

Combined qualitative and quantitative MRI reveals short-term progression in facioscapulohumeral muscular dystrophy

Journal of Neuromuscular Diseases

2026, Vol. 0(0) 1–16

© The Author(s) 2026

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/22143602261491510

journals.sagepub.com/home/jnd  IOS Press

Anna-Lena Mayer¹, Anika Starke², Valentina Schunk², Matthias Türk^{3,4}, Michael Uder², Arnd Dörfler¹, Armin M. Nagel^{2,5}, Pierre-Yves Baudin⁶, Elisabetta Gazzero⁷, and Teresa Gerhalter^{2,8} 

Abstract

Facioscapulohumeral muscular dystrophy (FSHD) is characterized by slowly progressive and heterogeneous muscle involvement; therefore, clinical measures show limited sensitivity to short-term change. Magnetic resonance imaging (MRI) provides useful biomarkers of muscle degeneration and disease activity. To investigate one-year longitudinal changes in lower leg muscles of patients with FSHD using qualitative (T1-weighted [T1w], STIR) and quantitative MRI (fat fraction [FF], water T2), and to assess their relationship with clinical function. Thirty-one patients with FSHD (21 men; age 44.2±17 years) and 30 matched controls underwent 3T MRI. Twenty-seven patients completed one-year follow-up. Semi-quantitative grading and quantitative measures were analyzed muscle-wise. Longitudinal changes and correlations with clinical scores were assessed. At baseline, T1w imaging showed marked inter- and intra-individual variability, with heterogeneous proximal–distal involvement in 22% of muscles and asymmetric involvement in 77% of patients. STIR hyperintensity was present in 25 patients. Over one-year, clinical decline was small (MRC: $\Delta = -0.26$, $p = 0.018$, SRM = -0.23; mobility: $\Delta = -0.41$, $p = 0.031$, SRM = -0.44). In contrast, quantitative MRI detected progression, with increases in FF (0.17±0.13 to 0.22±0.15, $p < 0.001$, SRM = 0.87), FF heterogeneity, and water T2 heterogeneity, while T1w scores and mean water T2 showed limited change. Progression was greatest in muscles with intermediate baseline FF (Δ FF/year: 0.049 vs 0.007, $p < 0.001$). Baseline STIR and water T2 abnormalities predicted increases in FF but not T1w scores. FF correlated more strongly with clinical measures than T1w grading. Quantitative MRI is more sensitive than qualitative grading for detecting short-term disease progression in FSHD. Measures of heterogeneity and disease activity predict structural progression and represent promising biomarkers for clinical trials.

Plain language summary

Facioscapulohumeral muscular dystrophy (FSHD) is a genetic muscle disease that causes gradual muscle weakness. It is one of the most common muscle dystrophies, affecting about 1 in 15,000 to 20,000 individuals in Europe. The severity and pattern of muscle involvement can vary widely between individuals, and even between different muscles in the same person.

¹Department of Neuroradiology, University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Erlangen, Germany

²Institute of Radiology, Universitätsklinikum Erlangen, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Erlangen, Germany

³Department of Neurology, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

⁴Centre for Rare Diseases Erlangen (ZSEER), University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

⁵Division of Medical Physics in Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁶Institute of Myology, Neuromuscular Investigation Center, NMR Laboratory, Paris, France

⁷University Outpatient Clinic for Muscle Disorders, ECRC, Charité and Max-Delbrück-Center for Molecular Medicine, Berlin, Germany

⁸Department of Neurology, Medical University of Graz, Graz, Austria

Corresponding author:

Teresa Gerhalter, Department of Neurology, Medical University of Graz, Auenbruggerplatz 22, 8036 Graz, Austria.

Email: teresa.gerhalter@medunigraz.at



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Detecting disease progression over short periods, such as one year, is challenging using standard clinical tests, which may not capture subtle changes. This makes it difficult to evaluate new treatments in clinical trials. In this study, we used magnetic resonance imaging (MRI) to examine muscle changes in the lower legs of 31 patients with FSHD. MRI can provide detailed images of muscle structure and detect both long-term damage (fatty replacement) and signs of active disease, such as inflammation or fluid-related (edematous) changes. Twenty-seven patients were re-examined after one year. We found that MRI could detect disease progression even when clinical muscle strength and mobility showed only small or no changes. In particular, quantitative MRI measurements showed a clear increase in fat fraction within muscles over time. Muscles that were moderately affected at the start of the study were most likely to worsen. We also observed that some muscles showed specific MRI signal changes (called STIR hyperintensities), which are thought to reflect active disease processes. In several cases, these changes appeared before muscles developed more fat replacement, suggesting they may help identify muscles at risk of future damage. Overall, our results show that MRI is a sensitive and objective tool for monitoring FSHD progression. It can detect early and subtle changes that are not captured by standard clinical tests. These findings are important for future clinical trials, as MRI could help evaluate the effectiveness of new treatments more accurately and within shorter time periods.

Keywords

FSHD, longitudinal study, T1w MRI, STIR, fat fraction, T2 mapping, MRC score

Received: 8 May 2026; Revised: 1 September 2026; Accepted: 7 September 2026

Introduction

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most common inherited muscular dystrophies, with a prevalence of approximately 1:8,000 to 1:20,000.¹ It is caused by epigenetic derepression of the DUX4 gene on chromosome 4q35, most commonly due to contraction of the D4Z4 repeat array (FSHD1) or mutations in chromatin modifiers such as *SMCHD1*, resulting in hypomethylation of the D4Z4 repeat.² Aberrant DUX4 expression induces myofiber damage through mechanisms including impaired myogenic differentiation, oxidative stress, and inflammatory signaling, ultimately leading to progressive muscle degeneration.³

Clinically, FSHD is characterized by slowly progressive and highly heterogeneous muscle involvement, with marked variability between patients and across individual muscles. Disease progression is often asymmetric and asynchronous, and functional decline may remain subtle over extended periods.⁴ Traditional clinical outcome measures, including manual muscle testing and timed functional tests, are influenced by motivation, pain, fatigue, and learning effects, limiting their sensitivity to short-term change. Consequently, there is a growing need for objective biomarkers that capture subclinical disease activity.⁵

Magnetic resonance imaging (MRI) provides a non-invasive means to assess structural and pathophysiological muscle changes.^{6,7} Conventional qualitative MRI using T1-weighted (T1w) and short-tau inversion recovery (STIR) sequences is widely used in clinical practice. T1w imaging reflects chronic muscle damage through fatty replacement and atrophy, whereas STIR hyperintensity is commonly interpreted as muscle edema, inflammation, or active myofiber damage.^{8,9} However, semi-quantitative grading systems are limited by inter-rater variability, ceiling effects, and reduced sensitivity to subtle or early changes.

Quantitative MRI (qMRI) techniques, such as chemical-shift encoded (CSE) fat fraction (FF) imaging and water T2 mapping, enable a reproducible assessment of muscle composition. These approaches improve sensitivity to early and longitudinal changes and allow detection of progression even when functional measures remain stable.^{10–13} Muscles with intermediate baseline FF appear particularly prone to progression, and composite MRI measures derived from multiple muscles showed increases over relatively short time intervals of 1–5 years.^{11–14}

In addition, STIR hyperintensity and elevated water T2 have been associated with muscles at risk of subsequent fatty degeneration, suggesting that they may reflect transitional disease states.^{10,15–17} Longitudinal studies indicate that such inflammatory-like changes frequently occur in mildly or intermediately affected muscles and may either resolve or evolve into irreversible fat replacement.^{16–18} Furthermore, analyses along the full muscle length have demonstrated spatially heterogeneous, wave-like patterns of disease progression that are not captured by single-slice assessments.^{10,17–19}

Despite these advances, short-term longitudinal studies integrating qualitative and quantitative MRI within the same cohort remain limited. The practical implementation of qMRI in clinical routine is challenged by technical complexity and the need for standardized analysis pipelines with automated segmentation and normative thresholds.^{20,21} Thus, the optimal integration of qualitative and quantitative MRI into routine practice and clinical trials remains an open question.

Therefore, this study aimed to characterize one-year longitudinal changes in leg muscles of patients with FSHD using combined qualitative (T1w and STIR grading) and quantitative MRI techniques (DIXON-derived fat fraction and water T2 mapping). Specifically, we investigated (i) the sensitivity of quantitative and qualitative MR to detect fatty degeneration over one year, (ii) the relationship between STIR/water T2 abnormalities and subsequent fatty degeneration, and (iii) the association between imaging findings and clinical function.

Materials & Methods

Study design and Participants

This prospective longitudinal study was reviewed and approved by the Institutional Review Board of Universitätsklinikum Erlangen (approval no. 291_20 B; amendment 20-291_4-B). The study was conducted in accordance with the Declaration of Helsinki and its later amendments. Written informed consent was obtained from all participants before inclusion; for minors, consent was obtained from a legal guardian and assent from the participant.

Patients were individuals aged 14 - 80 years with genetically confirmed FSHD. Inclusion required the ability to walk 20 metres, ensuring sufficient residual leg musculature was present for the MRI measurements. Technical aids such as canes, crutches and light knee or ankle orthoses were permitted. Exclusion criteria comprised contraindications to MRI (e.g. pacemaker, aneurysm clip, or claustrophobia). Age- and sex-matched healthy participants were recruited as controls via public advertisements. Inclusion criteria for healthy controls were the absence of a neuromuscular disorder and of diseases associated with muscle damage, and no contraindications to MRI.

The study was conducted between August 2023 and October 2025 at Universitätsklinikum Erlangen, comprising two time points (baseline and follow-up). At baseline, all participants underwent a review of their medical history, completed a questionnaire on disease progression and everyday mobility, underwent MRI of the lower legs, and had their muscle strength assessed. Patients with FSHD were invited for a follow-up visit one year later.

Clinical assessment

Clinically relevant data, including self-reported age at disease onset, total disease duration, duration of lower leg muscle weakness, and current medication, were collected through personal interviews and review of available medical records.

Lower limb function was assessed using a study-specific mobility questionnaire developed for this investigation. The questionnaire was designed as a pragmatic descriptive measure of overall mobility based on previously reported disease burden,²² and included independent walking, stair climbing, rising from floor or chair, as well as ability to walk on tiptoes and the presence of steppage gait and Gowers' sign. Each item was scored as follows: for difficulties walking on tiptoes, rising from the floor, and for the presence of steppage gait or Gowers' sign, 0 points were assigned; otherwise, 1 point was given. For the ability to rise from a chair, 2 points were assigned for "yes," 1 point for "with help," and 0 points for "not possible." For the ability to walk freely, 3 points were assigned for "yes," 2 points for "only a few meters," 1 point for "with rollator," and 0 points for "no." For stair climbing, 3 points were assigned for "yes," 2 points for "with banister," 1 point for "with help," and 0 points for "no." The resulting composite mobility score ranged from 0 to 12, with higher scores indicating better mobility and a score of 12 indicating no mobility impairment. This questionnaire was used to provide an overall assessment of mobility, including but not limited to lower-limb function, and was assessed as a complementary clinical measure to the MRI findings.

Muscle strength was assessed using the Medical Research Council (MRC) scale, ranging from 0 (no contraction) to 5 (normal strength). Muscle strength of the upper and lower extremities was assessed bilaterally by a trained examiner. Composite MRC scores for each lower limb (maximum 30 points per side) were calculated by summing the scores for knee flexion and extension, foot flexion and extension, and foot inversion and eversion.²³ Upper extremity strength and hip flexion and extension were assessed as part of the standardized neurological examination but were not included in the composite MRC score reported in this study. The same assessment protocol was applied at baseline and follow-up, and changes in the lower-limb composite MRC scores are reported.

MRI acquisitions

Measurements were performed on a 3T whole-body MR scanner (MAGNETOM Vida, Siemens Healthineers, Erlangen, Germany). Participants were positioned supine, feet-first. An 18-channel flexible surface coil was placed around the lower legs in combination with the built-in 32-channel spine coil (Siemens Healthineers).

Based on localizer images in three orientations, axial T1-weighted and fat-suppressed T2-weighted sequences were acquired on both legs using an identical imaging protocol for patients and controls. T1-weighted imaging consisted of an axial acquisition with the following parameters: field-of-view (FOV) = 400 x 242 mm², 3D stack with 90 slices, slice thickness = 3 mm, distance factor = 0%, acceleration factor 2, echo time (TE) = 9.9 ms, repetition time (TR) = 588 ms. Axial short-Tau-Inversion-Recovery T2-weighted (STIR-T2w) sequence was performed with the following parameters: FOV = 400 x 239 mm², 68 slices, slice thickness = 3 mm, distance factor = 0%, acceleration factor 2, TE = 75 ms, inversion time = 220 ms, TR = 5490 ms.

For quantitative MRI, one lower leg was selected for each participant. When participants reported a clear difference in lower limb involvement, the less affected leg was selected based on subjective assessment of side-to-side differences in symptoms and function. If no clear asymmetry was reported, the right leg was examined as a predefined default side.

Fat fraction (FF) was measured using a 3D six-point DIXON sequence with the following parameters: TR 21 ms, TEs 2.22/5.42/8.62/11.82/15.02/18.22 ms, flip angle = 6°, 50 slices with a 0.9x0.9x5 mm³ spatial resolution, acquisition time = 1 min 57 s. Data were processed using an in-house non-linear least-squares fitting method with T2* estimation and fat spectrum modeling.²⁴

Water T2 mapping was performed using a multi-slice-multi-echo (MSME) sequence with the following parameters: TR 3000 ms, 32 echoes with TEs from 9.5 ms to 304, 7 slices with a 1.3x1.3 mm² resolution, slice thickness = 10 mm, acquisition time = 3 min 29 s. Water T2 values were calculated using a tri-exponential fitting approach as previously described.²⁵

For longitudinal examinations, the imaging volume was positioned using an anatomical landmark-based approach. The largest calf circumference was selected as the center of the acquisition volume. The distance from the inferior border of the patella to this level was recorded at baseline and reproduced at follow-up using external measurements and skin markings before scanning.

MRI Data Evaluation

Semi-quantitative Assessment

Semi-quantitative image analysis was performed by a radiologist with six years of experience in musculoskeletal MRI. Seven lower leg muscles were evaluated bilaterally: tibialis anterior, extensor hallucis/digitorum longus, peroneal muscles, tibialis posterior, gastrocnemius medialis, gastrocnemius lateralis, and soleus.

Fatty infiltration was graded on axial T1-weighted images using the five-point Goutallier classification (0 = no fatty replacement; 4 = more fat than muscle tissue).²⁶ Muscle edema was assessed on axial STIR T2-weighted images using a three-point ordinal scale (0 = normal signal intensity; 0.5 = mild hyperintensity; 1 = marked hyperintensity). Grading was primarily based on signal intensity rather than spatial extent.

All available axial slices covering each muscle were reviewed (T1w TSE axial 90 slices, T2w TSE STIR axial 68 slices in total). Fatty infiltration and edema were assessed systematically across the entire muscle length. In cases of heterogeneous involvement, a differentiated evaluation was performed by assigning separate grades to proximal and distal muscle portions (origin and insertion, each 25 slices). The mean score across these segments was subsequently calculated and used for statistical analysis and comparison with qMRI and clinical scores. Additionally, the highest observed grade along the full muscle length was recorded and used as the representative score to describe the severity for visual grading. Muscle-wise values were averaged across the seven graded muscles to obtain subject-level composite wT1 and STIR scores for correlation analyses with clinical parameters.

Quantitative Fat Fraction and Water T2 Analysis

Quantitative segmentation was performed on the first TE image of the Dixon images using the Deep Anatomical Federated Network (DAFNE).²⁷ Regions of interest (ROIs) were automatically generated and manually reviewed for the following muscles: gastrocnemius medialis, gastrocnemius lateralis, soleus, peroneal muscles, extensor digitorum longus, tibialis anterior, and tibialis posterior. 23 slices per patient were evaluated on average between the distal end of the gastrocnemius lateralis muscle and the apex of the patella.

ROIs encompassed the full cross-sectional area of each muscle while excluding fascia, subcutaneous fat, and visible vascular structures. For each muscle, mean FF and water T2 values were extracted. To assess the intramuscle heterogeneity, the FF interquartile range per muscle was also extracted. For water T2 heterogeneity, we calculated the standard deviation of the ROI divided by the mean water T2. Muscle-wise values were averaged across the seven analyzed muscles to obtain subject-level composite FF and wT2 measures for correlation analyses with clinical parameters. Disease progression was assessed by baseline-year-1 changes in FF normalised by the time between examinations ($\Delta FF = (FF_{\text{visit2}} - FF_{\text{visit1}}) / \text{number of days between visits} * 365.25$ in %).

Statistical Analysis

All statistical analyses were performed in Python using SciPy and statsmodels. Data are presented as mean $\pm$ standard deviation (SD).

Due to the ordinal nature and non-normal distribution of clinical scores, longitudinal changes were assessed using Wilcoxon signed-rank tests. For consistency with MRI data, the MRC score of the clinically stronger side was used for analysis.

Parametric or non-parametric methods were applied depending on distributional assumptions of MR parameters, which were assessed using the Shapiro–Wilk test. Between-group comparisons at baseline (patients vs. controls) were performed using independent-samples t-tests or Mann–Whitney U tests as appropriate. Comparisons between FSHD subgroups (e.g., low FF < 0.13 vs. fat-replaced muscles) used unpaired tests. The threshold for low vs. fat-replaced muscles was set at 0.13, corresponding to the mean FF of all control muscles plus two times the standard deviation (mean + 2*SD).

Cross-sectional correlations between clinical scores and MRI measures at baseline (v1) were computed using Spearman rank correlation, given the ordinal nature of several clinical scales. MRI measures were averaged per patient for correlations with clinical outcomes. Muscle-wise correlations were performed for comparisons between MRI parameters (T1w score, FF, STIR, water T2, and water T2 heterogeneity) without adjustment for within-subject clustering, assuming muscle-independent progression.

Longitudinal changes were calculated as Δ qMRI per year to account for the variable time interval between baseline and follow-up scans. Longitudinal comparisons within the FSHD cohort (v1 vs. v2) were conducted using Wilcoxon signed-rank tests. Muscle-wise analyses were performed separately for each muscle. For multiple muscle comparisons, false discovery rate (FDR) correction was applied to control for multiple testing. Sensitivity to longitudinal change was assessed using the standardized response mean (SRM), calculated as the mean within-subject change between follow-up and baseline divided by the standard deviation of the within-subject change; SRMs were calculated for both semi-quantitative and quantitative MRI measures, with larger absolute values indicating greater responsiveness to change. To contextualize the magnitude of longitudinal FF changes, muscle-specific mean FF and the 5th–95th percentile range were calculated in healthy controls and compared descriptively with the annual Δ FF observed in patients with FSHD.

Associations between baseline MRI parameters and longitudinal progression were assessed using Spearman correlations on a muscle-wise basis. For multiple testing, FDR correction (Benjamini–Hochberg) was applied, with FDR-adjusted p-values reported.

Progression by baseline FF category was evaluated by stratifying muscles into low (<0.13), intermediate (0.13–0.60), and high (>0.60) baseline FF. Muscles with high FF were excluded from progression analyses due to ceiling effects. Differences in Δ FF/year and Δ T1w/year between low and intermediate baseline FF muscles were tested using the Mann–Whitney U test, and effect size was calculated as the rank-biserial correlation.

All tests were two-sided. A p-value < 0.05 was considered statistically significant after correction where applicable.

Results

Clinical Characteristics of the Study Cohort

Baseline cohort

Thirty-one patients with FSHD were included (21 men, 10 women; mean age 44.4 ± 17 years, range 14–79). Patients were all unrelated except for one mother and daughter as well as one father and son. Twenty-seven patients were able to walk independently for at least 20 meters, whereas four required assistive devices (crutches, cane, or ankle–foot orthosis). Difficulties walking on tiptoes were reported in 15 patients, and 14 exhibited a mild to advanced stepping gait. The more severely affected leg was the left in 7 patients, the right in 11, and bilateral or no clear asymmetry was observed in 13 patients.

At baseline, the lower-limb composite MRC score of the clinically stronger lower extremity was 27.5 ± 3.5 (maximum score 30) and of the weaker lower extremity 26.6 ± 4.0 , and the overall mobility score was 9.1 ± 2.8 (maximum score 12).

Thirty age- and sex-matched healthy individuals were included (20 men, 10 women; mean age 46.1 ± 17 years, range 14–81) for the control cohort.

Longitudinal cohort

Follow-up MRI and clinical assessment were completed in 27 patients after a mean interval of 54.2 ± 4.9 weeks (range 50–68). Four patients did not attend follow-up due to surgery (n = 1), lack of availability (n = 2), or loss to follow-up (n = 1).

Over one year, clinical measures showed small changes, with SRMs of -0.23 for the stronger-leg MRC score ($\Delta = -0.24 \pm 1.06$, $p = 0.23$), -0.07 for the weaker-leg MRC score ($\Delta = -0.06 \pm 0.85$, $p = 0.67$), and -0.44 for the overall mobility score ($\Delta = -0.41 \pm 0.93$, $p = 0.031$).

Average clinical characteristics are provided in [Supplemental Table 1](#), and individual patients' demographics in [Supplemental Table 2](#).

Radiological MRI Assessment of Fat Replacement and Edema

While three patients showed no fatty replacement at all, an asymmetric involvement in at least one muscle pair was found in two-thirds ($n = 24/31$) of all patients. Bilateral muscle involvement is shown in [Supplemental Figure 1](#).

Semi-quantitative T1-weighted assessment was performed in the less affected leg, resulting in the evaluation of 217 muscles. Fatty infiltration showed a broad spectrum of involvement, with marked inter- and intra-individual variability ([Figure 1A and B](#)). The gastrocnemius medialis and lateralis, soleus, and tibialis anterior were most frequently affected, whereas the tibialis posterior and deep flexor muscles showed relatively lower degrees of involvement. Fatty replacement was frequently heterogeneous within individual muscles, with differing proximal and distal involvement observed in 22% (48/217) of muscles out of 22 patients. Among these, 19 muscles showed a difference of one grade, while 28 demonstrated more pronounced heterogeneity (difference ≥ 2 grades).

Longitudinally, semi-quantitative T1w scores showed only limited progression over the one-year interval, with most muscles remaining stable. 10 patients showed mild progression of fatty infiltration in at least one of their legs. 4 of these 10 patients showed progression in the less affected leg at baseline. Importantly, progression occurred across different baseline Goutallier grades and was not limited to muscles with intermediate initial involvement. Overall, the average T1w scores did not change significantly over one year ($p = 0.18$, $\text{SRM} = 0.36$).

At baseline, STIR hyperintensity was present in 25 patients, affecting 91 muscles in total. Signal intensity and distribution varied considerably between patients and across muscles ([Figure 1C and D](#)). Mild (grade 0.5) and marked (grade 1) signal changes were most common in posterior calf muscles, including the gastrocnemius medialis ($n = 22/31$), tibialis anterior ($n = 18/31$) and gastrocnemius lateralis ($n = 16/31$). In contrast, the tibialis posterior ($n = 4/31$) and peroneal muscles ($n = 3/31$) were only rarely affected, indicating relative sparing of deep posterior and lateral compartments.

STIR abnormalities showed limited but detectable longitudinal variation. Using the ordinal STIR grading scale, 7 muscles demonstrated increased STIR scores and 5 muscles demonstrated decreased scores over one year, while 176 muscles remained unchanged. At the patient level, STIR intensity increased in at least one muscle in five patients (including one patient with simultaneous increases and decreases in different muscles), whereas 21 patients showed no longitudinal STIR change. In contrast, a subset of 9 patients demonstrated focal STIR hyperintensity at baseline in individual muscles that subsequently showed increased fatty replacement at one year (see [Figure 1](#)), suggesting that edematous changes may precede structural degeneration. Overall, the average STIR scores did not change significantly over one year ($p = 0.50$, $\text{SRM} = 0.17$). These findings were further explored using quantitative MRI measures.

Quantitative Fat Fraction and Water T2 Analysis

Of the 189 analyzed dystrophic muscles, 126 were classified as normal ($\text{FF} < 0.13$), 47 as mildly or intermediately fat-replaced ($\text{FF} 0.13 - 0.60$), and 16 as severely fat-replaced ($\text{FF} > 0.60$). Compared with controls, patients with FSHD demonstrated significantly higher mean fat fraction and greater fat fraction heterogeneity (interquartile range, IQR) across lower leg muscles (both $p < 0.001$, Mann–Whitney U test; [Figure 2](#)). In particular, the gastrocnemius medialis, gastrocnemius lateralis, soleus, tibialis anterior, and extensor digitorum longus exhibited significantly higher FF and FF IQR compared with controls.

Overall water T2 values were also increased in FSHD patients compared with controls ($p < 0.001$, Mann–Whitney U test), accompanied by increased water T2 heterogeneity ($p < 0.001$). However, muscle-wise analysis showed a significant increase in water T2 only in the soleus muscle. In contrast, water T2 heterogeneity was significantly higher in GM, SOL, EDL, and TA ([Figure 2](#)).

At follow-up, both mean FF and FF heterogeneity increased significantly overall (FF: 0.17 ± 0.13 to 0.22 ± 0.15 , $p = 0.0014$, $\text{SRM} = 0.87$; FF IQR: 0.11 ± 0.08 to 0.14 ± 0.09 , $p = 0.0014$, $\text{SRM} = 0.78$; both Wilcoxon signed-rank tests). Muscles that already exhibited elevated FF or FF IQR at baseline showed further increases over 1 year ([Figure 2](#)). Compared with the normative FF distribution in healthy controls, the mean annual increase in FF was substantial, particularly in GM (38.4%), GL (57.8%), EDL (34.2%), and TA (91.7% of the healthy-control mean FF; [Supplementary Table 4](#)). In contrast, no significant longitudinal changes were observed in overall water T2 ($p = 0.327$, $\text{SRM} = 0.08$) or at the individual muscle level.

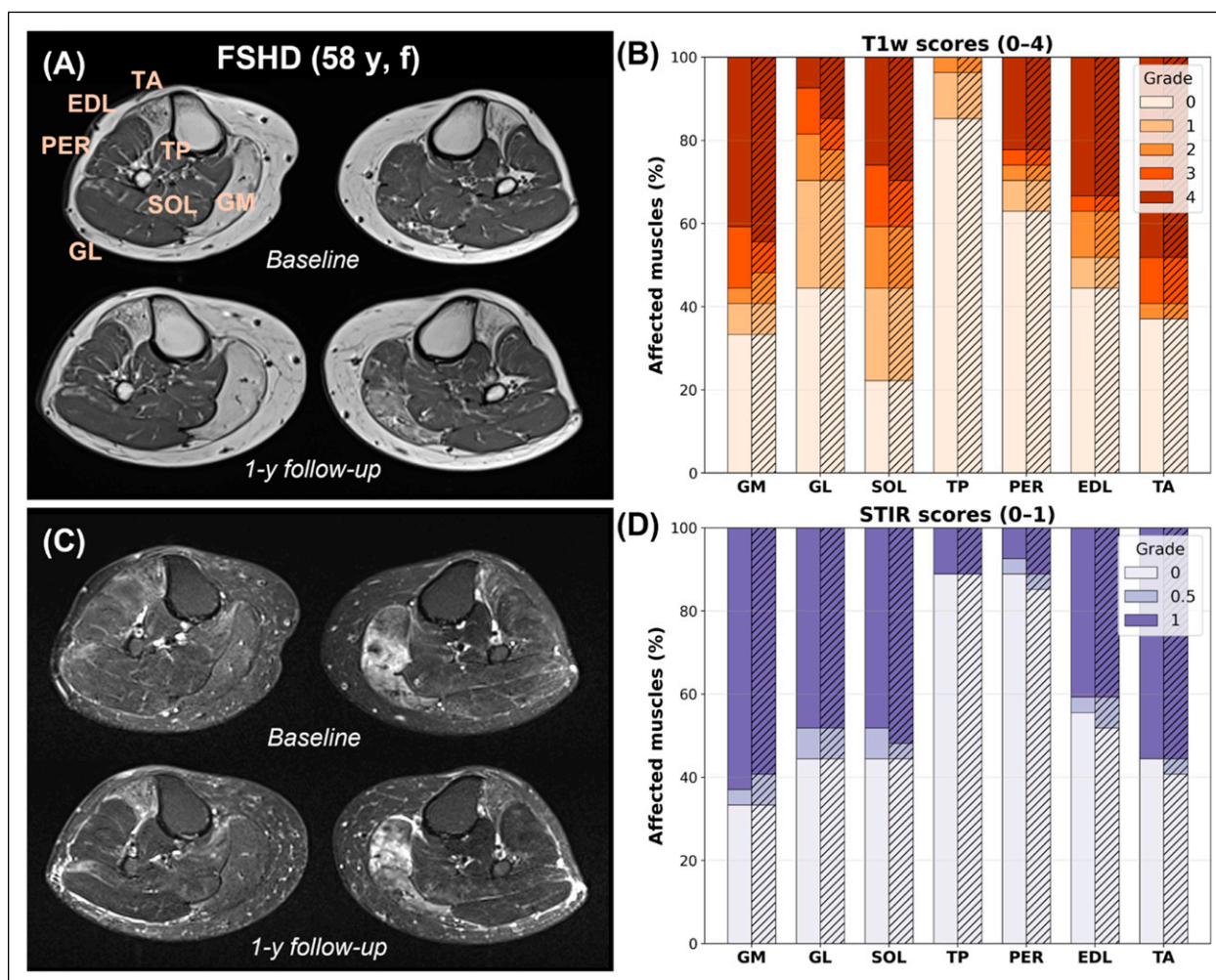


Figure 1. Qualitative MRI assessment of lower leg muscles in FSHD. **A.** Representative axial T1-weighted (T1w) images of a 58-year-old female patient with FSHD at baseline and 1-year follow-up. At baseline, low fatty replacement is present in the left GM (grade 2), the right GM shows strong fatty degeneration (grade 4). The left GL shows increased T1w signal at 1-y follow-up (grade 3). **B.** Distribution of maximum T1w grades (0–4) along the muscles in the FSHD cohort at baseline and follow-up (indicated by diagonal stripes). T1w grading was performed for different leg muscles using a semi-quantitative scale (0 = normal appearance; 1 = minimal fatty streaks; 2 = moderate fatty infiltration; 3 = severe fatty infiltration with residual muscle tissue; 4 = complete fatty replacement). The clinically less affected side was selected for analysis. **C.** Representative axial STIR images of the same patient at baseline and follow-up, demonstrating hyperintense signal alterations in the left GM and the right TA muscle (grade 1). **D.** Distribution of maximum STIR grades (0, 0.5, 1) along the muscles in the FSHD cohort at baseline and follow-up (indicated by diagonal stripes). STIR grading was performed for different leg muscles using a semi-quantitative three-level scale (0 = absent, 0.5 = focal or patchy hyperintensity, 1 = diffuse or extensive hyperintensity). As for T1w imaging, the clinically less affected side was selected for analysis. Abbreviations: GM, gastrocnemius medialis; GL, gastrocnemius lateralis; SOL, soleus; TA, tibialis anterior; TP, tibialis posterior; PER, peroneus longus; EDL, extensor digitorum longus.

However, overall water T2 heterogeneity increased (0.16 ± 0.04 to 0.17 ± 0.04 , $p = 0.025$, SRM = 0.12), with a significant increase also observed in the EDL muscle.

Muscles with FF > 0.13 generally exhibited higher water T2 values as well as greater water T2 heterogeneity compared with muscles with low FF (Figure 3). Only one muscle with low FF showed increased water T2 heterogeneity compared with controls. Muscles with FF > 0.13 were also more likely to demonstrate STIR hyperintensities.

Descriptive statistics for all MR-derived measures are provided in the [Supplemental Table 3](#).

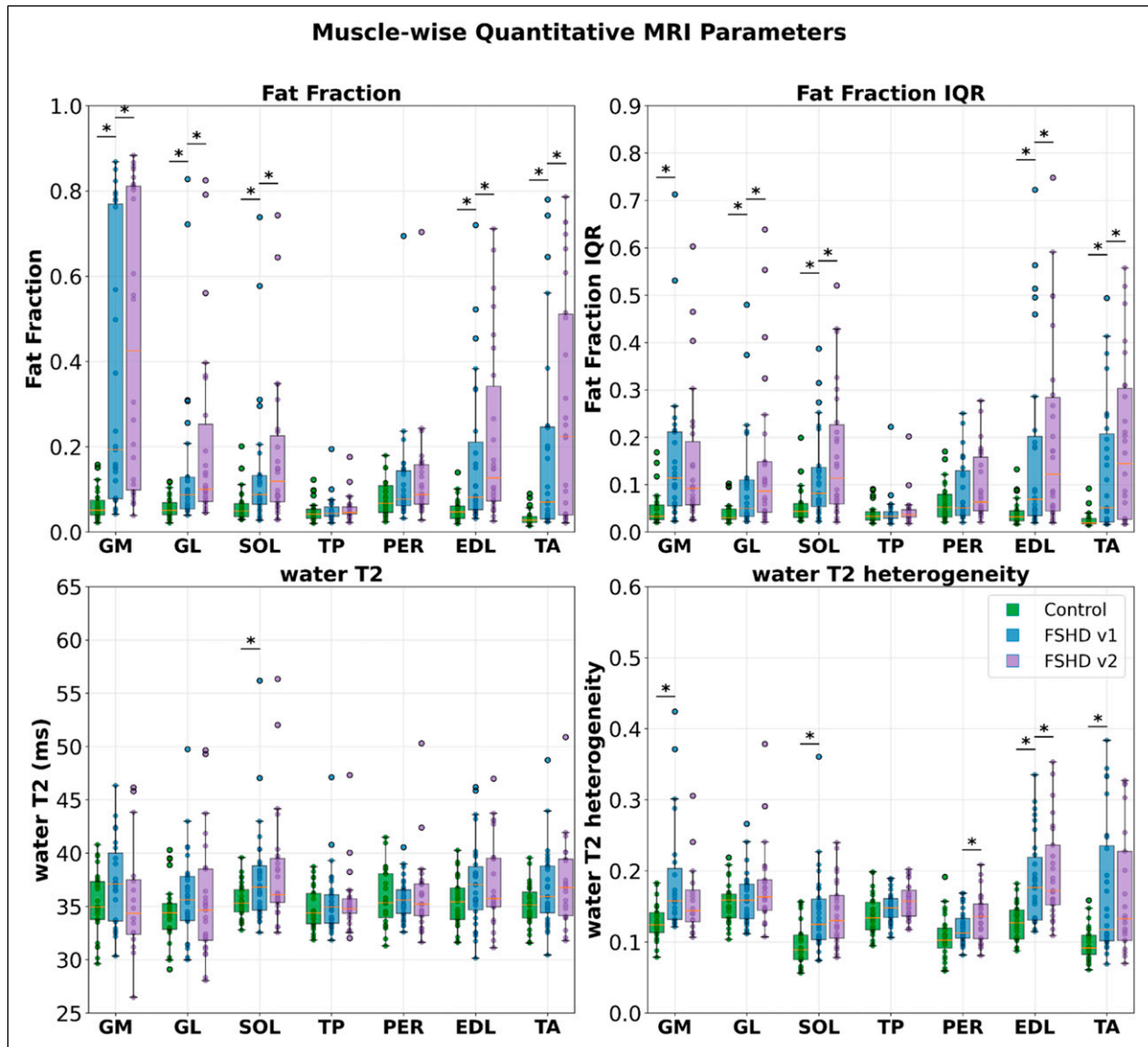


Figure 2. Quantitative MRI parameters of individual muscles, including fat fraction (FF), interquartile range (IQR) of FF as a measure of FF heterogeneity, water T2, and water T2 heterogeneity (expressed as standard deviation divided by mean), shown for healthy controls, FSHD patients at baseline (v1), and FSHD patients at follow-up (v2). Asterisks (*) indicate significant differences between controls and FSHD at v1 or between FSHD at v1 and v2. Muscles with FF > 0.6 (n = 16) were excluded from water T2 analyses due to unreliable curve fitting. Abbreviations: GM, gastrocnemius medialis; GL, gastrocnemius lateralis; SOL, soleus; TA, tibialis anterior; TP, tibialis posterior; PER, peroneus longus; EDL, extensor digitorum longus.

Association between MRI Findings and Clinical Measures

Figure 4 presents line plots comparing mean qualitative and quantitative MRI measures for baseline and follow-up. Quantitative fat fraction demonstrated a strong correlation with semi-quantitative T1-weighted grading consistent across visits (Spearman $\rho = 0.85$, $p < 0.001$). In contrast, water T2 showed only a moderate association with STIR-based scores ($\rho = 0.58$, $p < 0.001$).

Both quantitative FF and semi-quantitative T1w grading showed significant correlations with overall mobility score, lower-limb composite MRC score, and foot flexion and extension strength according to MRC score (Table 1(i)). Twelve of 17 patients with pronounced fatty replacement retained independent ambulation, and 8 of these patients also demonstrated relatively preserved muscle strength (composite MRC score > 25).

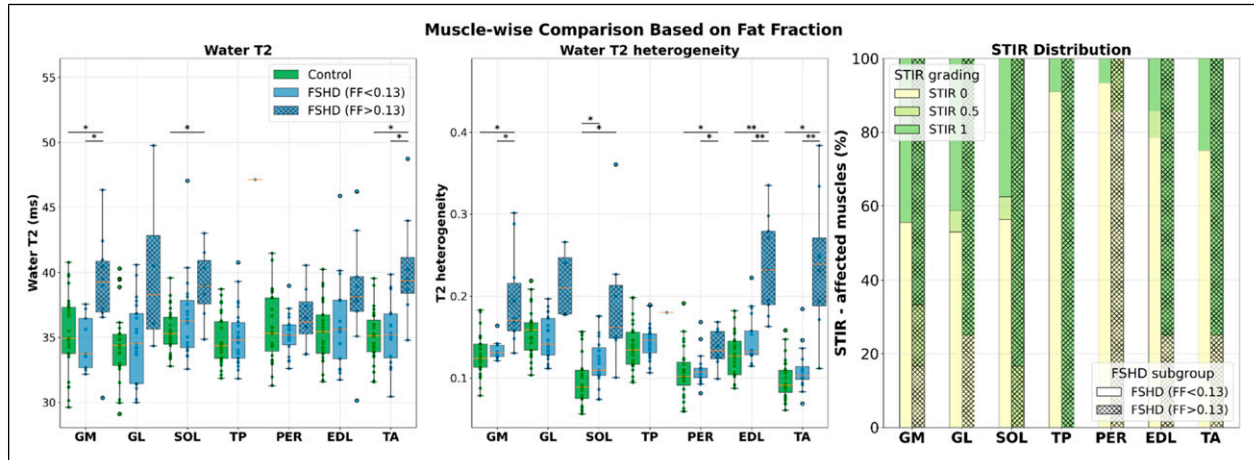


Figure 3. Early imaging markers beyond fat fraction. Muscles of FSHD patients were stratified according to fat fraction (FF) into low (FF < 0.13) and intermediate fat-replaced (0.13 < FF < 0.6) groups and compared with healthy controls. Left: Water T2 values. Middle: Water T2 heterogeneity. Right: Frequency of STIR hyperintensities (maximum score across the muscles). In gastrocnemius medialis (GM), soleus (SOL), and tibialis anterior (TA), water T2 values were significantly higher in fat-replaced muscles compared with controls and FSHD muscles with low FF, as indicated by asterisks (*, FDR-corrected $p < 0.05$; **, FDR-corrected $p < 0.001$). These muscles also showed increased water T2 heterogeneity in addition to PER and EDL. These muscles, as well as peroneus (PER) and extensor digitorum longus (EDL), also exhibited increased water T2 heterogeneity. Furthermore, SOL muscles with low FF showed increased water T2 heterogeneity compared with controls. Similarly, STIR hyperintensities were more frequently observed in muscles with higher FF compared to those with low FF. Abbreviations: GM, gastrocnemius medialis; GL, gastrocnemius lateralis; SOL, soleus; TA, tibialis anterior; TP, tibialis posterior; PER, peroneus longus; EDL, extensor digitorum longus.

Baseline lower-limb composite MRC and mobility scores showed only weak to moderate correlations with subsequent progression of fatty replacement, assessed either by semi-quantitative Goutallier grading or quantitative FF analysis (Table 1(ii)), with stronger associations observed for FF compared with T1w scores. Furthermore, changes in FF correlated moderately with baseline FF, water T2 values, and STIR grading ($\rho = 0.49$, $\rho = 0.45$, and $\rho = 0.45$, respectively; all $p <$

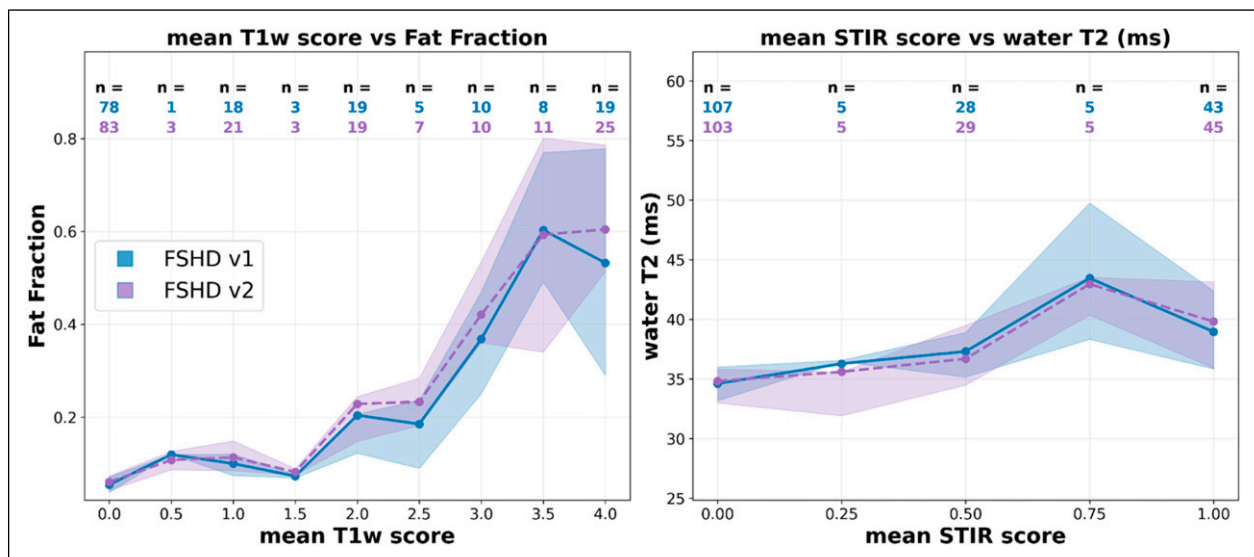


Figure 4. Comparison of qualitative and quantitative MRI metrics in lower leg muscles of FSHD patients at baseline (v1) and follow-up (v2). Left: Fat fraction (FF) stratified by mean T1-weighted (T1w) scores. Right: Water T2 stratified by mean STIR scores. Symbols represent group-wise mean values for each qualitative MRI category and visit. Lines connect mean values across categories for visualization of trends only and do not represent fitted regression models or smoothing functions. Shaded areas indicate the interquartile range (25th–75th percentile) of individual muscle values within each category and visit. The number of muscles contributing to each category is indicated above the x-axis.

Table 1. Correlation analysis in patients with FSHD at baseline (v1, left columns) and for longitudinal changes between baseline and follow-up (right columns). (i) Baseline analyses include associations between clinical scores and averaged MRI-derived measures of fat replacement (upper panel), as well as muscle-wise associations between fat fraction (FF) and water T2 or STIR grading (lower panel). (ii) Longitudinal analyses evaluated associations between baseline (v1) clinical scores and longitudinal changes in MRI-derived measures of fat replacement, expressed as (v2 – v1) normalized to the time interval between visits (per year). In addition, correlations were evaluated between changes in fat replacement and baseline MRI parameters, including fat fraction (FF), T1-weighted (T1w) scores, water T2, and STIR grading.

(i) Fat MRI measure at v1				(ii) Fat MRI measure change (v2 - v1) / time	
Upper panel: Clinical vs MRI (subject-level)					
MRI measure (average muscle values)	Functional test at v1	Spearman rho	P value (FDR corrected)	Spearman rho	P value (FDR corrected)
FF	Sum mobility	-0.77	<0.001	-0.59	0.051
	MRC score	-0.77	<0.001	-0.47	0.080
	Foot flexion	-0.67	<0.001	-0.36	0.157
	Foot extension	-0.70	<0.001	-0.30	0.216
	Foot inversion	-0.39	0.043	-0.30	0.216
	Foot elevation	-0.40	0.043	-0.48	0.080
T1w score	Sum mobility	-0.76	<0.001	-0.38	0.085
	MRC score	-0.74	<0.001	-0.43	0.080
	Foot flexion	-0.58	<0.001	-0.12	0.538
	Foot extension	-0.67	0.002	-0.20	0.334
	Foot inversion	-0.40	0.043	-0.38	0.085
	Foot elevation	-0.40	0.043	-0.41	0.083
Lower panel: MRI vs MRI (muscle-level)					
MRI measure (muscle-wise)	MRI measure at v1	Spearman rho	P value (FDR corrected)<	Spearman rho	P value (FDR corrected)
FF	FF	n.a.		0.49	<0.001
	STIR	0.52	<0.001	0.45	<0.001
	Water T2	0.57	<0.001	0.45	<0.001
T1w score	T1w score	n.a.		0.07	0.383
	STIR	0.61	<0.001	0.06	0.383
	Water T2	0.54	<0.001	0.09	0.266

For correlations involving functional measures, MRI parameters were averaged across the assessed lower-leg muscles at the subject level, whereas MRI-to-MRI correlations were performed on a muscle-wise basis.

0.001). Stratified analyses indicate that T2 mean is most predictive in muscles with low baseline FF, whereas T2 heterogeneity shows the strongest association in muscles with intermediate FF, suggesting complementary roles of these metrics across disease stages (Figure 5).

Muscles with intermediate baseline fat fraction (FF 0.13–0.60) showed significantly greater progression compared with muscles with low baseline FF (<0.13) ($\Delta\text{FF}/\text{year}$: 0.049 ± 0.066 vs 0.007 ± 0.014 , $p < 0.001$), indicating that disease progression is most pronounced in moderately affected muscles (Figure 5).

Discussion

In this longitudinal MRI study, we systematically characterized structural muscle degeneration and markers of disease activity in the lower legs of patients with FSHD over a one-year interval. By directly comparing qualitative and quantitative MRI approaches within the same cohort, we investigated (i) the sensitivity of qualitative and quantitative MRI to detect fatty degeneration over one year, (ii) the relationship between STIR/water T2 abnormalities and subsequent degeneration, and (iii) the association between imaging findings and clinical function. Three main findings emerge.

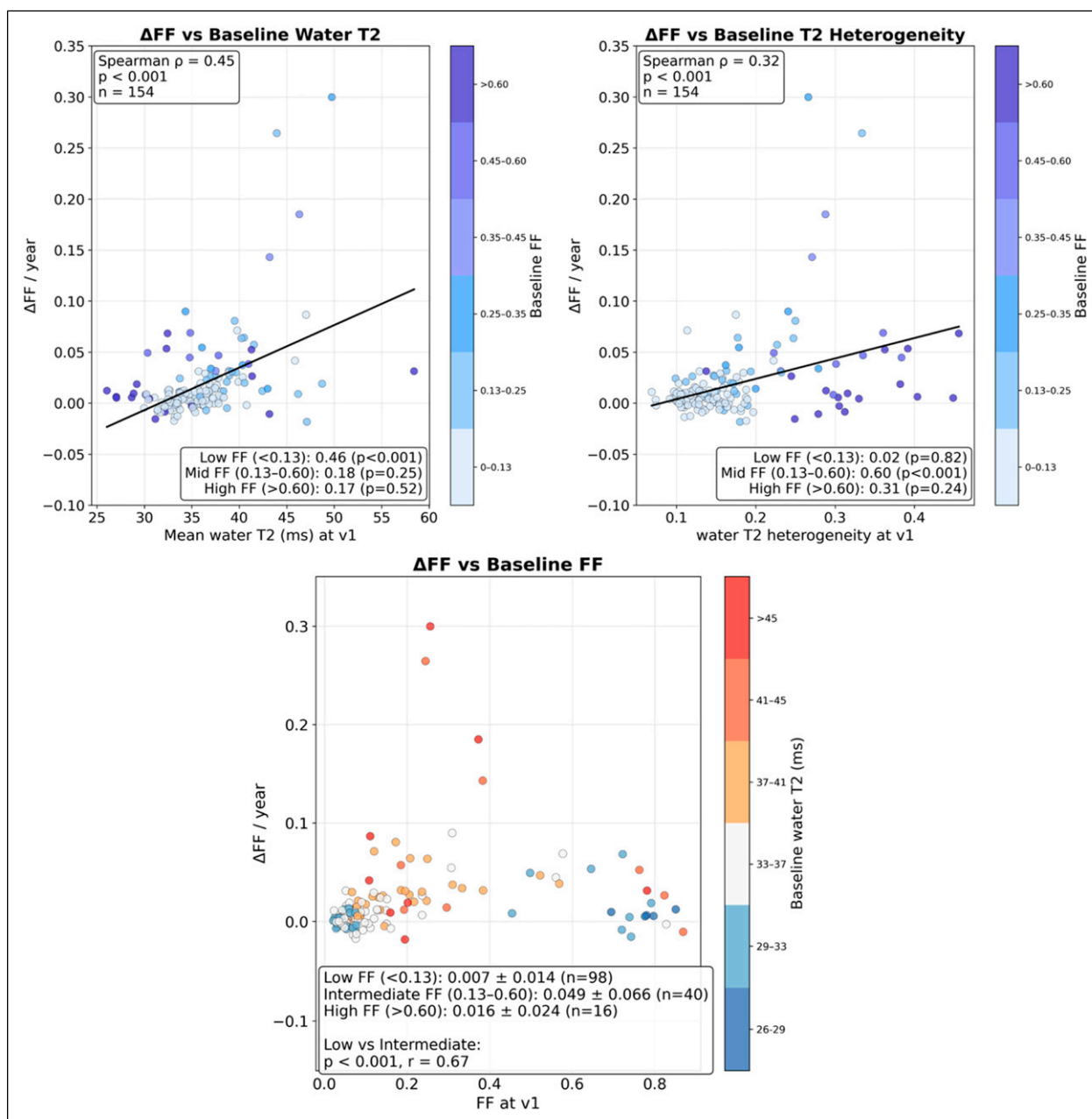


Figure 5. Association between baseline water T2 metrics (upper row), baseline FF (lower row) and subsequent fat fraction (FF) progression in FSHD. **Upper row:** Scatter plots show the relationship between Δ FF per year and baseline values, with points color-coded by baseline FF. Both baseline T2 mean ($\rho = 0.45$, $p < 0.001$) and T2 heterogeneity ($\rho = 0.32$, $p < 0.001$) were associated with subsequent FF increase. Stratified analyses indicate that T2 mean is most predictive in muscles with low baseline FF, whereas T2 heterogeneity shows the strongest association in muscles with intermediate FF, suggesting complementary roles of these metrics across disease stages. **Lower row:** Scatter plot showing the relationship between baseline fat fraction (FF at v1) and annual change in fat fraction (Δ FF/year) across individual muscles. Each point represents a single muscle, color-coded by baseline water T2 values. Muscles were stratified into three categories based on baseline FF: low (<0.13), intermediate ($0.13-0.60$), and high (>0.60). Mean Δ FF/year ($\pm$ standard deviation) and sample size are displayed for each subgroup. Muscles with intermediate baseline FF exhibited the highest progression rates compared to low FF muscles (Mann-Whitney U test), while muscles with high baseline FF showed limited further progression, consistent with a ceiling effect.

Sensitivity of MRI to Detect Disease Progression

Quantitative MRI proved more sensitive than semi-quantitative grading in capturing short-term disease progression. While T1w scores showed only limited change over one year, significant increases were observed in FF, FF heterogeneity (FF IQR), and water T2 heterogeneity. These findings indicate ongoing structural and compositional muscle changes despite minimal visual progression.

These findings are in line with previous longitudinal MRI studies in FSHD, which demonstrated that quantitative FF detects progression over relatively short intervals even when clinical measures remain stable.^{10,11,14} Similar observations have been made across neuromuscular disorders, where MRI-based biomarkers outperform strength and functional tests in sensitivity and reproducibility.^{28–31}

Progression was not uniform across disease stages. Muscles with intermediate baseline FF showed the greatest increase in fat replacement, whereas muscles with low FF remained largely stable, in line with prior reports identifying intermediate-stage muscles as most prone to progression.^{11–14}

Importantly, water T2 mean values did not change significantly over time, whereas water T2 heterogeneity increased and was already elevated in muscles with low FF. This suggests that averaging water T2 may mask localized changes, while heterogeneity captures spatially variable disease activity. These findings extend prior work focusing primarily on average T2 measures by highlighting the added value of heterogeneity metrics.^{18,32}

STIR and Water T2 as Markers of Disease Activity

Markers of disease activity at baseline, including STIR hyperintensity and elevated water T2, were preferentially observed in muscles with greater structural involvement, indicating that inflammatory-like processes preferentially occur in already affected muscles. This is consistent with previous imaging and histopathological studies linking STIR hyperintensity to inflammatory changes and early dystrophic alterations.^{33,34} Importantly, baseline STIR and water T2 abnormalities were associated with subsequent increases in quantitative FF, but not with changes in T1w scores. This supports the concept that STIR and water T2 reflect transitional disease states that precede measurable fat accumulation, as suggested by prior longitudinal imaging studies.^{10,15–17} Our data refine this model by demonstrating that these associations are more robust when progression is assessed quantitatively. The absolute water T2 values observed in our cohort were consistent with previous quantitative MRI studies. In healthy controls, mean water T2 values were 35.2 ± 2 ms using a three-exponential fitting model, comparable to values reported using similar T2 mapping approaches.^{32,35} Although absolute water T2 values depend on acquisition parameters, echo times, fat suppression techniques, and fitting methodology, increased water T2 values in STIR-positive FSHD muscles have been consistently reported across different approaches,^{18,32} supporting the interpretation of elevated water T2 as a marker of active muscle involvement.

However, the precise relationship between STIR abnormalities and the natural history of FSHD has remained unclear, as STIR lesions have been reported to persist without subsequent fat replacement in some muscles, while other muscles progress without a detectable STIR phase.^{36,37} Together, these findings support a stage-dependent interpretation of STIR abnormalities, with higher-grade or persistent changes indicating muscles at risk of near-term degeneration. Studies with extended follow-up periods may further refine our understanding of the temporal dynamics of STIR and water T2 abnormalities.

Relationship Between Imaging and Clinical Function

Both FF and T1w scores correlated with clinical measures at baseline, but associations were consistently stronger for FF. In particular, FF showed a strong association with the lower-limb composite MRC score and with individual lower-limb strength measures, supporting a relationship between quantitative muscle composition and strength of the assessed lower limb. This aligns with prior studies demonstrating correlations between muscle fat fraction and strength measures in FSHD.^{11,18,38}

FF and T1w scores were also associated with the overall mobility score. However, these associations should be interpreted cautiously because mobility represents a broader measure of functional performance and depends on multiple muscle groups beyond the lower leg assessed by MRI, including proximal lower-limb, trunk, and upper-limb muscles, as well as compensatory mechanisms.

Structural muscle involvement was also not invariably accompanied by proportional impairment in the clinical measures assessed. Some patients with advanced fatty replacement retained ambulation and relatively preserved strength, consistent with previous observations of discordance between imaging findings and functional performance.^{39,40} However, this observation should be interpreted in the context of the limited anatomical coverage of the MRI and the multifactorial

determinants of mobility and strength. Compensatory mechanisms and the non-linear relationship between muscle composition and function may contribute to this apparent dissociation.

Longitudinally, clinical changes were subtle. When assessed per limb, lower-limb MRC scores did not change significantly; however, stratification into stronger and weaker legs revealed a small but significant decline in the stronger leg only. This supports prior evidence that clinical measures are relatively insensitive to short-term progression, particularly in more severely affected muscles.^{11,12,14} The limited magnitude of longitudinal clinical change further supports the potential value of quantitative MRI for detecting muscle-level progression over short intervals.

Implications for Biomarker Development and Clinical Trials

Together, these findings highlight the limitations of semi-quantitative MRI and clinical measures for short-term monitoring in FSHD. Quantitative MRI, and particularly FF, provides a sensitive and objective measure of structural progression. In addition, water T2 heterogeneity represents a promising exploratory measure that may capture early, spatially heterogeneous disease activity not reflected by mean T2 values.

With emerging therapies targeting DUX4 and related pathways,⁴¹ sensitive biomarkers are essential to monitor treatment response. Our findings support quantitative MRI, including FF and heterogeneity measures, as robust outcome measures for both natural history studies and clinical trials. This is consistent with recent recommendations emphasizing MRI as a key biomarker in FSHD research.⁵

Clinical outcome measures such as MRC scores remain limited by variability and low sensitivity to change. Findings from placebo arms of prior trials, showing paradoxical short-term improvements in strength,^{42–44} further underscore their limitations for short-term studies. Incorporating MRI-based endpoints may therefore improve sensitivity, reduce variability, and enhance trial efficiency.

Limitations

Our study has several limitations. First, it was conducted at a single center with a relatively modest sample size, which may limit the generalizability of the findings. In addition, the cohort showed a predominance of male participants. Sex-related differences in skeletal muscle composition and response to degeneration have been reported,⁴⁵ and emerging evidence suggests that hormonal factors may influence disease severity and muscle regeneration pathways in FSHD.^{46,47} Larger, sex-balanced longitudinal cohorts would be needed to investigate in depth whether sex and hormones influence imaging biomarkers of disease progression.

Second, healthy controls underwent MRI only at baseline and were not examined longitudinally. However, previous reproducibility studies have demonstrated excellent longitudinal stability of quantitative muscle MRI parameters in healthy individuals,^{48–50} suggesting that the observed changes in the FSHD cohort are unlikely to be explained by measurement variability alone.

Third, quantitative MRI was acquired only in the clinically less affected leg to limit examination time and participant burden. Consequently, bilateral quantitative assessment of disease progression was not possible, although qualitative MRI was performed bilaterally. While this approach may increase sensitivity for detecting changes in relatively preserved muscle tissue, it does not capture the full extent of bilateral disease burden.

Fourth, muscle-wise analyses were performed under the simplifying assumption that individual muscles represented independent observations. Since muscles within the same patient may exhibit correlated disease evolution, particularly during active phases of FSHD, this approach may underestimate variability and increase the risk of inflated statistical significance. Although FDR correction was applied to account for multiple comparisons, it does not address within-patient clustering. Larger longitudinal studies should apply mixed-effects models or other hierarchical approaches to account for within-patient clustering.

Fifth, semi-quantitative scoring was based on averaged values across muscle segments. While this approach better reflects spatial heterogeneity, it may underestimate focal or peak abnormalities and limits direct comparability with studies that use maximum grading methods.

Conclusion

Quantitative MRI detects progressive muscle degeneration in FSHD over one year, whereas qualitative grading shows limited sensitivity. Baseline disease activity markers predict progression of fat fraction but not semi-quantitative scores. Measures of heterogeneity, particularly water T2 heterogeneity, provide additional sensitivity to early and spatially variable disease changes and may serve as valuable biomarkers in future studies and clinical trials.

Acknowledgement

The authors would like to thank all the participants who volunteered for this study.

ORCID iD

Teresa Gerhalter  <https://orcid.org/0000-0001-9734-4632>

Ethical considerations

All human studies were approved by the German ethics committee of Universitätsklinikum Erlangen (approval no. 20-291_4-B) in August 2023. An amendment allowing the inclusion of minors was approved in October 2023. The study has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. Written informed consent was obtained from all participants prior to inclusion in the study. In the case of minors, written informed consent was obtained from the legal guardian, and assent was obtained from the minor following an age-appropriate information and consent discussion with both the minor and the guardian.

Consent to participate

All participants gave their informed written consent prior to their inclusion in the study.

Author contributions

TG and ALM contributed to the conception and design of the study. ALM, MT, and EG recruited participants. ALM, TG, VS, AS were involved in data acquisition. VS and AS outlined the ROIs. ALM, TG performed data analysis. PYB developed and provided software for data analysis. ALM and TG drafted the manuscript and prepared the figures. AN, EG, MT, and PYB critically revised the manuscript for important intellectual content. MU and AD provided resources and infrastructure.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was funded by the Deutsche Forschungsgemeinschaft (DFG Nr. 497320374 and 500888779 / RU5534 MR biosignatures at UHF).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

De-identified data that support the findings of this study are available from the corresponding author upon reasonable request.

Supplemental material

Supplemental material for this article is available online.

References

1. Deenen JCW, Horlings CGC, Verschuuren JJGM, et al. The Epidemiology of Neuromuscular Disorders: A Comprehensive Overview of the Literature. *J Neuromuscul Dis* 2015; 2: 73–85.
2. Statland JM and Tawil R. Risk of functional impairment in Facioscapulohumeral muscular dystrophy. *Muscle & Nerve* 2014; 49: 520–527. <https://doi.org/10.1002/mus.23949>
3. Lemmers RJLF, Tawil R, Petek LM, et al. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet* 2012; 44: 1370–1374. <https://doi.org/10.1038/ng.2454>
4. Sacconi S, Salviati L and Desnuelle C. Facioscapulohumeral muscular dystrophy. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 2015; 1852: 607–614. <https://doi.org/10.1016/j.bbadis.2014.05.021>
5. Monforte M, Attarian S, Vissing J, et al. 265th ENMC International Workshop: Muscle imaging in Facioscapulohumeral Muscular Dystrophy (FSHD): relevance for clinical trials. 22–24 April 2022, Hoofddorp, The Netherlands. *Neuromuscular Disorders* 2023; 33: 65–75. <https://doi.org/10.1016/j.nmd.2022.10.005>
6. Murphy W, Totty WG and Carroll JE. MRI of normal and pathological skeletal muscle. *American Journal of Radiology* 1986; 146: 565–574. <https://doi.org/10.2214/ajr.146.3.565>
7. Vincenten SC, Teeselink S, Mul K, et al. Muscle imaging in facioscapulohumeral muscular dystrophy research: A scoping review and expert recommendations. *Neuromuscular Disorders* 2025; 47: 105274. <https://doi.org/10.1016/j.nmd.2025.105274>, Epub ahead of print 1.

8. Wattjes MP, Kley Ra. and Fischer D. Neuromuscular imaging in inherited muscle diseases. *European Radiology* 2010; 20: 2447–2460. <https://doi.org/10.1007/s00330-010-1799-2>
9. Poliachik SL, Friedman SD, Carter GT, et al. Skeletal Muscle Edema in Muscular Dystrophy: Clinical and Diagnostic Implications. *Physical Medicine and Rehabilitation Clinics of North America* 2012; 23: 107–122. <https://doi.org/10.1016/j.pmr.2011.11.016>
10. Janssen BH, Voet NBM, Nabuurs CI, et al. Distinct disease phases in muscles of facioscapulohumeral dystrophy patients identified by MR detected fat infiltration. *PloS one* 2014; 9: e85416. <https://doi.org/10.1371/journal.pone.0085416>
11. Andersen G, Dahlqvist JR, Vissing CR, et al. MRI as outcome measure in facioscapulohumeral muscular dystrophy: 1-year follow-up of 45 patients. *Journal of Neurology* 2017; 264: 438–447. <https://doi.org/10.1007/s00415-016-8361-3>
12. Wang LH, Shaw DWW, Faino A, et al. Longitudinal study of MRI and functional outcome measures in facioscapulohumeral muscular dystrophy. *BMC Musculoskeletal Disorders* 2021; 22: 262. <https://doi.org/10.1186/s12891-021-04134-7>
13. Mellion ML, Widholm P, Karlsson M, et al. Quantitative Muscle Analysis in FSHD Using Whole-Body Fat-Referenced MRI: Composite Scores for Longitudinal and Cross-sectional Analysis. *Neurology* 2022; 99: e877–e889. <https://doi.org/10.1212/wnl.0000000000200757>
14. Vincenten SCC, Mul K, van As D, et al. Five-year follow-up study on quantitative muscle magnetic resonance imaging in facioscapulohumeral muscular dystrophy: The link to clinical outcome. *Journal of Cachexia, Sarcopenia and Muscle* 2023; 14: 1695–1706. <https://doi.org/10.1002/jcsm.13250>, Epub ahead of print 1 August 2023.
15. Mul K, Vincenten SCC, Voermans NC, et al. Adding quantitative muscle MRI to the FSHD clinical trial toolbox. *Neurology* 2017; 89: 2057–2065. <https://doi.org/10.1212/WNL.0000000000004647>
16. Dahlqvist JR, Poulsen NS, Østergaard ST, et al. Evaluation of inflammatory lesions over 2 years in facioscapulohumeral muscular dystrophy. *Neurology*; 95: e1211–e1221. <https://doi.org/10.1212/WNL.0000000000010155>, Epub ahead of print September 2020.
17. Heskamp L, Ogier A, Bendahan D, et al. Whole-muscle fat analysis identifies distal muscle end as disease initiation site in facioscapulohumeral muscular dystrophy. *Communications Medicine* 2022; 2: 155. <https://doi.org/10.1038/s43856-022-00217-1>
18. Paoletti M, Monforte M, Barzaghi L, et al. Natural history of facioscapulohumeral muscular dystrophy evaluated by multiparametric quantitative MRI: a prospective cohort study. *Journal of Neurology*; 272: 306. <https://doi.org/10.1007/s00415-025-13062-8>, Epub ahead of print 1 April 2025.
19. van den Heuvel A, Lassche S, Mul K, et al. Facioscapulohumeral dystrophy transcriptome signatures correlate with different stages of disease and are marked by different MRI biomarkers. *Scientific Reports* 2022; 12: 1426. <https://doi.org/10.1038/s41598-022-04817-8>
20. Schlaffke L, Rehmann R, Güttches A-K, et al. Evaluation of Neuromuscular Diseases and Complaints by Quantitative Muscle MRI. *JCM* 2024; 13: 1958. <https://doi.org/10.3390/jcm13071958>
21. Reyngoudt H, Marty B, Boisserie J-M, et al. Global versus individual muscle segmentation to assess quantitative MRI-based fat fraction changes in neuromuscular diseases. *European Radiology* 2021; 31: 4264–4276.
22. Johnson NE, Quinn C, Eastwood E, et al. Patient identified disease burden in facioscapulohumeral muscular dystrophy. *Muscle and Nerve* 2012; 46: 948–950. <https://doi.org/10.1002/mus.23529>
23. Dyck PJ, Boes CJ, Mulder D, et al. History of standard scoring, notation, and summation of neuromuscular signs. A current survey and recommendation. *Journal of the peripheral nervous system : JPNS* 2005; 10: 158–173. <https://doi.org/10.1111/j.1085-9489.2005.0010206.x>
24. Baudin P, Reyngoudt H, Schunk V, et al. Effects of Lipid Induced Magnetic Microstructure on Fat Fraction Quantification in Muscular Dystrophies. *Magnetic Resonance in Med* 2026: 70565. <https://doi.org/10.1002/mrm.70565>, Epub ahead of print.
25. Azzabou N, Loureiro de Sousa P, Caldas E, et al. Validation of a generic approach to muscle water T2 determination at 3T in fat-infiltrated skeletal muscle. *Journal of Magnetic Resonance Imaging* 2015; 41: 645–653. <https://doi.org/10.1002/jmri.24613>
26. Goutallier D, Postel JM, Bernageau J, et al. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. *Clinical orthopaedics and related research* 1994; 304: 78–83.
27. Santini F, Wasserthal J, Agosti A, et al. Deep Anatomical Federated Network (Dafne): An open client-server framework for the continuous, collaborative improvement of deep learning-based medical image segmentation. *Radiol Artif Intell* 2025; 7: e240097. <https://doi.org/10.1148/ryai.240097>. Epub ahead of print 23.
28. Morrow JM, Sinclair CDJ, Fischmann A, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. *The Lancet Neurology* 2016; 15: 65–77. [https://doi.org/10.1016/S1474-4422\(15\)00242-2](https://doi.org/10.1016/S1474-4422(15)00242-2)
29. Barnard AM, Willcocks RJ, Triplett WT, et al. MR biomarkers predict clinical function in Duchenne muscular dystrophy. *Neurology* 2020; 94: e897–e909. <https://doi.org/10.1212/WNL.0000000000009012>
30. Barp A, Bello L, Caumo L, et al. Muscle MRI and functional outcome measures in Becker muscular dystrophy. *Scientific Reports* 2017; 7: 1–11. <https://doi.org/10.1038/s41598-017-16170-2>
31. Otto LAM, Froeling M, Van Eijk RPA, et al. Quantification of disease progression in spinal muscular atrophy with muscle MRI—a pilot study. *NMR in Biomedicine* 2021; 34: e4473. <https://doi.org/10.1002/nbm.4473>
32. Gerhalter T, Marty B, Gast LV, et al. Quantitative ¹H and ²³Na muscle MRI in Facioscapulohumeral muscular dystrophy patients. *Journal of Neurology* 2021; 268: 1076–1087.

33. Tasca G, Monforte M, Corbi M, et al. Muscle Microdialysis to Investigate Inflammatory Biomarkers in Facioscapulohumeral Muscular Dystrophy. *Molecular Neurobiology* 2018; 55: 2959–2966. <https://doi.org/10.1007/s12035-017-0563-x>
34. Frisullo G, Frusciante R, Nociti V, et al. CD8+ T Cells in Facioscapulohumeral Muscular Dystrophy Patients with Inflammatory Features at Muscle MRI. *J Clin Immunol* 2011; 31: 155–166. <https://doi.org/10.1007/s10875-010-9474-6>
35. Reyngoudt H, Baudin P, Carlier PG, et al. New Insights into the Spread of MRS Based Water T2 Values Observed in Highly Fatty Replaced Muscles. *Magnetic Resonance Imaging* 2023; 58: 1557–1568. <https://doi.org/10.1002/jmri.28669>
36. Friedman SD, Poliachik SL, Otto RK, et al. Longitudinal features of stir bright signal in FSHD. *Muscle and Nerve* 2014; 49: 257–260. <https://doi.org/10.1002/mus.23911>
37. Monforte M, Laschena F, Ottaviani P, et al. Tracking muscle wasting and disease activity in facioscapulohumeral muscular dystrophy by qualitative longitudinal imaging. *Journal of Cachexia, Sarcopenia and Muscle* 2019; 10: 1258–1265. <https://doi.org/10.1002/jcsm.12473>
38. Kan HE, Scheenen TWJ, Wohlgemuth M, et al. Quantitative MR imaging of individual muscle involvement in facioscapulohumeral muscular dystrophy. *Neuromuscular Disorders* 2009; 19: 357–362. <https://doi.org/10.1016/j.nmd.2009.02.009>
39. Regula JU, Jestaedt L, Jende F, et al. Clinical Muscle Testing Compared with Whole-Body Magnetic Resonance Imaging in Facioscapulo-humeral Muscular Dystrophy. *Clin Neuroradiol* 2016; 26: 445–455. <https://doi.org/10.1007/s00062-015-0386-y>
40. Tasca G, Monforte M, Ottaviani P, et al. Magnetic resonance imaging in a large cohort of facioscapulohumeral muscular dystrophy patients: Pattern refinement and implications for clinical trials. *Annals of Neurology* 2016; 79: 854–864. <https://doi.org/10.1002/ana.24640>
41. Ghasemi M, Emerson CP and Hayward LJ. Outcome Measures in Facioscapulohumeral Muscular Dystrophy Clinical Trials. *Cells* 2022; 11: 687. <https://doi.org/10.3390/cells11040687>
42. Statland JM, McDermott MP, Heatwole C, et al. Reevaluating measures of disease progression in facioscapulohumeral muscular dystrophy. *Neuromuscular Disorders* 2013; 23: 306–312. <https://doi.org/10.1016/j.nmd.2013.01.008>
43. Kissel JT, McDermott MP, Mendell JR, et al. Randomized, double-blind, placebo-controlled trial of albuterol in facioscapulohumeral dystrophy. *Neurology* 2001; 57: 1434–1440. <https://doi.org/10.1212/wnl.57.8.1434>
44. Van Der Kooi EL, Vogels OJM, Van Asseldonk RJGP, et al. Strength training and albuterol in facioscapulohumeral muscular dystrophy. *Neurology* 2004; 63: 702–708. <https://doi.org/10.1212/01.wnl.0000134660.30793.1f>
45. Emmert ME, Emmert AS, Goh Q, et al. Sexual dimorphisms in skeletal muscle: current concepts and research horizons. *Journal of Applied Physiology* 2024; 137: 274–299. <https://doi.org/10.1152/japplphysiol.00529.2023>
46. Stunnenberg BC, Raaphorst J, Deenen JCW, et al. Prevalence and mutation spectrum of skeletal muscle channelopathies in the Netherlands. *Neuromuscular Disorders* 2018; 28: 402–407. <https://doi.org/10.1016/j.nmd.2018.03.006>
47. Maiullari S, Mele G, Calandra P, et al. Estrogen rescues muscle regeneration impaired by DUX4 in a humanized xenograft mouse model. *Cell Death Dis* 2025; 16: 508. <https://doi.org/10.1038/s41419-025-07827-2>
48. Forsting J, Rehmann R, Rohm M, et al. Prospective longitudinal cohort study of quantitative muscle magnetic resonance imaging in a healthy control population. *NMR in Biomedicine* 2024; 37: e5214. <https://doi.org/10.1002/nbm.5214>
49. Grimm A, Meyer H, Nickel MD, et al. Repeatability of Dixon magnetic resonance imaging and magnetic resonance spectroscopy for quantitative muscle fat assessments in the thigh. *J cachexia sarcopenia muscle* 2018; 9: 1093–1100. <https://doi.org/10.1002/jcsm.12343>
50. Schlaffke L, Rehmann R, Rohm M, et al. Multi center evaluation of stability and reproducibility of quantitative MRI measures in healthy calf muscles. *NMR in Biomedicine* 2019; 32: e4119. <https://doi.org/10.1002/nbm.4119>