinal muscular atrophy: Baseline data NatHis-SMA study', *PLoS ONE*, 13(7):e0201004.

Chan, W. P. and Liu, G. C. (2002) 'MR imaging of primary skeletal muscle diseases in children', *American Journal of Roentgenology*, 179(4), 989–997.

Chow S, Shao J, Wang H. 2008. *Sample Size Calculations in Clinical Research*. 2nd Ed. Chapman & Hall/CRC Biostatistics Series. page 51.

Chung, K. W., Suh, B. C., Shy, M. E., Cho, S. Y., Yoo, J. H., Park, S. W., Moon, H., Park, K. D., Choi, K. G., Kim, S., Kim, S. B., Shim, D. S., Kim, S. M., Sunwoo, I. N., and Choi, B. O. (2008) 'Different clinical and magnetic resonance imaging features between Charcot-Marie-Tooth disease type 1A and 2A', *Neuromuscular Disorders*, 18(8), 610–618.

Cirak, S., Arechavala-Gomeza, V., Guglieri, M., Feng, L., Torelli, S., Anthony, K., Abbs, S., Garralda, M. E., Bourke, J., Wells, D. J., Dickson, G., Wood, M. J., Wilton, S. D., Straub, V., Kole, R., Shrewsbury, S. B., Sewry, C., Morgan, J. E., Bushby, K., *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', *The Lancet*, 378(9791), 595–605.

Copay, A. G., Subach, B. R., Glassman, S. D., Polly, D. W., and Schuler, T. C. (2007) 'Understanding the minimum clinically important difference: a review of concepts and methods', *Spine Journal*, 541–546.

D'Ydewalle, C., Krishnan, J., Chiheb, D. M., Van Damme, P., Irobi, J., Kozikowski, A. P., Berghe, P. Vanden, Timmerman, V., Robberecht, W., and Van Den Bosch, L. (2011) 'HDAC6 inhibitors reverse axonal loss in a mouse model of mutant HSPB1-induced Charcot-Marie-Tooth disease', *Nature Medicine*, 17(8), 968–974.

Dahlqvist, J. R., Vissing, C. R., Thomsen, C., and Vissing, J. (2014) 'Severe paraspinal muscle involvement in facioscapulohumeral muscular dystrophy', *Neurology*, 83(13), 1178–1183.

Dahlqvist, J. R., Fornander, F., de Stricker Borch, J., Oestergaard, S. T., Poulsen, N. S., and Vissing, J. (2018) 'Disease progression and outcome measures in spinobulbar muscular atrophy', *Annals of Neurology*, 84(5), 754–765.

Dahlqvist, J. R., Oestergaard, S. T., Poulsen, N. S., Thomsen, C., and Vissing, J. (2019) 'Refining the spinobulbar muscular atrophy phenotype by quantitative MRI and clinical assessments', *Neurology*, 92(6), e548–e559.

ten Dam, L., van der Kooi, A. J., Verhamme, C., Wattjes, M. P., and de Visser, M. (2016) 'Muscle imaging in inherited and acquired muscle diseases', *European Journal of Neurology*, 688–703.

Damon, B. M., Heemskerk, A. M., and Ding, Z. (2012) 'Polynomial fitting of DT-MRI fiber tracts allows accurate estimation of muscle architectural parameters', *Magnetic Resonance Imaging*, 30(5), 589–600.

Degardin, A., Morillon, D., Lacour, A., Cotten, A., Vermersch, P., and Stojkovic, T. (2010) 'Morphologic imaging in muscular dystrophies and inflammatory myopathies', *Skeletal Radiology*, 39(12), 1219–1227.

Deoni, S. C. L., Rutt, B. K., and Peters, T. M. (2003) 'Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state', *Magnetic Resonance in Medicine*,

49(3), 515–526.

Diaz-Manera, J., Fernandez-Torron, R., Llauger, J., James, M. K., Mayhew, A., Smith, F. E., Moore, U. R., Blamire, A. M., Carlier, P. G., Rufibach, L., Mittal, P., Eagle, M., Jacobs, M., Hodgson, T., Wallace, D., Ward, L., Smith, M., Stramare, R., Rampado, A., *et al.* (2018) 'Muscle MRI in patients with dysferlinopathy: Pattern recognition and implications for clinical trials', *Journal of Neurology, Neurosurgery and Psychiatry*, 89(10), 1071–1081.

Dixon, W. T. (1984) 'Simple proton spectroscopic imaging', *Radiology*, 153(1), 189–194.

Eichinger, K., Burns, J., Cornett, K., Bacon, C., Shepherd, M. L., Mountain, J., Sowden, J., Shy, R., Shy, M. E., and Herrmann, D. N. (2018) 'The Charcot-Marie-Tooth functional outcome measure (CMT-FOM)', *Neurology*, 91(15), E1381–E1384.

Eichinger, K. J., Burns, J., Cornett, K., Bacon, C., Shepherd, M., Mountain, J., Sowden, J., Shy, R., Shy, M. E., and Herrmann, D. N. (2017) '2017 Peripheral Nerve Society Meeting July 8-12, 2017 Sitges, Barcelona, Spain', *Journal of the Peripheral Nervous System*.

Eichler, F. S., Hornemann, T., McCampbell, A., Kuljis, D., Penno, A., Vardeh, D., Tamrazian, E., Garofalo, K., Lee, H. J., Kini, L., Selig, M., Frosch, M., Gable, K., Von Eckardstein, A., Woolf, C. J., Guan, G., Harmon, J. M., Dunn, T. M., and Brown, R. H. (2009) 'Overexpression of the wild-type SPT1 subunit lowers desoxysphingolipid levels and rescues the phenotype of HSAN1', *Journal of Neuroscience*, 29(46), 14646–14651.

Fanin, M., Nardetto, L., Nascimbeni, A. C., Tasca, E., Spinazzi, M., Padoan, R., and Angelini, C. (2007) 'Correlations between clinical severity, genotype and muscle pathology in limb girdle muscular dystrophy type 2A', *Journal of Medical Genetics*, 44(10), 609–614.

Faridian-Aragh, N., Wagner, K. R., Leung, D. G., and Carrino, J. A. (2014) 'Magnetic resonance imaging phenotyping of Becker muscular dystrophy', *Muscle and Nerve*, 50(6), 962–967.

Feng, X., Luo, S., Li, J., Yue, D., Xi, J., Zhu, W., Gao, X., Guan, X., Lu, J., Liang, Z., and Zhao, C. (2018) 'Fatty infiltration evaluation and selective pattern characterization of lower limbs in limb-girdle muscular dystrophy type 2A by muscle magnetic resonance imaging', *Muscle and Nerve*, 58(4), 536–541.

Figueroa-Bonaparte, S., Llauger, J., Segovia, S., Belmonte, I., Pedrosa, I., Montiel, E., Montesinos, P., Sánchez-González, J., Alonso-Jiménez, A., Gallardo, E., Illa, I., Barba-Romero, M. A., Barcena, J., Carbonell, P., Carzorla, M. R., Creus, C., Coll-Cantí, J., Díaz, M., Domínguez, C., *et al.* (2018) 'Quantitative muscle MRI to follow up late onset Pompe patients: A prospective study', *Scientific Reports*, 18;8(1):10898.

Finkel, R. S., Mercuri, E., Darras, B. T., Connolly, A. M., Kuntz, N. L., Kirschner, J., Chiriboga, C. A., Saito, K., Servais, L., Tizzano, E., Topaloglu, H., Tulinius, M., Montes, J., Glanzman, A. M., Bishop, K., Zhong, Z. J., Gheuens, S., Bennett, C. F., Schneider, E., *et al.* (2017) 'Nusinersen versus sham control in infantile-onset spinal muscular atrophy', *New England Journal of Medicine*, 377(18), 1723–1732.

Fischer, D., Kley, R. A., Strach, K., Meyer, C., Sommer, T., Eger, K., Rolfs, A., Meyer, W., Pou, A., Pradas, J., Heyer, C. M., Grossmann, A., Huebner, A., Kress, W., Reimann, J., Schroder, R., Eymard, B., Fardeau, M., Udd, B., *et al.* (2008) 'Distinct muscle imaging patterns in myofibrillar myopathies', *Neurology*, 71(10), 758–765.

Fischmann, A., Gloor, M., Bieri, O., Fasler, S., Studler, U., and Fischer, D. (2011) 'Abstracts', *Clinical Neuroradiology*, 21(3), 185–196.

Fischmann, A., Hafner, P., Fasler, S., Gloor, M., Bieri, O., Studler, U., and Fischer, D. (2012) 'Quantitative MRI can detect subclinical disease progression in muscular dystrophy', *Journal of Neurology*, 259(8), 1648–1654.

Fischmann, A., Hafner, P., Gloor, M., Schmid, M., Klein, A., Pohlman, U., Waltz, T.,

Gonzalez, R., Haas, T., Bieri, O., and Fischer, D. (2013) 'Quantitative MRI and loss of free ambulation in Duchenne muscular dystrophy', *Journal of Neurology*, 260(4), 969–974.

Fischmann, A., Morrow, J. M., Sinclair, C. D. J., Reilly, M. M., Hanna, M. G., Yousry, T., and Thornton, J. S. (2014) 'Improved anatomical reproducibility in quantitative lower-limb muscle MRI', *Journal of Magnetic Resonance Imaging*, 39(4), 1033–1038.

Fledrich, R., Mannil, M., Leha, A., Ehbrecht, C., Solari, A., Pelayo-Negro, A. L., Berciano, J., Schlotter-Weigel, B., Schnizer, T. J., Prukop, T., Garcia-Angarita, N., Czesnik, D., Haberlová, J., Mazanec, R., Paulus, W., Beissbarth, T., Walter, M. C., Hogrel, J. Y., Dubourg, O., *et al.* (2017) 'Biomarkers predict outcome in Charcot-Marie-Tooth disease 1A', *Journal of Neurology, Neurosurgery and Psychiatry*, 88(11), 941–952.

Forbes, S. C., Willcocks, R. J., Triplett, W. T., Rooney, W. D., Lott, D. J., Wang, D. J., Pollaro, J., Senesac, C. R., Daniels, M. J., Finkel, R. S., Russman, B. S., Byrne, B. J., Finanger, E. L., Tennekoon, G. I., Walter, G. A., Sweeney, H. L., and Vandenborne, K. (2014) 'Magnetic resonance imaging and spectroscopy assessment of lower extremity skeletal muscles in boys with duchenne muscular dystrophy: A multicenter cross sectional study', *PLoS ONE*, 9(9): e106435.

de Freitas, M., Vidal, C., Dias, J. C., Bittar, C., Escada, T., Nevares, M. T., Nascimento, O. J. M., Marques, W., and Domingues, R. C. (2015) 'Calf hypertrophy in charcot-marie-tooth disease: report of nine cases', *Journal of the Neurological Sciences*, 357, e332.

Fridman, V., Bundy, B., Reilly, M. M., Pareyson, D., Bacon, C., Burns, J., Day, J., Feely, S., Finkel, R. S., Grider, T., Kirk, C. A., Herrmann, D. N., Laurá, M., Li, J., Lloyd, T., Sumner, C. J., Muntoni, F., Piscosquito, G., Ramchandren, S., *et al.* (2015) 'CMT subtypes and disease burden in patients enrolled in the Inherited Neuropathies Consortium natural history study: A cross-sectional analysis', *Journal of Neurology, Neurosurgery and Psychiatry*, 86(8), 873–878.

Fridman, V., Suriyanarayanan, S., Novak, P., David, W., Macklin, E. A., McKenna-Yasek, D., Walsh, K., Aziz-Bose, R., Oaklander, A. L., Brown, R., Hornemann, T., and Eichler, F. (2019) 'Randomized trial of l -serine in patients with hereditary sensory and autonomic neuropathy type 1', *Neurology*, 92(4), E359–E370.

Fridman, V., Sillau, S., Acsadi, G., Bacon, C., Dooley, K., Burns, J., Day, J., Feely, S., Finkel, R. S., Grider, T., Gutmann, L., Herrmann, D. N., Kirk, C. A., Knause, S. A., Laurá, M., Lewis, R. A., Li, J., Lloyd, T. E., Moroni, I., *et al.* (2020) 'A longitudinal study of CMT1A using Rasch analysis based CMT neuropathy and examination scores', *Neurology*, e884–e896.

Friedman, S. D., Poliachik, S. L., Otto, R. K., Carter, G. T., Budech, C. B., Bird, T. D., Miller, D. G., and Shaw, D. W. W. (2014) 'Longitudinal features of stir bright signal in FSHD1', *Muscle and Nerve*, 49(2), 257–260.

Gadermayr, M., Disch, C., Müller, M., Merhof, D., and Gess, B. (2018) 'A comprehensive study on automated muscle segmentation for assessing fat infiltration in neuromuscular diseases', *Magnetic Resonance Imaging*, 48, 20–26.

Gaeta, M., Mileto, A., Mazzeo, A., Minutoli, F., Di Leo, R., Settineri, N., Donato, R., Ascenti, G., and Blandino, A. (2012) 'MRI findings, patterns of disease distribution, and muscle fat fraction calculation in five patients with Charcot-Marie-Tooth type 2 F disease', *Skeletal Radiology*, 41(5), 515–524.

Gallardo, E., García, A., Combarros, O., and Berciano, J. (2006) 'Charcot-Marie-Tooth disease type 1A duplication: Spectrum of clinical and magnetic resonance imaging features in leg and foot muscles', *Brain*, 129(2), 426–437.

Gallardo, E., Claeys, K. G., Nelis, E., García, A., Canga, A., Combarros, O., Timmerman, V., De Jonghe, P., and Berciano, J. (2008) 'Magnetic resonance imaging findings of leg musculature in Charcot-Marie-Tooth disease type 2 due to dynamin 2 mutation', *Journal of Neurology*, 255, 986–992.

Gallardo, E., García, A., Ramón, C., Maraví, E., Infante, J., Gastón, I., Alonso, Á., Combarros, O., De Jonghe, P., and Berciano, J. (2009) 'Charcot-Marie-Tooth disease type 2J with MPZ Thr124Met mutation: Clinico-electrophysiological and MRI study of a family', *Journal of Neurology*, 256(12), 2061–2071.

Garofalo, K., Penno, A., Schmidt, B. P., Lee, H. J., Frosch, M. P., Von Eckardstein, A., Brown, R. H., Hornemann, T., and Eichler, F. S. (2011a) 'Oral L-serine supplementation reduces production of neurotoxic deoxysphingolipids in mice and humans with hereditary sensory autonomic neuropathy type 1', *Journal of Clinical Investigation*, 121(12), 4735–4745.

Gess, B., Baets, J., De Jonghe, P., Reilly, M. M., Pareyson, D., and Young, P. (2015) 'Ascorbic acid for the treatment of Charcot-Marie-Tooth disease', *Cochrane Database of Systematic Reviews*.

Glover, G. H. and Schneider, E. (1991) 'Three- point dixon technique for true water/fat decomposition with B0 inhomogeneity correction', *Magnetic Resonance in Medicine*, 18(2), 371–383.

Del Grande, F., Carrino, J. A., Del Grande, M., Mammen, A. L., and Stine, L. C. (2011) 'Magnetic resonance imaging of inflammatory myopathies', *Topics in Magnetic Resonance Imaging*, 39–43.

Grimm, A., Meyer, H., Nickel, M. D., Nittka, M., Raithel, E., Chaudry, O., Friedberger, A., Uder, M., Kemmler, W., Engelke, K., and Quick, H. H. (2019) 'A Comparison between 6-point Dixon MRI and MR Spectroscopy to Quantify Muscle Fat in the Thigh of Subjects with Sarcopenia', *The Journal of frailty & aging*, 8(1), 21–26.

Gutmann, L. and Shy, M. (2015) 'Update on Charcot-Marie-Tooth disease', *Current Opinion in Neurology*, 28(5), 462–467.

Guyatt, G., Walter, S., and Norman, G. (1987) 'Measuring change over time: Assessing the usefulness of evaluative instruments', *Journal of Chronic Diseases*, 40(2), 171–178.

Haas, M., Vlcek, V., Balabanov, P., Salmonson, T., Bakchine, S., Markey, G., Weise, M., Schlosser-Weber, G., Brohmann, H., Yerro, C. P., Mendizabal, M. R., Stoyanova-Beninska, V., and Hillege, H. L. (2015) 'European medicines agency review of ataluren for the treatment of ambulant patients aged 5 years and older with Duchenne muscular dystrophy resulting from a nonsense mutation in the dystrophin gene', *Neuromuscular Disorders*, 25(1), 5–13.

Halpern, S. D., Karlawish, J. H. T., and Berlin, J. A. (2002) 'The continuing unethical conduct of underpowered clinical trials', *Journal of the American Medical Association*, 288(3), 358–362.

Hankiewicz, K., Carlier, R. Y., Lazaro, L., Linzoain, J., Barnerias, C., Gómez-Andrés, D., Avila-Smirnow, D., Ferreiro, A., Estournet, B., Guicheney, P., Germain, D. P., Richard, P., Bulacio, S., Mompoint, D., and Quijano-Roy, S. (2015) 'Whole-body muscle magnetic resonance imaging in SEPN1-related myopathy shows a homogeneous and recognizable pattern', *Muscle and Nerve*, 52(5), 728–735.

Herlidou, S., Rolland, Y., Bansard, J. Y., Le Rumeur, E., and De Certaines, J. D. (1999) 'Comparison of automated and visual texture analysis in MRI: Characterization of normal and diseased skeletal muscle', *Magnetic Resonance Imaging*, 17(9), 1393–1397.

Hiba, B., Richard, N., Hébert, L. J., Coté, C., Nejjari, M., Vial, C., Bouhour, F., Puymirat, J., and Janier, M. (2012) 'Quantitative assessment of skeletal muscle degeneration in patients with myotonic dystrophy type 1 using MRI', *Journal of Magnetic Resonance Imaging*, 35(3), 678–685.

Hogrel, J. Y., Wary, C., Moraux, A., Azzabou, N., Decostre, V., Ollivier, G., Canal, A., Lilien, C., Ledoux, I., Annoussamy, M., Reguiba, N., Gidaro, T., Le Moing, A. G., Cardas, R., Voit, T., Carlier, P. G., and Servais, L. (2016) 'Longitudinal functional and NMR assessment of upper limbs in Duchenne muscular dystrophy', *Neurology*, 86(11), 1022–

1030.

Hollingsworth, K. G., de Sousa, P. L., Straub, V., and Carlier, P. G. (2012) 'Towards harmonization of protocols for MRI outcome measures in skeletal muscle studies: Consensus recommendations from two TREAT-NMD NMR workshops, 2 May 2010, Stockholm, Sweden, 1-2 October 2009, Paris, France', *Neuromuscular Disorders*, 22(SUPPL. 2).

Hollingsworth, K. G., Garrood, P., Eagle, M., Bushby, K., and Straub, V. (2013) 'Magnetic resonance imaging in duchenne muscular dystrophy: Longitudinal assessment of natural history over 18 months', *Muscle and Nerve*, 48(4), 586–588.

Hollingsworth, K. G. (2014) 'Quantitative MRI in muscular dystrophy: An indispensable trial endpoint?', *Neurology*, 956–957.

Hooijmans, M. T., Niks, E. H., Burakiewicz, J., Anastasopoulos, C., van den Berg, S. I., van Zwet, E., Webb, A. G., Verschuuren, J. J. G. M., and Kan, H. E. (2017) 'Non-uniform muscle fat replacement along the proximodistal axis in Duchenne muscular dystrophy', *Neuromuscular Disorders*, 27(5), 458–464.

Houlden, H., King, R., Blake, J., Groves, M., Love, S., Woodward, C., Hammans, S., Nicoll, J., Lennox, G., O'Donovan, D. G., Gabriel, C., Thomas, P. K., and Reilly, M. M. (2006) 'Clinical, pathological and genetic characterization of hereditary sensory and autonomic neuropathy type I (HSAN I)', *Brain*, 129(2), 411–425.

Houlden, H., Blake, J., and Reilly, M. M. (2004) 'Hereditary sensory neuropathies', *Current Opinion in Neurology*, 569–577.

Hu, H. H., Börnert, P., Hernando, D., Kellman, P., Ma, J., Reeder, S., and Sirlin, C. (2012) 'ISMRM workshop on fat-water separation: Insights, applications and progress in MRI', *Magnetic Resonance in Medicine*, 68(2), 378–388.

Huang, Y., Majumdar, S., Genant, H. K., Chan, W. P., Sharma, K. R., Yu, P., Mynhier, M., and Miller, R. G. (1994) 'Quantitative MR relaxometry study of muscle composition and function in duchenne muscular dystrophy', *Journal of Magnetic Resonance Imaging*, 4(1), 59–64.

Husted, J. A., Cook, R. J., Farewell, V. T., and Gladman, D. D. (2000) 'Methods for assessing responsiveness: A critical review and recommendations', *Journal of Clinical Epidemiology*, 53(5), 459–468.

Jacobson, N. S., Roberts, L. J., Berns, S. B., and McGlinchey, J. B. (1999) 'Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives', *Journal of Consulting and Clinical Psychology*, 300–307.

Jaeschke, R., Singer, J., and Guyatt, G. H. (1989) 'Measurement of health status. Ascertaining the minimal clinically important difference', *Controlled Clinical Trials*, 10(4), 407–415.

Janiczek, R. L., Gambarota, G., Sinclair, C. D. J., Yousry, T. A., Thornton, J. S., Golay, X., and Newbould, R. D. (2011) 'Simultaneous T 2 and lipid quantitation using IDEAL-CPMG', *Magnetic Resonance in Medicine*, 66(5), 1293–1302.

Janssen, B. H., Voet, N. B. M., Nabuurs, C. I., Kan, H. E., De Rooy, J. W. J., Geurts, A. C., Padberg, G. W., Van Engelen, B. G. M., and Heerschap, A. (2014) 'Distinct disease phases in muscles of facioscapulohumeral dystrophy patients identified by MR detected fat infiltration', *PLoS ONE*, 9(1): e85416.

Jerath, N. U. and Shy, M. E. (2015) 'Hereditary motor and sensory neuropathies: Understanding molecular pathogenesis could lead to future treatment strategies', *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 667–678.

Jungbluth, H. (2017) 'Myopathology in times of modern imaging', *Neuropathology and Applied Neurobiology*, 24–43.

Juras, V., Mlynarik, V., Szomolanyi, P., Valkovič, L., and Trattnig, S. (2019) 'Magnetic

Resonance Imaging of the Musculoskeletal System at 7T: Morphological Imaging and Beyond', *Topics in Magnetic Resonance Imaging*, 28(3), 125–135.

Jurkat-Rott, K., Weber, M. A., Fauler, M., Guo, X. H., Holzherr, B. D., Paczulla, A., Nordsborg, N., Joechle, W., and Lehmann-Horn, F. (2009) 'K+-dependent paradoxical membrane depolarization and Na + overload, major and reversible contributors to weakness by ion channel leaks', *Proceedings of the National Academy of Sciences of the United States of America*, 106(10), 4036–4041.

Kan, H. E., Scheenen, T. W. J., Wohlgemuth, M., Klomp, D. W. J., van Loosbroek-Wagenmans, I., Padberg, G. W., and Heerschap, A. (2009) 'Quantitative MR imaging of individual muscle involvement in facioscapulohumeral muscular dystrophy', *Neuromuscular Disorders*, 19(5), 357–362.

Karampinos, D. C., Yu, H., Shimakawa, A., Link, T. M., and Majumdar, S. (2011) 'T 1-corrected fat quantification using chemical shift-based water/fat separation: Application to skeletal muscle', *Magnetic Resonance in Medicine*, 66(5), 1312–1326.

Karlsson, A., Rosander, J., Romu, T., Tallberg, J., Grönqvist, A., Borga, M., and Dahlqvist Leinhard, O. (2015) 'Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI', *Journal of Magnetic Resonance Imaging*, 41(6), 1558–1569.

Katirji, M. B. (1988) 'Aids to the Examination of the Peripheral Nervous System', *Neurology*, 38(10), 1663–1663.

Kemnitz, J., Eckstein, F., Culvenor, A. G., Ruhdorfer, A., Dannhauer, T., Ring-Dimitriou, S., Sänger, A. M., and Wirth, W. (2017) 'Validation of an active shape model-based semi-automated segmentation algorithm for the analysis of thigh muscle and adipose tissue cross-sectional areas', *Magnetic Resonance Materials in Physics, Biology and Medicine*, 30(5), 489–503.

Keszei, A. P., Novak, M., and Streiner, D. L. (2010) 'Introduction to health measurement scales', *Journal of Psychosomatic Research*, 319–323.

Kim, H. K., Laor, T., Horn, P. S., and Wong, B. (2010) 'Quantitative assessment of the T2 relaxation time of the gluteus muscles in children with duchenne muscular dystrophy: A comparative study before and after steroid treatment', *Korean Journal of Radiology*, 11(3), 304–311.

Kim, H. K., Laor, T., Horn, P. S., Racadio, J. M., Wong, B., and Dardzinski, B. J. (2010) 'T2 mapping in Duchenne muscular dystrophy: Distribution of disease activity and correlation with clinical assessments', *Radiology*, 255(3), 899–908.

Kim, H. S., Yoon, Y. C., Choi, B. O., Jin, W., and Cha, J. G. (2019) 'Muscle fat quantification using magnetic resonance imaging: case–control study of Charcot-Marie-Tooth disease patients and volunteers', *Journal of Cachexia, Sarcopenia and Muscle*, 10(3), 574–585.

Kirshner, B. and Guyatt, G. (1985) 'A methodological framework for assessing health indices', *Journal of Chronic Diseases*, 38(1), 27–36.

Krajewski, K. M., Lewis, R. A., Fuerst, D. R., Turansky, C., Hinderer, S. R., Garbern, J., Kamholz, J., and Shy, M. E. (2000) 'Neurological dysfunction and axonal degeneration in Charcot-Marie-Tooth disease type 1A', *Brain*, 123(7), 1516–1527.

Kugathasan, U., Evans, M. R. B., Morrow, J. M., Sinclair, C. D. J., Thornton, J. S., Yousry, T. A., Hornemann, T., Suriyanarayanan, S., Owusu-Ansah, K., Lauria, G., Lombardi, R., Polke, J. M., Wilson, E., Bennett, D. L. H., Houlden, H., Hanna, M. G., Blake, J. C., Laura, M., and Reilly, M. M. (2019) 'Development of MRC Centre MRI calf muscle fat fraction protocol as a sensitive outcome measure in Hereditary Sensory Neuropathy Type 1', *Journal of Neurology, Neurosurgery and Psychiatry*, 90(8):895-906

Kurihara, S., Adachi, Y., Wada, K., Awaki, E., Harada, H., and Nakashima, K. (2002) 'An epidemiological genetic study of Charcot-Marie-Tooth disease in western Japan',

Neuroepidemiology, 21(5), 246–250.

Lareau-Trudel, E., Troter, A. Le, Ghattas, B., Pouget, J., Attarian, S., Bendahan, D., and Salort-Campana, E. (2015) 'Muscle quantitative MR imaging and clustering analysis in patients with facioscapulohumeral muscular dystrophy type 1', *PLoS ONE*, 10(7): e0132717.

Lefter, S., Hardiman, O., and Ryan, A. M. (2017) 'A population-based epidemiologic study of adult neuromuscular disease in the Republic of Ireland', *Neurology*, 88(3), 304–313.

Lehman, L. A. and Velozo, C. A. (2010) 'Ability to detect change in patient function: Responsiveness designs and methods of calculation', *Journal of Hand Therapy*, 23(4), 361–371.

Lencioni, T., Piscosquito, G., Rabuffetti, M., Bovi, G., Di Sipio, E., Diverio, M., Moroni, I., Padua, L., Pagliano, E., Schenone, A., Pareyson, D., and Ferrarin, M. (2017) 'Responsiveness of gait analysis parameters in a cohort of 71 CMT subjects', *Neuromuscular Disorders*, 27(11), 1029–1037.

Lewis, R. A., McDermott, M. P., Herrmann, D. N., Hoke, A., Clawson, L. L., Siskind, C., Feely, S. M. E., Miller, L. J., Barohn, R. J., Smith, P., Luebbe, E., Wu, X., and Shy, M. E. (2013) 'High-dosage ascorbic acid treatment in charcot-marie-tooth disease type 1A results of a randomized, double-masked, controlled trial', *JAMA Neurology*, 70(8), 981–987.

Liang, M. H., Fossel, A. H., and Larson, M. G. (1990) *Comparisons of five health status instruments for orthopedic evaluation*, *Medical Care*.

Liu, G. C., Jong, Y. J., Chiang, C. H., and Jaw, T. S. (1993) 'Duchenne muscular dystrophy: MR grading system with functional correlation', *Radiology*, 186(2), 475–480.

Liu, M., Chino, N., and Ishihara, T. (1993) 'Muscle damage progression in Duchenne muscular dystrophy evaluated by a new quantitative computed tomography method', *Archives of Physical Medicine and Rehabilitation*, 74(5), 507–514.

Lodi, R., Muntoni, F., Taylor, J., Kumar, S., Sewry, C. A., Blamire, A., Styles, P., and Taylor, D. J. (1997) 'Correlative MR imaging and 31P-MR spectroscopy study in sarcoglycan deficient limb girdle muscular dystrophy', *Neuromuscular Disorders*, 7(8), 505–511.

Lu, C. H., Kalmar, B., Malaspina, A., Greensmith, L., and Petzold, A. (2011) 'A method to solubilise protein aggregates for immunoassay quantification which overcomes the neurofilament "hook" effect', *Journal of Neuroscience Methods*, 195(2), 143–150.

Lupski, J. R., de Oca-Luna, R. M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B. J., Saucedo-Cardenas, O., Barker, D. F., Killian, J. M., Garcia, C. A., Chakravarti, A., and Patel, P. I. (1991) 'DNA duplication associated with Charcot-Marie-Tooth disease type 1A', *Cell*, 66(2), 219–232.

Ma, J. (2008) 'Dixon techniques for water and fat imaging', *Journal of Magnetic Resonance Imaging*, 543–558.

Madelin, G. and Regatte, R. R. (2013) 'Biomedical applications of sodium MRI in vivo', *Journal of Magnetic Resonance Imaging*, 38(3), 511–529.

Maillard, S. M., Jones, R., Owens, C., Pilkington, C., Woo, P., Wedderburn, L. R., and Murray, K. J. (2004) 'Quantitative assessment of MRI T2 relaxation time of thigh muscles in juvenile dermatomyositis', *Rheumatology*, 43(5), 603–608.

Makrogiannis, S., Serai, S., Fishbein, K. W., Schreiber, C., Ferrucci, L., and Spencer, R. G. (2012) 'Automated quantification of muscle and fat in the thigh from water-, fat-, and nonsuppressed MR images', *Journal of Magnetic Resonance Imaging*, 35(5), 1152–1161.

Mankodi, A., Bishop, C. A., Auh, S., Newbould, R. D., Fischbeck, K. H., and Janiczek,

R. L. (2016) ‘Quantifying disease activity in fatty-infiltrated skeletal muscle by IDEAL-CPMG in Duchenne muscular dystrophy’, *Neuromuscular Disorders*, 26(10), 650–658.

Mannil, M., Solari, A., Leha, A., Pelayo-Negro, A. L., Berciano, J., Schlotter-Weigel, B., Walter, M. C., Rautenstrauss, B., Schnizer, T. J., Schenone, A., Seeman, P., Kadian, C., Schreiber, O., Angarita, N. G., Fabrizi, G. M., Gemignani, F., Padua, L., Santoro, L., Quattrone, A., *et al.* (2014) ‘Selected items from the Charcot-Marie-Tooth (CMT) Neuropathy Score and secondary clinical outcome measures serve as sensitive clinical markers of disease severity in CMT1A patients’, *Neuromuscular Disorders*, 24(11), 1003–1017.

Marden, F. A., Connolly, A. M., Siegel, M. J., and Rubin, D. A. (2005) ‘Compositional analysis of muscle in boys with Duchenne muscular dystrophy using MR imaging’, *Skeletal Radiology*, 34(3), 140–148.

McDaniel, J. D., Ulmer, J. L., Prost, R. W., Franczak, M. B., Jaradeh, S., Hamilton, C. A., and Mark, L. P. (1999) ‘Magnetization transfer imaging of skeletal muscle in autosomal recessive limb girdle muscular dystrophy’, *Journal of Computer Assisted Tomography*, 23(4), 609–614.

McDonald, C. M., Campbell, C., Torricelli, R. E., Finkel, R. S., Flanigan, K. M., Goemans, N., Heydemann, P., Kaminska, A., Kirschner, J., Muntoni, F., Osorio, A. N., Schara, U., Sejersen, T., Shieh, P. B., Sweeney, H. L., Topaloglu, H., Tulinius, M., Vilchez, J. J., Voit, T., *et al.* (2017) ‘Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial’, *The Lancet*, 390(10101), 1489–1498.

Mercuri, E., Talim, B., Moghadaszadeh, B., Petit, N., Brockington, M., Counsell, S., Guicheney, P., Muntoni, F., and Merlini, L. (2002) ‘Clinical and imaging findings in six cases of congenital muscular dystrophy with rigid spine syndrome linked to chromosome 1p (RSMD1)’, *Neuromuscular Disorders*, 12(7–8), 631–638.

Mercuri, E., Bushby, K., Ricci, E., Birchall, D., Pane, M., Kinali, M., Allsop, J., Nigro, V., Sáenz, A., Nascimbeni, A., Fulizio, L., Angelini, C., and Muntoni, F. (2005) ‘Muscle MRI findings in patients with limb girdle muscular dystrophy with calpain 3 deficiency (LGMD2A) and early contractures’, *Neuromuscular Disorders*, 15(2), 164–171.

Mercuri, E., Pichiecchio, A., Allsop, J., Messina, S., Pane, M., and Muntoni, F. (2007) ‘Muscle MRI in inherited neuromuscular disorders: Past, present, and future’, *Journal of Magnetic Resonance Imaging*, 433–440.

Mercuri, E., Darras, B. T., Chiriboga, C. A., Day, J. W., Campbell, C., Connolly, A. M., Iannaccone, S. T., Kirschner, J., Kuntz, N. L., Saito, K., Shieh, P. B., Tulinius, M., Mazzone, E. S., Montes, J., Bishop, K. M., Yang, Q., Foster, R., Gheuens, S., Bennett, C. F., *et al.* (2018) ‘Nusinersen versus sham control in later-onset spinal muscular atrophy’, *New England Journal of Medicine*, 378(7), 625–635.

Micallef, J., Attarian, S., Dubourg, O., Gonnaud, P. M., Hogrel, J. Y., Stojkovic, T., Bernard, R., Jouve, E., Pitel, S., Vacherot, F., Remec, J. F., Jomir, L., Azabou, E., Al-Moussawi, M., Lefebvre, M. N., Attolini, L., Yaici, S., Tanesse, D., Fontes, M., *et al.* (2009) ‘Effect of ascorbic acid in patients with Charcot-Marie-Tooth disease type 1A: a multicentre, randomised, double-blind, placebo-controlled trial’, *The Lancet Neurology*, 1103–1110.

Morrow, J. M., Pitceathly, R. D. S., Quinlivan, R. M., and Yousry, T. A. (2013) ‘Muscle MRI in Bethlem myopathy’, *BMJ Case Reports*.

Morrow, J. M., Sinclair, C. D. J., Fischmann, A., Reilly, M. M., Hanna, M. G., Yousry, T. A., and Thornton, J. S. (2014) ‘Reproducibility, and age, body-weight and gender dependency of candidate skeletal muscle MRI outcome measures in healthy volunteers’, *European Radiology*, 24(7), 1610–1620.

Morrow, J. M., Sinclair, C. D. J., Fischmann, A., Machado, P. M., Reilly, M. M., Yousry, T. A., Thornton, J. S., and Hanna, M. G. (2016) ‘MRI biomarker assessment of

neuromuscular disease progression: A prospective observational cohort study', *The Lancet Neurology*, 15(1), 65–77.

Morrow, J. M., Evans, M. R. B., Grider, T., Sinclair, C. D. J., Thedens, D., Shah, S., Yousry, T. A., Hanna, M. G., Nopoulos, P., Thornton, J. S., Shy, M. E., and Reilly, M. M. (2018) 'Validation of MRC Centre MRI calf muscle fat fraction protocol as an outcome measure in CMT1A', *Neurology*, 91(12), E1125–E1129.

Murphy, A. P., Morrow, J., Dahlqvist, J. R., Stojkovic, T., Willis, T. A., Sinclair, C. D. J., Wastling, S., Yousry, T., Hanna, M. S., James, M. K., Mayhew, A., Eagle, M., Lee, L. E., Hogrel, J. Y., Carlier, P. G., Thornton, J. S., Vissing, J., Hollingsworth, K. G., and Straub, V. (2019) 'Natural history of limb girdle muscular dystrophy R9 over 6 years: searching for trial endpoints', *Annals of Clinical and Translational Neurology*, 6(6), 1033–1045.

Murphy, S. M., Herrmann, D. N., McDermott, M. P., Scherer, S. S., Shy, M. E., Reilly, M. M., and Pareyson, D. (2011) 'Reliability of the CMT neuropathy score (second version) in Charcot-Marie-Tooth disease', *Journal of the Peripheral Nervous System*, 16(3), 191–198.

Nicholson, G. A., Dawkins, J. L., Blair, I. P., Auer-Grumbach, M., Brahmbhatt, S. B., and Hulme, D. J. (2001) 'Hereditary sensory neuropathy type I: Haplotype analysis shows founders in southern England and Europe', *American Journal of Human Genetics*, 69(3), 655–659.

O'Brien, M.D. (2000). Aids to the Examination of the Peripheral Nervous System (Elsevier Health Sciences).

Oddy, M. J., Brown, C., Mistry, R., and Eastwood, D. M. (2006) 'Botulinum toxin injection site localization for the tibialis posterior muscle', *Journal of Pediatric Orthopaedics Part B*, 15(6), 414–417.

Ouvrier, R. (2010) 'What can we learn from the history of Charcot-Marie-Tooth disease?', *Developmental Medicine and Child Neurology*, 405–406.

Paradas, C., Llauger, J., Diaz-Manera, J., Rojas-García, R., De Luna, N., Iturriaga, C., Márquez, C., Usón, M., Hankiewicz, K., Gallardo, E., and Illa, I. (2010) 'Redefining dysferlinopathy phenotypes based on clinical findings and muscle imaging studies', *Neurology*, 75(4), 316–323.

Pareyson, D., Reilly, M. M., Schenone, A., Fabrizi, G. M., Cavallaro, T., Manganelli, L., Vita, G., Quattrone, A., Padua, L., Gemignani, F., Visioli, F., Laurà, M., Radice, D., Calabrese, D., Hughes, R. A. C., Solari, A., Salsano, C., Nanetti, L., Marelli, C., *et al.* (2011) 'Ascorbic acid in charcot-marie-tooth disease type 1A (CMTTRIAAL and CMT-TRAUK): A double-blind randomised trial', *The Lancet Neurology*, 10(4), 320–328.

Patzkó, Á., Bai, Y., Saporta, M. A., Katona, I., Wu, X., Vizzuso, D., Feltri, M. L., Wang, S., Dillon, L. M., Kamholz, J., Kirschner, D., Sarkar, F. H., Wrabetz, L., and Shy, M. E. (2012) 'Curcumin derivatives promote Schwann cell differentiation and improve neuropathy in R98C CMT1B mice', *Brain*, 135(12), 3551–3566.

Pelayo-Negro, A. L., Gallardo, E., García, A., Sánchez-Juan, P., Infante, J., and Berciano, J. (2014) 'Evolution of Charcot-Marie-Tooth disease type 1A duplication: A 2-year clinico-electrophysiological and lower-limb muscle MRI longitudinal study', *Journal of Neurology*, 261, 675–685.

Penno, A., Reilly, M. M., Houlden, H., Laurá, M., Rentsch, K., Niederkofler, V., Stoeckli, E. T., Nicholson, G., Eichler, F., Brown, R. H., Von Eckardstein, A., and Hornemann, T. (2010) 'Hereditary sensory neuropathy type 1 is caused by the accumulation of two neurotoxic sphingolipids', *Journal of Biological Chemistry*, 285(15), 11178–11187.

Phoenix, J., Betal, D., Roberts, N., Helliwell, T. R., and Edwards, R. H. T. (1996) 'Objective quantification of muscle and fat in human dystrophic muscle by magnetic resonance image analysis', *Muscle and Nerve*, 19(3), 302–310.

Pichiecchio, A., Poloni, G. U., Ravaglia, S., Ponzio, M., Germani, G., Maranzana, D.,

Costa, A., Repetto, A., Tavazzi, E., Danesino, C., Moglia, A., and Bastianello, S. (2009) 'Enzyme replacement therapy in adult-onset glycogenosis II: Is quantitative muscle MRI helpful?', *Muscle and Nerve*, 40(1), 122–125.

Pichiecchio, A., Alessandrino, F., Bortolotto, C., Cerica, A., Rosti, C., Raciti, M. V., Rossi, M., Berardinelli, A., Baranello, G., Bastianello, S., and Calliada, F. (2018) 'Muscle ultrasound elastography and MRI in preschool children with Duchenne muscular dystrophy', *Neuromuscular Disorders*, 28(6), 476–483.

Piscosquito, G., Reilly, M. M., Schenone, A., Fabrizi, G. M., Cavallaro, T., Santoro, L., Manganelli, F., Vita, G., Quattrone, A., Padua, L., Gemignani, F., Visioli, F., Laurà, M., Calabrese, D., Hughes, R. A. C., Radice, D., Solari, A., and Pareyson, D. (2015) 'Responsiveness of clinical outcome measures in Charcot-Marie-Tooth disease', *European Journal of Neurology*, 22(12), 1556–1563.

van der Ploeg, A., Carlier, P. G., Carlier, R. Y., Kissel, J. T., Schoser, B., Wenninger, S., Pestronk, A., Barohn, R. J., Dimachkie, M. M., Goker-Alpan, O., Mozaffar, T., Pena, L. D. M., Simmons, Z., Straub, V., Guglieri, M., Young, P., Boentert, M., Baudin, P. Y., Wens, S., *et al.* (2016) 'Prospective exploratory muscle biopsy, imaging, and functional assessment in patients with late-onset Pompe disease treated with alglucosidase alfa: The EMBASSY Study', *Molecular Genetics and Metabolism*, 119(1–2), 115–123.

Poliachik, S. L., Friedman, S. D., Carter, G. T., Parnell, S. E., and Shaw, D. W. (2012) 'Skeletal Muscle Edema in Muscular Dystrophy: Clinical and Diagnostic Implications', *Physical Medicine and Rehabilitation Clinics of North America*, 107–122.

Ponrartana, S., Andrade, K. E., Wren, T. A. L., Ramos-Platt, L., Hu, H. H., Bluml, S., and Gilsanz, V. (2014) 'Repeatability of chemical-shift-encoded water-fat MRI and diffusion-tensor imaging in lower extremity muscles in children', *American Journal of Roentgenology*, 202(6):W567-73.

Ponrartana, S., Ramos-Platt, L., Wren, T. A. L., Hu, H. H., Perkins, T. G., Chia, J. M., and Gilsanz, V. (2015) 'Effectiveness of diffusion tensor imaging in assessing disease severity in Duchenne muscular dystrophy: preliminary study', *Pediatric Radiology*, 45(4), 582–589.

Positano, V., Christiansen, T., Santarelli, M. F., Ringgaard, S., Landini, L., and Gastaldelli, A. (2009) 'Accurate segmentation of subcutaneous and intermuscular adipose tissue from MR images of the thigh', *Journal of Magnetic Resonance Imaging*, 29(3), 677–684.

Prentice, R. L. (1989) 'Surrogate endpoints in clinical trials: Definition and operational criteria', *Statistics in Medicine*, 8(4), 431–440.

Qi, J., Olsen, N. J., Price, R. R., Winston, J. A., and Park, J. H. (2008) 'Diffusion-weighted imaging of inflammatory myopathies: Polymyositis and dermatomyositis', *Journal of Magnetic Resonance Imaging*, 27(1), 212–217.

R., R. and A., P. (2017) 'Texture-based segmentation of skeletal muscle in Dixon MRI images', *Journal of Neuromuscular Diseases*, 4, S27–S28. Available at: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L619419347.

Raeymaekers, P., Timmerman, V., Nelis, E., De Jonghe, P., Hoogenduk, J. E., Baas, F., Barker, D. F., Martin, J. J., De Visser, M., Bolhuis, P. A., and Van Broeckhoven, C. (1991) 'Duplication in chromosome 17p11.2 in Charcot-Marie-Tooth neuropathy type 1a (CMT 1a)', *Neuromuscular Disorders*, 1(2), 93–97.

Reeder, S. B., McKenzie, C. A., Pineda, A. R., Yu, H., Shimakawa, A., Brau, A. C., Hargreaves, B. A., Gold, G. E., and Brittain, J. H. (2007) 'Water-fat separation with IDEAL gradient-echo imaging', *Journal of Magnetic Resonance Imaging*, 25(3), 644–652.

Regula, J. U., Jestaedt, L., Jende, F., Bartsch, A., Meinck, H. M., and Weber, M. A.

(2016) ‘Clinical Muscle Testing Compared with Whole-Body Magnetic Resonance Imaging in Facio-scapulo-humeral Muscular Dystrophy’, *Clinical Neuroradiology*, 26(4), 445–455.

Reilly, M. M., Shy, M. E., Muntoni, F., and Pareyson, D. (2010) ‘168th ENMC International workshop: Outcome measures and clinical trials in Charcot-Marie-Tooth disease (CMT)’, *Neuromuscular Disorders*, 20(12), 839–846.

Reilly, M. M., Jonghe, P. de, and Pareyson, D. (2006) ‘136th ENMC International Workshop: Charcot-Marie-Tooth Disease Type 1A (CMT1A)8-10 April 2005, Naarden, The Netherlands’, *Neuromuscular Disorders*, 16(6), 396–402.

Reilly, M. M., Murphy, M., and Laura, M. (2011) ‘2009 PERIPHERAL NERVE SOCIETY MEETING Charcot-Marie-Tooth disease’, *Journal of the Peripheral Nervous System*, 14, 1–14.

Ricotti, V., Evans, M. R. B., Sinclair, C. D. J., Butler, J. W., Ridout, D. A., Hogrel, J. Y., Emira, A., Morrow, J. M., Reilly, M. M., Hanna, M. G., Janiczek, R. L., Matthews, P. M., Yousry, T. A., Muntoni, F., and Thornton, J. S. (2016) ‘Upper limb evaluation in duchenne muscular dystrophy: Fat-water quantification by MRI, muscle force and function define endpoints for clinical trials’, *PLoS ONE*, 11(9): e0162542.

Riggs, B. L., Hodgson, S. F., O’fallon, W. M., Chao, E. Y. s., Wahner, H. W., Muhs, J. M., Cedel, S. L., and Melon, L. J. (1990) ‘Effect of Fluoride Treatment on the Fracture Rate in Postmenopausal Women with Osteoporosis’, *New England Journal of Medicine*, 322(12), 802–809.

Ringleb, S. I., Bensamoun, S. F., Chen, Q., Manduca, A., An, K. N., and Ehman, R. L. (2007) ‘Applications of magnetic resonance elastography to healthy and pathologic skeletal muscle’, *Journal of Magnetic Resonance Imaging*, 301–309.

Rossor, A. M., Polke, J. M., Houlden, H., and Reilly, M. M. (2013) ‘Clinical implications of genetic advances in charcot-marie-tooth disease’, *Nature Reviews Neurology*, 9(10), 562–571.

Rossor, A. M., Liu, C. H., Petzold, A., Malaspina, A., Laura, M., Greensmith, L., and Reilly, M. M. (2016) ‘Plasma neurofilament heavy chain is not a useful biomarker in Charcot-Marie-Tooth disease’, *Muscle and Nerve*, 53(6), 972–975.

Sadjadi, R., Reilly, M. M., Shy, M. E., Pareyson, D., Laura, M., Murphy, S., Feely, S. M. E., Grider, T., Bacon, C., Piscosquito, G., Calabrese, D., and Burns, T. M. (2014) ‘Psychometrics evaluation of Charcot-Marie-Tooth Neuropathy Score (CMTNSv2) second version, using Rasch analysis’, *Journal of the Peripheral Nervous System*, 19(3), 192–196.

Salerno, D. F., Werner, R. A., Albers, J. W., Becker, M. P., Armstrong, T. J., and Franzblau, A. (1999) ‘Reliability of nerve conduction studies among active workers’, *Muscle and Nerve*, 22(10), 1372–1379.

Sandelius, Å., Zetterberg, H., Blennow, K., Adiutori, R., Malaspina, A., Laura, M., Reilly, M. M., and Rossor, A. M. (2018) ‘Plasma neurofilament light chain concentration in the inherited peripheral neuropathies’, *Neurology*, 90(6), e518–e524.

Sandell, S. M., Mahjneh, I., Palmio, J., Tasca, G., Ricci, E., and Udd, B. A. (2013) ‘“Pathognomonic” muscle imaging findings in DNAJB6 mutated LGMD1D’, *European Journal of Neurology*, 20(12), 1553–1559.

Saporta, M. A., Dang, V., Volfson, D., Zou, B., Xie, X. S., Adebola, A., Liem, R. K., Shy, M., and Dimos, J. T. (2015) ‘Axonal Charcot-Marie-Tooth disease patient-derived motor neurons demonstrate disease-specific phenotypes including abnormal electrophysiological properties’, *Experimental Neurology*. Elsevier Inc., 263, 190–199.

Schlaeger, S., Klupp, E., Weidlich, D., Cervantes, B., Foreman, S. C., Deschauer, M., Schoser, B., Katemann, C., Kooijman, H., Rummeny, E. J., Zimmer, C., Kirschke, J. S., and Karampinos, D. C. (2018) ‘T2-weighted dixon turbo spin echo for accelerated

simultaneous grading of whole-body skeletal muscle fat infiltration and edema in patients with neuromuscular diseases', *Journal of Computer Assisted Tomography*, 42(4), 574–579.

Schuhfried, O., Herceg, M., Pieber, K., and Paternostro-Sluga, T. (2017) 'Interrater Repeatability of Motor Nerve Conduction Velocity of the Ulnar Nerve', *American Journal of Physical Medicine and Rehabilitation*, 96(1), 45–49.

Sereda, M. W., Meyer Zu Hörste, G., Suter, U., Uzma, N., and Nave, K. A. (2003) 'Therapeutic administration of progesterone antagonist in a model of Charcot-Marie-Tooth disease (CMT-1A)', *Nature Medicine*, 9(12), 1533–1537.

Shahabpour, M., Kichouh, M., Laridon, E., Gielen, J. L., and De Mey, J. (2008) 'The effectiveness of diagnostic imaging methods for the assessment of soft tissue and articular disorders of the shoulder and elbow', *European Journal of Radiology*, 65(2), 194–200.

Shy, M. E., Lupski, J. R., Chance, P. F., Klein, C. J., and Dyck, P. J. (2005) 'Hereditary Motor and Sensory Neuropathies: An Overview of Clinical, Genetic, Electrophysiologic, and Pathologic Features', in *Peripheral Neuropathy*.

Shy, M. E., Blake, J., Krajewski, K., Fuerst, D. R., Laura, M., Hahn, A. F., Li, J., Lewis, R. A., and Reilly, M. (2005) 'Reliability and validity of the CMT neuropathy score as a measure of disability', *Neurology*, 64(7), 1209–1214.

Shy, M. E., Chen, L., Swan, E. R., Taube, R., Krajewski, K. M., Herrmann, D., Lewis, R. A., and McDermott, M. P. (2008) 'Neuropathy progression in Charcot-Marie-Tooth disease type 1A', *Neurology*, 70(5), 378–383.

Sinclair, C. D. J., Morrow, J. M., Miranda, M. A., Davagnanam, I., Cowley, P. C., Mehta, H., Hanna, M. G., Koltzenburg, M., Yousry, T. A., Reilly, M. M., and Thornton, J. S. (2012) 'Skeletal muscle MRI magnetisation transfer ratio reflects clinical severity in peripheral neuropathies', *Journal of Neurology, Neurosurgery and Psychiatry*, 83(1), 29–32.

Skre, H. (1974) 'Genetic and clinical aspects of Charcot-Marie-Tooth's disease', *Clinical Genetics*, 6(2), 98–118.

Slabaugh, M. A., Friel, N. A., Karas, V., Romeo, A. A., Verma, N. N., and Cole, B. J. (2012) 'Interobserver and intraobserver reliability of the goutallier classification using magnetic resonance imaging: Proposal of a simplified classification system to increase reliability', *American Journal of Sports Medicine*, 40(8), 1728–1734.

Sookhoo, S., MacKinnon, I., Bushby, K., Chinnery, P. F., and Birchall, D. (2007) 'MRI for the demonstration of subclinical muscle involvement in muscular dystrophy', *Clinical Radiology*, 62(2), 160–165.

Spector, S. A., Lemmer, J. T., Koffman, B. M., Fleisher, T. A., Feuerstein, I. M., Hurley, B. F., and Dalakas, M. C. (1997) 'Safety and efficacy of strength training in patients with sporadic inclusion body myositis', *Muscle and Nerve*, 20(10), 1242–1248.

Sproule, D. M., Montgomery, M. J., Punyanitya, M., Shen, W., Dashnaw, S., Montes, J., Dunaway, S., Finkel, R., Darras, B., De Vivo, D. C., and Kaufmann, P. (2011) 'Thigh muscle volume measured by magnetic resonance imaging is stable over a 6-month interval in spinal muscular atrophy', *Journal of Child Neurology*, 26(10), 1252–1259.

Stilwell, G., Kilcoyne, R. F., and Sherman, J. L. (1995) 'Patterns of muscle atrophy in the lower limbs in patients with Charcot-Marie-Tooth disease as measured by magnetic resonance imaging.', *The Journal of foot and ankle surgery : official publication of the American College of Foot and Ankle Surgeons*, 34, 583–6; discussion 596.

Stramare, R., Beltrame, V., Dal Borgo, R., Gallimberti, L., Frigo, A. C., Pegoraro, E., Angelini, C., Rubaltelli, L., and Feltrin, G. P. (2010) 'La risonanza magnetica nella valutazione delle patologie muscolari: confronto tra distrofia dei cingoli, miopatia a corpi ialini e distrofia miotonica', *Radiologia Medica*, 115(4), 585–599.

Straub, V., Carlier, P. G., and Mercuri, E. (2012) 'TREAT-NMD workshop: Pattern recognition in genetic muscle diseases using muscle MRI. 25-26 February 2011, Rome, Italy', *Neuromuscular Disorders*, 22(SUPPL. 2).

Streiner, D. L. and Norman, G. R. (2008) *Health Measurement Scales: A practical guide to their development and use*, *Health Measurement Scales: A Practical Guide to their Development and Use*.

Syed, Y. Y. (2016) 'Eteplirsen: First Global Approval', *Drugs*, 76(17), 1699–1704.

Szarek, M., White, H. D., Schwartz, G. G., Alings, M., Bhatt, D. L., Bittner, V. A., Chiang, C. E., Diaz, R., Edelberg, J. M., Goodman, S. G., Hanotin, C., Harrington, R. A., Jukema, J. W., Kimura, T., Kiss, R. G., Lecorps, G., Mahaffey, K. W., Moryusef, A., Pordy, R., *et al.* (2019) 'Alirocumab Reduces Total Nonfatal Cardiovascular and Fatal Events: The ODYSSEY OUTCOMES Trial', *Journal of the American College of Cardiology*, 73(4), 387–396.

Tasca, G., Iannaccone, E., Monforte, M., Masciullo, M., Bianco, F., Laschena, F., Ottaviani, P., Pelliccioni, M., Pane, M., Mercuri, E., and Ricci, E. (2012) 'Muscle MRI in Becker muscular dystrophy', *Neuromuscular Disorders*, 22(SUPPL. 2).

Timmerman, V., Nelis, E., Van Hul, W., Nieuwenhuijsen, B. W., Chen, K. L., Wang, S., Othman, K. Ben, Cullen, B., Leach, R. J., Hanemann, C. O., De Jonghe, P., Raeymaekers, P., van Ommen, G. J. B., Martin, J. J., Müller, H. W., Vance, J. M., Fischbeck, K. H., and Van Broeckhoven, C. (1992) 'The peripheral myelin protein gene PMP–22 is contained within the Charcot–Marie–Tooth disease type 1A duplication', *Nature Genetics*, 1(3), 171–175.

Tomas, X., Milisenda, J. C., Garcia-Diez, A. I., Prieto-Gonzalez, S., Faruch, M., Pomes, J., and Grau-Junyent, J. M. (2019) 'Whole-body MRI and pathological findings in adult patients with myopathies', *Skeletal Radiology*, 653–676.

Tomasová Studýnková, J., Charvát, F., Jarošová, K., and Vencovský, J. (2007) 'The role of MRI in the assessment of polymyositis and dermatomyositis', *Rheumatology*, 46(7), 1174–1179.

Torriani, M., Townsend, E., Thomas, B. J., Bredella, M. A., Ghomi, R. H., and Tseng, B. S. (2012) 'Lower leg muscle involvement in Duchenne muscular dystrophy: An MR imaging and spectroscopy study', *Skeletal Radiology*, 41(4), 437–445.

Verhamme, C., de Haan, R. J., Vermeulen, M., Baas, F., de Visser, M., and van Schaik, I. N. (2009) 'Oral high dose ascorbic acid treatment for one year in young CMT1A patients: A randomised, double-blind, placebo-controlled phase II trial', *BMC Medicine*, 7:70.

Verhamme, C., Van Schaik, I. N., Koelman, J. H. T. M., De Haan, R. J., and De Visser, M. (2009) 'The natural history of Charcot-Marie-Tooth type 1A in adults: A 5-year follow-up study', *Brain*, 132(12), 3252–3262.

Walsh, M., Bell, K. M., Chong, B., Creed, E., Brett, G. R., Pope, K., Thorne, N. P., Sadedin, S., Georgeson, P., Phelan, D. G., Day, T., Taylor, J. A., Sexton, A., Lockhart, P. J., Kiers, L., Fahey, M., Macciocca, I., Gaff, C. L., Oshlack, A., *et al.* (2017) 'Diagnostic and cost utility of whole exome sequencing in peripheral neuropathy', *Annals of Clinical and Translational Neurology*, 4(5), 318–325.

Wang, H., Davison, M., Wang, K., Xia, T. H., Kramer, M., Call, K., Luo, J., Wu, X., Zuccarino, R., Bacon, C., Bai, Y., Moran, J. J., Gutmann, L., Feely, S. M. E., Grider, T., Rossor, A. M., Reilly, M. M., Svaren, J., and Shy, M. E. (2020) 'Transmembrane protease serine 5: a novel Schwann cell plasma marker for CMT1A', *Annals of Clinical and Translational Neurology*, 7(1), 69–82.

Wattjes, M. P., Kley, R. A., and Fischer, D. (2010) 'Neuromuscular imaging in inherited muscle diseases', *European Radiology*, 20(10), 2447–2460.

Willcocks, R. J., Arpan, I. A., Forbes, S. C., Lott, D. J., Senesac, C. R., Senesac, E.,

Deol, J., Triplett, W. T., Baligand, C., Daniels, M. J., Sweeney, H. L., Walter, G. A., and Vandenborne, K. (2014) 'Longitudinal measurements of MRI-T2 in boys with Duchenne muscular dystrophy: Effects of age and disease progression', *Neuromuscular Disorders*, 24(5), 393–401.

Willcocks, R. J., Rooney, W. D., Triplett, W. T., Forbes, S. C., Lott, D. J., Senesac, C. R., Daniels, M. J., Wang, D. J., Harrington, A. T., Tennekoon, G. I., Russman, B. S., Finanger, E. L., Byrne, B. J., Finkel, R. S., Walter, G. A., Sweeney, H. L., and Vandenborne, K. (2016) 'Multicenter prospective longitudinal study of magnetic resonance biomarkers in a large duchenne muscular dystrophy cohort', *Annals of Neurology*, 79(4), 535–547.

Willis, T. A., Hollingsworth, K. G., Coombs, A., Sveen, M. L., Andersen, S., Stojkovic, T., Eagle, M., Mayhew, A., de Sousa, P. L., Dewar, L., Morrow, J. M., Sinclair, C. D. J., Thornton, J. S., Bushby, K., Lochmüller, H., Hanna, M. G., Hogrel, J. Y., Carlier, P. G., Vissing, J., *et al.* (2013) 'Quantitative Muscle MRI as an Assessment Tool for Monitoring Disease Progression in LGMD2I: A Multicentre Longitudinal Study', *PLoS ONE*, 8(8):e70993.

Willis, T. A., Hollingsworth, K. G., Coombs, A., Sveen, M. L., Andersen, S., Stojkovic, T., Eagle, M., Mayhew, A., De Sousa, P. L., Dewar, L., Morrow, J. M., Sinclair, C. D. J., Thornton, J. S., Bushby, K., Lochmuller, H., Hanna, M. G., Hogrel, J. Y., Carlier, P. G., Vissing, J., *et al.* (2014) 'Quantitative magnetic resonance imaging in limb-girdle muscular dystrophy 2i: A multinational cross-sectional study', *PLoS ONE*, 9(2): e90377.

Wokke, B. H., Bos, C., Reijnierse, M., Van Rijswijk, C. S., Eggers, H., Webb, A., Verschuuren, J. J., and Kan, H. E. (2013) 'Comparison of dixon and T1-weighted MR methods to assess the degree of fat infiltration in duchenne muscular dystrophy patients', *Journal of Magnetic Resonance Imaging*, 38(3), 619–624.

Wokke, B. H., van den Bergen, J. C., Versluis, M. J., Niks, E. H., Milles, J., Webb, A. G., van Zwet, E. W., Aartsma-Rus, A., Verschuuren, J. J., and Kan, H. E. (2014) 'Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy', *Neuromuscular Disorders*, 24(5), 409–416.

Wren, T. A. L., Bluml, S., Tseng-Ong, L., and 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. American journal of roentgenology*, 190(1), W8–W12