Quantitative Muscle Analysis in FSHD Using Whole-Body Fat-Referenced MRI Composite Scores for Longitudinal and Cross-sectional Analysis

Michelle L. Mellion, MD,* Per Widholm, MD,* Markus Karlsson, PhD, André Ahlgren, PhD, Rabi Tawil, MD, Kathryn R. Wagner, MD, PhD, Jeffrey M. Statland, MD, Leo Wang, MD, PhD, Perry B. Shieh, MD, PhD, Baziel G.M. van Engelen, MD, PhD, Joost Kools, MD, Lucienne Ronco, PhD, Adebowo Odueyungbo, PhD, John Jiang, PhD, Jay J. Han, MD, Maya Hatch, PhD, Jeanette Towles, Olof Dahlqvist Leinhard, PhD, and Diego Cadavid, MD

Neurology® 2022;99:e877-e889. doi:10.1212/WNL.0000000000200757

Abstract

Background and Objectives
Facioscapulohumeral muscular dystrophy (FSHD) is a rare, debilitating disease characterized by progressive muscle weakness. MRI is a sensitive assessment of disease severity and progression. We developed a quantitative whole-body (WB) musculoskeletal MRI (WB-MSK-MRI) protocol analyzing muscles in their entirety. This study aimed to assess WB-MSK-MRI as a potential imaging biomarker providing reliable measurements of muscle health that capture disease heterogeneity and clinically meaningful composite assessments correlating with severity and more responsive to change in clinical trials.

Methods
Participants aged 18–65 years, with genetically confirmed FSHD1, clinical severity 2 to 4 (Ricci scale, range 0–5), and ≥1 short tau inversion recovery–positive lower extremity muscle eligible for needle biopsy, enrolled at 6 sites and were imaged twice 4–12 weeks apart. Volumetric analysis of muscle fat infiltration (MFI), muscle fat fraction (MFF), and lean muscle volume (LMV) in 18 (36 total) muscles from bilateral shoulder, proximal arm, trunk, and legs was performed after automated atlas-based segmentation, followed by manual verification. A WB composite score, including muscles at highest risk for progression, and functional cross-sectional composites for correlation with relevant functional outcomes including timed up and go (TUG), FSHD-TUG, and reachable workspace (RWS), were developed.

Results
Seventeen participants enrolled in this study; 16 follow-up MRIs were performed at 52 days (range 36–85 days). Functional cross-sectional composites (MFF and MFI) showed moderate to strong correlations: TUG (ρ = 0.71, ρ = 0.83), FSHD-TUG (ρ = 0.73, ρ = 0.73), and RWS (left arm: ρ = −0.71, ρ = −0.53; right arm: ρ = −0.61, ρ = −0.65). WB composite variability: LMVtot coefficient of variation (CV) 1.9% and 3.4%; MFFtot, within-subject SD (S_w) 0.5% and 1.5%; and MFItot (S_w), 0.3% and 0.4% for normal and intermediate muscles, respectively. CV and S_w were higher in intermediate (MFI ≥0.10; MFF <0.50) than in normal (MFI <0.10, MFF <0.50) muscles.

Discussion
We developed a WB-MSK-MRI protocol and composite measures that capture disease heterogeneity and assess muscle involvement as it correlates with FSHD-relevant clinical endpoints. Functional composites robustly correlate with functional assessments. Stability of the WB composite shows that it could be an assessment of change in therapeutic clinical trials.

*These authors contributed equally to this work as co-first authors.

From Fulcrum Therapeutics (M.L.M., L.R., A.O., J.J., D.C.), Cambridge, MA; AMRA Medical AB (P.W., M.K., A.A., O.D.L.), Linköping; Departments of Radiology and Health (P.W.), Medicine and Caring Sciences, Linköping University; Centre for Medical Image Science and Visualization (CMIV) (P.W.), Linköping University, Sweden; University of Rochester Medical Center (R.T.), Rochester, NY; Kennedy Krieger Institute (K.R.W.), Johns Hopkins School of Medicine, Baltimore, MD; University of Kansas Medical Center (J.M.S.), Kansas City; University of Washington (L.W.), Seattle; Ronald Reagan UCLA Medical Center (P.B.S.), Los Angeles, CA; Radboud University Medical Center (B.G.M.E., J.K.), Nijmegen, the Netherlands; University of California-Irvine (J.J.H., M.H.), Orange; Synterex Inc. (J.T.), Dedham, MA; and Linköping University (O.D.L.), Sweden. Dr. Kools is currently at the Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Copyright © 2022 American Academy of Neurology

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.
Facioscapulohumeral muscular dystrophy (FSHD) is a rare, debilitating disease characterized clinically by skeletal muscle weakness affecting the face, shoulders, upper and lower extremities, trunk, and pelvic girdle. Identification of a biomarker or surrogate that is reasonably correlated with clinical outcome assessments such as reachable workspace (RWS) and timed up and go (TUG) will benefit the assessment of therapeutics for FSHD. Clinical outcomes are limited because they exclude nonambulatory participants, fail to capture the heterogeneity of individual muscle involvement, and may not be sensitive enough to assess the progressive changes typical of FSHD over the period(s) in which most clinical trials are conducted.1

Skeletal muscle MRI is noninvasive and nonirradiating and can be performed in most participants, irrespective of disease severity. Furthermore, in addition to volumetric measurements of muscle bulk, MRI can also measure fatty infiltration, fatty replacement of muscle, and lean muscle volume.1

In FSHD, several groups have confirmed strong correlations between skeletal muscle fat fraction (MFF) measured by MRI and clinical outcome assessments such as muscle strength in knee flexion and extension, distance walked and physical activity measured by accelerometry, motor function measure score, 6-minute walk test, and clinical severity scores.1-4 Strong correlations between fat replacement and muscle strength in FSHD have also been reported in cross-sectional studies using whole-body (WB) MRI qualitatively scored by visual inspection.3-5 A recent longitudinal study of the rate of progression of skeletal muscle replacement by fat in FSHD suggests that measurements made with MRI may be a biomarker that might be useful for measuring disease worsening over 6–12 months.3 Yearly progression in muscle fat replacement was observed using a composite score calculated from 12 different regions in the legs and thorax, with an excellent reproducibility of R = 0.99. Another MRI study limited to the lower extremities showed change in composite fat fraction at 1 year only in muscles in the intermediate stage of fat replacement at baseline.2

There is evidence that muscles with intermediate fat fractions between 0.10 and 0.70, or higher T2/short tau inversion recovery (STIR) burden, are most likely to progress over a short-term follow-up.2,3 Changes may be small, and a ceiling effect has been observed for progression in individual muscles with baseline fat fractions >0.60. In addition, these studies are limited to a single slice of specific regions of muscles and fail to provide a comprehensive skeletal muscle and WB assessment.

We hypothesized that this WB musculoskeletal MRI (WB-MSK-MRI) protocol imaging muscles in their entirety provides a reliable and reproducible wholistic evaluation of skeletal musculature capturing disease heterogeneity and contributes to the assessment of disease severity as it correlates with disease-relevant clinical outcome assessments (COAs). We developed a quantitative WB-MSK-MRI protocol analyzing muscles in their entirety. We sought to generate measurements of not only individual muscles but also WB and functional composite measurements that can accurately capture disease severity at the individual and population level in FSHD and show that they correlate with disease-relevant COAs. The WB composite score includes muscles at highest risk for progression, and the functional cross-sectional composites include relevant muscles for correlation with corresponding functional outcomes including TUG, FSHD-TUG, and RWS. This study aimed to assess this WB-MSK-MRI as a potential imaging biomarker providing reliable measurements of muscle health that capture disease heterogeneity and clinically meaningful composite assessments that correlate with severity and may be responsive to change in clinical trials.

Methods

Study Design and Participants
This was a prospective, observational, multicenter study conducted at 6 sites, 5 in the United States (University of Rochester, Kennedy Kreiger Institute, University of Washington, University of California Los Angeles, and University of Kansas Medical Center), and 1 in the Netherlands (Radboud University Medical Center).

This study aimed to enroll approximately 20 participants aged 18–65 years with genetically confirmed FSHD type 1, clinical
severity between 2 and 4 (on Ricci scale, range 0–5), and with at least 1 STIR-positive lower extremity muscle eligible for needle biopsy. Exclusion criteria included contraindications to muscle biopsies and/or musculoskeletal MRI.

Because FSHD is characterized by slow, variable progression and clinical studies often stretch over long periods of time, we aimed at describing the long-term reproducibility to account for normal physiologic changes not related to disease progression. Participants were asked to undergo WB-MSK-MRI scans at 2 visits approximately 4–12 weeks apart and complete standardized functional assessments including classic TUG, FSHD-TUG, and RWS.

### Standard Protocol Approvals, Registrations, and Participant Consents

Approval by an institutional review board/independent ethics committee was obtained at each site before study initiation. This study was consistent with the moral, ethical, and scientific principles governing clinical research, as set out in the Declaration of Helsinki and principles of Good Clinical Practice. Participants were screened for eligibility after signing the most current approved informed consent form and after the principal investigator or his/her designee provided a clear understanding of the study goals, demands from the schedule of assessments, and potential risks and mitigations.

### MRI Acquisition

All participants were scanned using 3T MRI scanners (Siemens, Erlangen, Germany, or Philips, Best, the Netherlands). Images were acquired in different sections including neck and shoulders, torso, unilateral arms, and legs. The MRI protocol was originally implemented with combined unilateral arm and torso imaging. To improve image quality and increase the consistency of participant positioning in the scanner, the protocol was updated to include separate torso and arm acquisitions.

The MRI Dixon technique was used to acquire images (Siemens: Dixon-Vibe; Philips: mDixon). Details on image acquisition parameters are provided in Supplementary Materials. These images were analyzed using AMRA Researcher (AMRA Medical, Linköping, Sweden). In brief, the image analysis consisted of the following parts, as reported separately: (1) automatic calibration using fat-referenced MRI, which converts the fat images into quantitative fat concentration maps; (2) automatic generation of labeled 3-dimensional muscle segmentations using nonrigid image registration to atlases; (3) manual quality control of image quality and segmentations; and (4) automatic calculation of muscle measurements (Table 1).

### Muscle Analysis

Three different measures were calculated for each muscle:

1. **MFF**: Relative amount of fat within the muscle fascia, calculated as the total volume of fat divided by the total muscle volume (TMV).
2. **Lean muscle volume (LMV)**: Amount of muscle tissue within the muscle, similar to contractile cross-sectional area. Calculated by subtracting the volume of fat from the TMV.
3. **Muscle fat infiltration (MFI)**: Fat fraction calculated within voxels with less than 50% fat. This measurement is a representation of the fat content in parts of the muscle that are not fat replaced and is analogous to the use of MR spectroscopy to measure fat fraction in homogenous muscle tissue.

Image and segmentation quality control was performed by trained operators, grading image quality and adjusting the segmentations to ensure anatomic correctness. Quality control of each data set was reviewed and approved by a second trained operator, with additional operators as arbitrators if needed.

<table>
<thead>
<tr>
<th>Muscle group</th>
<th>Muscle name</th>
<th>TUG</th>
<th>FSHD-TUG</th>
<th>RWS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotator cuff</td>
<td>Supraspinatus</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Infraspinatus</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subscapularis</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teres minor</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arm</td>
<td>Deltoideus</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biceps brachii</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Triceps brachii</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torso</td>
<td>Trapezius</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rhomboida</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Serratus anterior</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latissimus dorsib</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Pectoralis major</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Spinal erectors</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thigh</td>
<td>Quadricepsd</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hamstringse</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adductorsf</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Lower leg</td>
<td>Gastrocnemius medialis</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Tibialis anterior</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: FSHD = facioscapulohumeral muscular dystrophy; RWS = reachable workspace; TUG = timed up and go.

Note: Also shown are which muscles were included in which functional cross-sectional composites for correlation with the functional tests.

* Includes rhomboid major and rhomboid minor.

† Includes teres major.

‡ Includes erector spinae and transverso spinalis.

§ Includes rectus femoris and vastus lateralis/medialis/intermedius.

¶ Includes semimembranosus, semitendinosus, and biceps femoris.

‖ Includes brevis/longus/magnus/minimus, pectineus, and obturatorius externus.

Neurology.org/N Neurology | Volume 99, Number 9 | August 30, 2022
Signal quality issues (SQIs) were graded as either “minor” or “major.” A minor signal quality issue indicated that measurements for the muscle could be obtained but with decreased precision. A major signal quality issue indicated that no measurement could be obtained. Muscles with major and minor SQIs were excluded from the WB composite analysis.

**Muscle Categorization**

Based on previous literature, individual muscles were graded according to their calculated baseline MFF and MFI as normal-appearing (category A, MFI <0.10, MFF <0.50), intermediate (category B, MFI ≥0.10; MFF <0.50), and late-stage (category C, MFF ≥0.50). For defining the lower range, data from general population research in the UK Biobank indicate that normal MFI values in thigh muscles are lower than 10%.14 For defining the upper range, previous studies have reported a high progression rate in muscles with MFF up to 50%–60%.2,3,15 In addition, studies of both FSHD and Duchenne muscular dystrophy have shown that MFF >50% is associated with loss of muscle function.2,16,17 Muscle segmentation methodologies and quantification techniques across protocols likely account for the slight differences in the upper limit definition for intermediate muscles.

**Composite Scores**

The measurements of the individual muscles can be combined into composite scores, describing the total LMV, MFF, or MFI for a group of muscles:

\[
\text{LMV}_{\text{tot}} = \sum_i \text{LMV}_i \\
\text{MFF}_{\text{tot}} = 1 - \frac{\sum_i \text{LMV}_i}{\sum_i \text{TMV}_i} \\
\text{MFI}_{\text{tot}} = \frac{\sum_i (\text{MFI}_i \times \text{LMV}_i)}{\sum_i \text{LMV}_i}
\]

where TMV is the total muscle volume, segmented inside the muscle fascia.

**WB Composite Score**

The purpose of the WB composite was to include an individualized set of muscles at high risk for progression (“B” muscles) from each participant to measure change over time. The WB composite score includes only those muscles that were measurable at both observation time points and had no signal quality issues. Figure 1 shows that while the same muscles were imaged by MRI, the muscles and number of muscles included for analysis may differ among participants.2,3

When evaluating change of the WB composite, we use absolute change (in percentage points) for MFF\textsubscript{tot} and MFI\textsubscript{tot} while relative changes were used for LMV\textsubscript{tot}. The rationale for using relative changes in LMV\textsubscript{tot} is that different participants will have different amounts of intermediate muscles included for longitudinal analysis, so comparison of absolute changes in LMV\textsubscript{tot} is not meaningful.
The use of relative change for LMV\textsubscript{tot} would cause the effects of measurement noise to be amplified if a participant only had a few small intermediate affected muscles. Therefore, a criterion for excluding participants with too small LMV\textsubscript{tot} was adopted as baseline LMV\textsubscript{tot} <25 cL or LMV\textsubscript{tot} <50 cL if the participant had only 1 intermediate muscle. The rationale for this exclusion criterion is that smaller muscles have a larger coefficient of variation (CV), reducing the overall reliability of the composite LMV change from baseline endpoint.

**Functional Cross-sectional Composite Score**

Functional cross-sectional composite scores are composed of muscles that are involved in a specific functional assessment of interest. Muscles associated with each functional assessment (TUG, FSHD-TUG, and RWS) were identified in consultation with a physical therapist (M.H.) (Table 1). Measurements for muscles with major quality issues (i.e., muscles with missing values) were imputed. Details of the imputation are provided in Supplementary Materials.

The functional cross-sectional composite scores were correlated with relevant clinical outcome assessments at baseline.

**Functional Tests**

**Classic TUG and FSHD-TUG**

The classic TUG test is used to assess a person’s mobility and requires both static and dynamic balance. The classic TUG measures the time that a person takes to rise from a chair, walk 3 meters, turn around, walk back to the chair, and sit down. The FSHD-TUG test adds to the classic TUG a supine-to-sit component (from a therapy mat in the clinic) at the start of the test and a sit-to-supine component at the end of the test.\textsuperscript{8,18,19} Each TUG test type (i.e., classic and FSHD) was conducted twice per visit.

**RWS**

For the RWS evaluation, participants followed previously published protocols.\textsuperscript{20-22} In brief, participants were seated in front of the Microsoft Kinect sensor 2.0 (Redmond, WA) and asked to perform an upper extremity movement protocol while following video instructions under the supervision of a study clinical evaluator. Participants were asked to complete the protocol with and without 500 g wrist weights and on both the right and left arms. On completion of the protocol, total relative surface area (RSA) of an individual’s range of motion was determined from the motion sensor and software algorithm (surface envelope). The RWS software automatically divides the total reachable space into 4 quadrants (upper medial and lateral, lower medial and lateral), with the shoulder joint serving as the origin. Absolute total and quadrant RWS surface envelopes (m\textsuperscript{2}) were normalized to each individual’s arm length to obtain RSAs for comparison among participants.

For correlation analyses with MRI data, muscle groupings were based on known biomechanical models for the shoulder complex/shoulder movement.\textsuperscript{23,24} Because the movements within the RWS outcome are largely flexion-based, extension-based, and scaption-based (scapular plane elevation) arm tasks, the scapulothoracic muscles were selected as the most appropriate for this imaging/function analysis in the shoulder girdle (Table 1).

**Statistical Analyses**

There was no formal sample size calculation for this biomarker preparatory study. The sample size was chosen to provide a sufficient number of skeletal muscle MRI examinations to estimate reliability and variability under natural history conditions.

Statistical analyses were performed using R Studio version 1.2.5033. Reproducibility of the longitudinal composite scores was evaluated using estimates of variance including within-subject SD (\(s_w\))\textsuperscript{25} which is the variation within individual participants and within subject CV.\textsuperscript{26} \(s_w\) was estimated according to

\[
s_w = \frac{1}{\sqrt{2}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \text{Diff}_i^2},
\]

where \(\text{Diff}_i\) is the difference between scans for participant \(i\) and \(N\) is the total number of participants.

---

**Table 2** Number and Percentages of Muscles in the Different MRI Categories for the Different Muscle Groups

<table>
<thead>
<tr>
<th>Muscle group</th>
<th>Normal</th>
<th>Intermediate</th>
<th>Late-stage*</th>
<th>Minor SQI</th>
<th>Major SQI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotator cuff</td>
<td>110 (81)</td>
<td>6 (4)</td>
<td>3 (2)</td>
<td>17 (13)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Arms</td>
<td>64 (63)</td>
<td>11 (11)</td>
<td>13 (13)</td>
<td>10 (10)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Torso</td>
<td>12 (6)</td>
<td>47 (23)</td>
<td>127 (62)</td>
<td>11 (5)</td>
<td>7 (3)</td>
</tr>
<tr>
<td>Thighs</td>
<td>20 (20)</td>
<td>34 (33)</td>
<td>35 (34)</td>
<td>10 (10)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Lower leg</td>
<td>11 (16)</td>
<td>27 (40)</td>
<td>25 (37)</td>
<td>3 (4)</td>
<td>2 (3)</td>
</tr>
</tbody>
</table>

Abbreviation: SQI = signal quality issue. Data are reported as n (%).

*Muscles that could not be analyzed because of high fat replacement are included in the late-stage category.
CV was estimated by calculating the coefficient of variation for each participant and then taking the root mean square of those:

$$CV = \sqrt{\frac{1}{N} \sum_{i=1}^{N} CV_i^2},$$

where $CV_i$ is the coefficient of variation for participant $i$.

Correlation between functional tests and functional cross-sectional composite scores was assessed using the Spearman correlation coefficient.

**Data Availability**

Data from this study will not be made available.

**Results**

**Study Population**

Seventeen participants enrolled in this study. Seventeen baseline and 16 follow-up MRIs were available for analysis. Follow-up scans were performed on average at 52 days (range 36–85 days). One participant discontinued after the first visit because they did not want to complete the clinical procedures for the second visit. The mean age of participants entering this study was 49.4 years (range: 23–65 years), 70.6% were male, and they had a mean (SD) Ricci score of 3.0 (0.71), with a Ricci score of 2 for 1 participant, 2.5 for 8 participants, 3 for 3 participants, and 4 for 5 participants.

**Muscle Categorization**

Twenty-one scans were performed under the original MRI protocol; 12 scans were performed with the updated MRI protocol. Nine participants were scanned twice with the original MRI protocol, 1 subject was scanned only once with the original MRI protocol and aborted this study before the second scan, 5 participants were scanned twice with the updated MRI protocol, and 2 participants were scanned with the original MRI protocol at the first visit and with the updated protocol at the second visit.

In total, 612 muscles were included for analysis. Table 2 presents how many muscles were categorized into the different categories. Sixteen muscles (3%) were not analyzed because of technical issues (Table 2; major SQI), and 51 muscles (8%) were analyzed but still had some technical issues, excluding the muscles from further longitudinal analysis. In addition, another 67 muscles (11%) were not analyzed because the muscles were too heavily fat replaced, including all trapezius muscles; these muscles are classified as late-stage. On average, each participant had 12.8 normal muscles (range 7–21), 7.4 intermediate muscles (range 2–12), 10.0 late-stage muscles (range 5–19), and 3.9 (range 0–9) muscles that were excluded due to image artifacts. Figure 2 shows examples of muscles of the different categories.

Figure 3C shows schematic illustrations of the analyzed muscles in each participant, colored according to muscle category, and Figure 4 shows the distribution of MFI and MFF for each muscle. These figures illustrate that the MRI protocol and quantitative assessment capture the heterogeneity of muscles affected in FSHD. In addition, the pattern of muscle involvement is as expected, with the rotator cuff muscles being the least affected.

**Analysis of Reproducibility**

**WB Composite Analysis**

All 16 participants with 2 MRI scans had intermediate muscles without any SQI. On average, participants had an LMV$_{tot}$ of 68.4% (range 40.0%–90.0%), with 11 participants having intermediate muscle categories (11.2% (range 7.7%–16.5%)) and 5 participants with late-stage muscle categories (4.9% (range 0.0%–11.4%)). The mean (SD) MFI was 1.6 (0.5), with a range of 0.4 to 4.2. The mean (SD) MFF was 2.0 (0.9), with a range of 0.1 to 5.2.

The reproducibility of the WB composite analysis was high, with a correlation coefficient of 0.97 (95% CI 0.96–0.98) for the lateral and 0.86 (95% CI 0.80–0.90) for the medial group. The mean (SD) difference was 0.06 (0.13) for the lateral and 0.25 (0.30) for the medial group.

**Discussion**

The results of this study indicate that the MRI protocol and quantitative assessment capture the heterogeneity of muscles affected in FSHD. The pattern of muscle involvement is as expected, with the rotator cuff muscles being the least affected. The reproducibility of the WB composite analysis was high, with a correlation coefficient of 0.97 (95% CI 0.96–0.98) for the lateral and 0.86 (95% CI 0.80–0.90) for the medial group. The mean (SD) difference was 0.06 (0.13) for the lateral and 0.25 (0.30) for the medial group.

**Conclusion**

This study provides evidence for the reproducibility and validity of the MRI protocol and quantitative assessment for measuring muscle involvement in FSHD. The results suggest that the MRI protocol and quantitative assessment capture the heterogeneity of muscles affected in FSHD, with the rotator cuff muscles being the least affected. The reproducibility of the WB composite analysis was high, with a correlation coefficient of 0.97 (95% CI 0.96–0.98) for the lateral and 0.86 (95% CI 0.80–0.90) for the medial group. The mean (SD) difference was 0.06 (0.13) for the lateral and 0.25 (0.30) for the medial group.
Each participant’s baseline results are schematically presented with a colored box for each muscle where the color represents the muscle category (C). The figure also contains a summarizing bar plot, showing how many muscles of each category each participant has (B), as well as a bar plot showing the $LMV_{\text{tot}}$ of each participant’s normal and intermediate muscles (A). The figure also shows the Ricci score as well as TUG and FSHD-TUG results for each participant.

Add = adductors; Bi = biceps brachii; Delt = deltoideus; FSHD = facioscapulohumeral muscular dystrophy; GM = gastrocnemius medialis; Ham = hamstrings; Inf = infraspinatus; Lats = latissimus dorsi; $LMV_{\text{tot}}$ = total lean muscle volume; MFF = muscle fat fraction; MFI = muscle fat infiltration; Pec = pectoralis; Quad = quadriceps; Rho = rhomboideus; SA = serratus anterior; SE = spinal erector; Sub = subscapularis; Sup = supraspinatus; TA = tibialis anterior; TM = teres minor; Trap = trapezius; Tri = triceps brachii; TUG = timed up and go.
normal muscles of 308 cL (range 66–981 cL) and an LMV<sub>tot</sub> of intermediate muscles of 222 cL in “B” muscles (range 7–577 cL). Two participants did not meet the criteria for minimum LMV<sub>tot</sub> of intermediate muscles; these participants were excluded from longitudinal composite test-retest reliability analysis. In total, 14 participants were included in this analysis.

Bland Altman plots are shown in Figure 4. For LMV<sub>tot</sub>, the CV was 3.4% for the intermediate muscles and 1.9% for the normal muscles. The s<sub>w</sub> for MFF<sub>tot</sub> was 0.5% and 1.5% and for MFI<sub>tot</sub> 0.3% and 0.4% for normal and intermediate muscles, respectively.

To investigate the effect of excluding muscles with minor SQI from the WB longitudinal composites, a sensitivity analysis was conducted calculating the composites also including the muscles with minor SQIs (51 muscles). There was no major difference in observed test-retest reliability for LMV<sub>tot</sub>, MFF<sub>tot</sub>, or MFI<sub>tot</sub> and only small differences in MFF and MFI magnitudes.
**Functional Cross-sectional Composite Correlations With Clinical Outcome Assessments**

On cross-sectional analysis, the MFF\textsubscript{tot} and MFI\textsubscript{tot} showed mostly strong correlations with functional tests (Figure 5). Strong correlations were observed between MFF\textsubscript{tot} and MFI\textsubscript{tot} of the leg muscles and TUG (ρ = 0.71 and ρ = 0.83, both p < 0.01; Figure 5, A and B), between MFF\textsubscript{tot} and MFI\textsubscript{tot} of all muscles and FSHD-TUG (ρ = 0.73 and ρ = 0.73, both p < 0.01; Figure 5, C and D), between MFF\textsubscript{tot} and MFI\textsubscript{tot} of the left scapula thoracic muscles and the RWS RSA for the left arm (ρ = −0.71 [p < 0.01] and ρ = −0.53 [p = 0.04]; Figure 5, E and F), and between MFF\textsubscript{tot} and MFI\textsubscript{tot} of the right scapula thoracic muscles and the RWS RSA for the right arm (ρ = −0.61 [p = 0.02] and ρ = −0.65 [p = 0.01]; Figure 5, G and H).

One participant was excluded from the RWS analysis because the participant did not correctly follow the instructions.

No functional correlations were performed with LMV\textsubscript{tot} because muscle size is so heavily affected by body size, age, and sex. Further work with muscle size normalization is ongoing to enable such analysis.

We developed a WB-MSK-MRI protocol and composite analyses that capture the heterogeneity of the disease at both the participant and skeletal muscle levels and provide important information about disease severity relative to FSHD clinical endpoints. Muscles for the volumetric analysis were identified by manually verified automated atlas-based segmentation. Endpoints included MFI, MFF, and LMV. This WB-MSK-MRI protocol and composite measures can capture disease heterogeneity and assess muscle involvement as it correlates with FSHD-relevant clinical endpoints. The functional composites are robustly correlated with functional assessments, and WB composite can be used to measure longitudinal changes in therapeutic clinical trials.

**Classification of Evidence**

This study provides Class II evidence that quantitative WB-MSK-MRI findings associate with FSHD1 severity measured using established functional assessments.

**Discussion**

In this study, we developed a WB-MSK-MRI protocol and composite analyses that capture the heterogeneity of the disease at the participant and skeletal muscle levels and provide important information about disease severity relative to FSHD clinical endpoints. The WB composite for measurement of changes was found to be highly reproducible on repeated testing. The functional cross-sectional composite showed strong correlations with FSHD-relevant clinical outcome measures. These data support that these MRI assessments may be possible surrogate biomarkers for clinical function in FSHD and should be investigated further in longitudinal clinical trials.

The muscle classification scheme differs from most previous studies in that it is able to quantitatively capture the heterogeneity of muscle involvement in FSHD and does so across the WB—not only in selected cross-sections in the legs. One of the most striking examples of this is participant 2 and participant 4 in Figure 3C. Both participants have the same Ricci score, but all of participant 4’s thigh muscles are classified as normal, whereas none of participant 2’s thigh muscles are classified as normal. This highlights the advantage of using the WB-MSK-MRI approach because all major muscles can be imaged to build a composite score of the muscles with active disease, regardless of where the muscles are located.

Skeletal muscles were classified per baseline MFI and MFF. Only the intermediate (“B”) muscles were included in the WB composite because these are the muscles at highest risk for progression. It is important that, and relevant to our selected cutoffs for intermediate muscles, in a recent study, the fastest progressing muscles had a baseline MFF between 10% and 50%. Higher variability in intermediate muscles is expected because these muscles are actively affected by disease-causing pathology of muscle structure as compared with normal-appearing muscles. Most participants in this study had intermediate muscles that met criteria to be included in the WB composite analysis despite not having an eligibility criterion specific for this condition. The muscles included and the number of muscles included in the composite may differ between participants, although the same muscles and the same number of muscles were imaged at each time point for each participant. The purpose of this composite assessment was to include an individualized set of muscles from each participant enriched to measure change over time at the individual and population level. Although this approach includes different muscles for each participant to be followed over time, it addresses an important challenge of disease heterogeneity and the difficulties of assessing for changes in slowly progressing diseases such as FSHD and other neuromuscular disorders (NMDs). In this study, stability of the WB composite over this relatively short duration shows that it could be a viable assessment of change in therapeutic clinical trials. It is critical to include only those muscles that are likely to progress because inclusion of muscles that are unlikely to change over the observation period may obfuscate measurement of disease progression and efficacy assessment of candidate therapeutics. Longitudinal studies are necessary to further evaluate this assumption.

In addition to assessment of WB composite measurement stability, it was important to determine the correlations of skeletal muscle MRI with disease-relevant clinical outcome assessments to understand the relevance of these MRI assessments in participant performance and to provide evidence of surrogacy. Only TUG, FSHD-TUG, and RWS all measure key functional impairments in FSHD. Our results showed moderate to strong correlations between MFF\textsubscript{tot} and MFI\textsubscript{tot} and TUG, FSHD-TUG, and RWS, supporting that these
Figure 5 Correlation Between MRI Biomarkers and Functional Tests

(A) TUG vs MFF$_{\text{total}}$. (B) TUG vs MFI$_{\text{total}}$. (C) FSHD-TUG vs MFF$_{\text{total}}$. (D) FSHD-TUG vs MFI$_{\text{total}}$. (E) Left arm RWS vs MFF$_{\text{total}}$. (F) Left arm RWS vs MFI$_{\text{total}}$. (G) Right arm RWS vs MFF$_{\text{total}}$. (H) Right arm RWS vs MFI$_{\text{total}}$. Note that one of the participants (red dot) had a TUG/FSHD-TUG time that was outside the range of the y axis. Correlation coefficient is Spearman $\rho$. There are strong cross-sectional correlations of MFI$_{\text{total}}$ and MFF$_{\text{total}}$ with TUG, FSHD-TUG, and RWS. FSHD = facioscapulohumeral muscular dystrophy; MFI$_{\text{total}}$ = total muscle fat infiltration; MFF$_{\text{total}}$ = total muscle fat fraction; RSA = relative surface envelope area; RWS = reachable workspace; RWS$_{\text{left}}$ = left arm reachable workspace; RWS$_{\text{right}}$ = right arm reachable workspace; TUG = timed up and go.
measures may be possible surrogates for clinical function in FSHD. For cross-sectional analysis, MFItot or MFFtot is probably a better assessment than LMVtot. LMVtot could be more meaningful once normative data are available (e.g., age, sex, and body mass index). Furthermore, MFItot and MFFtot are measures that have absolute meaning at any time point because they are reported as percentages at that time point. Similar correlations between MRI measurements and functional tests have also been shown by others, not only in FSHD but also in various other NMDs.12

Limitations of this study include the relatively small sample size and the inability to analyze certain muscles because of technical issues and late-stage status (i.e., too heavily fat replaced), although sensitivity analysis including those muscles with minor SQI showed that there was no difference. In addition, the MRI protocol was updated during this study to obtain better images. These limitations will be taken into account as protocol optimization continues.

In this study, we performed a WB-MSK-MRI protocol analyzing upper/lower extremity and trunk muscles in their entirety using 3-dimensional imaging and volumetric analysis in a multicenter setting. We developed a WB-MSK-MRI protocol and composite measures that capture disease hetrogeneity and assess muscle involvement as it correlates with FSHD-relevant clinical endpoint. Functional composites are robustly correlated with functional assessments. Stability of the WB composite shows that it could be an assessment of change in therapeutic clinical trials. These data support the potential of WB-MSK-MRI as a useful tool for assessing disease progression and to study treatment effects. This WB-MSK-MRI protocol and composites provide the ability to study participants across a wide spectrum of disability and are currently being used in 2 FSHD clinical trials to assess disease progression and correlations with selected clinical outcome assessments during treatment with losmapimod, a small molecule inhibitor of p38 MAPK α/β, to reduce DUX4 expression (NCT04264442; NCT04004000).28

Study Funding
This study was funded by Fulcrum Therapeutics.

Disclosure
M.L. Mellion is an employee of Fulcrum Therapeutics at the time of this work and may hold stock or stock options. P. Widholm is an employee of AMRA Medical AB and holds stock or stock options. M. Karlsson is an employee of AMRA Medical AB and holds stock or stock options. A. Ahlgren is an employee of AMRA Medical AB and holds stock or stock options. R. Tawil serves as a consultant for Fulcrum Therapeutics, Dyne Therapeutics, MT Pharma, Arrowhead Pharma, and Avidity Biosciences. K.R. Wagner received research support from Fulcrum Therapeutics, Pfizer, Sarepta Therapeutics, Catabasis, Fulcrum Therapeutics, and PTC Therapeutics; grant support from National Institute of Neurological Disorders and Stroke and NIAMS; and consulting fees from Dynacure, PTC Therapeutics, Casma Therapeutics, Asklepios BioPharmaceutical, Inc., Santhera, Regenxbio, and Sarepta Therapeutics; served on data monitoring committee for Fibrogen, Inc.; served on advisory boards for FSHD Society and Vita Therapeutics; receives royalties from Jones and Barlett Learning; and holds a patent for “Isolation Of Fusion-competent Myoblasts And Therapeutic Applications Thereof Related To Muscular Dystrophy” (patent number: 15/781709). J. Statland received grant and/or research funding from Fulcrum Therapeutics, Dyne Therapeutics, Avidity, MT Pharma, Sarepta, Acceleron, National Institute of Neurological Disorders and Stroke, MDA, FSHD Society, and Friends of FSH Research. L.H. Wang received grant funding from NIH and serves as a consultant for Fulcrum Therapeutics and AskBio Pharmaceuticals. P. Shieh received grant support from Fulcrum Therapeutics and received consulting/speaking fees and/or grant funding to his institution (outside the submitted work) from Sarepta Therapeutics, AveXis, PTC Therapeutics, Genentech, Pfizer, Biogen, Ra Pharma, Argex, Catalyst Pharmaceuticals, Fulcrum Therapeutics, Sanofi, Santhera, Grifols, and CSL Behring. B.G.M. van Engelen received grant support from Global FSH, Stichting Sperien voor Sperien, Princes Beatrix Spierfonds and Dutch FSHD Foundation; served on an advisory board for Fulcrum Therapeutics; received grant support (outside the submitted work) from European Union’s Horizon 2020 research and innovation programme (Murab), Netherlands Organisation for Scientific Research (NWO), The Netherlands Organisation for Health Research and Development (ZonMw), and Association Francaise contre les Myopathies; and holds a patent for “Myositis” (patent number: EP2012740236). J. Kools received grants and nonfinancial support from Fulcrum Therapeutics. L. Ronco is an employee of Fulcrum Therapeutics at the time of this work and may hold stock or stock options. A. Odueyungbo is an employee of Fulcrum Therapeutics at the time of this work and may hold stock or stock options. J. Jiang is an employee of Fulcrum Therapeutics at the time of this work and may hold stock or stock options. J. Han received research support from Fulcrum Therapeutics. M. Hatch received research support from Fulcrum Therapeutics, served as consultant for Bioniks, and served as part of the research team that developed the reachable workspace. J. Towles served as consultant for Fulcrum Therapeutics. O. Dahlqvist Leinhard is an employee of AMRA Medical AB and holds stock or stock options. D. Cadavid is an employee of Fulcrum Therapeutics at the time of this work and may hold stock or stock options. Go to Neurology.org/N for full disclosures.

Publication History
Received by Neurology September 13, 2021. Accepted in final form April 6, 2022. Submitted and externally peer reviewed. The handling editors were Anthony Amato, MD, FAAN.
### Appendix Authors

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michelle L. Mellion, MD</td>
<td>Fulcrum Therapeutics, Cambridge, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data; other</td>
</tr>
<tr>
<td>Per Widholm, MD</td>
<td>AMRA Medical AB, Linköping; Departments of Radiology and Health, Medicine and Caring Sciences, Linköping University; Centre for Medical Image Science and Visualization (CMIV), Linköping University, Sweden</td>
<td>Study concept or design; analysis or interpretation of data; other</td>
</tr>
<tr>
<td>Markus Karlsson, PhD</td>
<td>AMRA Medical AB, Linköping, Sweden</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; study concept or design</td>
</tr>
<tr>
<td>André Ahlgren, PhD</td>
<td>AMRA Medical AB, Linköping, Sweden</td>
<td>Major role in the acquisition of data</td>
</tr>
<tr>
<td>Rabi Tawil, MD</td>
<td>University of Rochester Medical Center, Rochester, NY</td>
<td>Study concept or design; other</td>
</tr>
<tr>
<td>Kathryn R. Wagner, MD, PhD</td>
<td>Kennedy Krieger Institute, Johns Hopkins School of Medicine, Baltimore, MD</td>
<td>Major role in the acquisition of data</td>
</tr>
<tr>
<td>Jeffrey M. Statland, MD</td>
<td>University of Kansas Medical Center, Kansas City</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design</td>
</tr>
<tr>
<td>Leo Wang, MD, PhD</td>
<td>University of Washington, Seattle</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design</td>
</tr>
<tr>
<td>Perry B. Shieh, MD, PhD</td>
<td>Ronald Reagan UCLA Medical Center, Los Angeles, CA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design</td>
</tr>
<tr>
<td>Baziil G.M. van Engelen, MD, PhD</td>
<td>Radboud University Medical Center, Nijmegen, the Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design</td>
</tr>
<tr>
<td>Joost Kools, MD</td>
<td>Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design</td>
</tr>
<tr>
<td>Lucienne Ronco, PhD</td>
<td>Fulcrum Therapeutics, Cambridge, MA</td>
<td>Study concept or design; analysis or interpretation of data; other</td>
</tr>
<tr>
<td>Adefowope Oduyeungbo, PhD</td>
<td>Fulcrum Therapeutics, Cambridge, MA</td>
<td>Analysis or interpretation of data; other</td>
</tr>
</tbody>
</table>

### Appendix (continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>John Jiang, PhD</td>
<td>Fulcrum Therapeutics, Cambridge, MA</td>
<td>Analysis or interpretation of data; other</td>
</tr>
<tr>
<td>Jay J. Han, MD</td>
<td>University of California-Irvine, Orange</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data; other</td>
</tr>
<tr>
<td>Maya Hatch, PhD</td>
<td>University of California-Irvine, Orange</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data; other</td>
</tr>
<tr>
<td>Jeanette Towles</td>
<td>Synterex Inc., Dedham, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content</td>
</tr>
<tr>
<td>Olof Dahlqvist Leinhard, PhD</td>
<td>AMRA Medical AB, Linköping; Linköping University, Sweden</td>
<td>Study concept or design; analysis or interpretation of data; other</td>
</tr>
<tr>
<td>Diego Cadavid, MD</td>
<td>Fulcrum Therapeutics, Cambridge, MA</td>
<td>Study concept or design; analysis or interpretation of data; other</td>
</tr>
</tbody>
</table>

### References


Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.


Quantitative Muscle Analysis in FSHD Using Whole-Body Fat-Referenced MRI: Composite Scores for Longitudinal and Cross-sectional Analysis
Michelle L. Mellion, Per Widholm, Markus Karlsson, et al.
Neurology 2022;99;e877-e889 Published Online before print June 24, 2022
DOI 10.1212/WNL.0000000000200757

This information is current as of June 24, 2022

Updated Information & Services
including high resolution figures, can be found at:
http://n.neurology.org/content/99/9/e877.full

References
This article cites 22 articles, 5 of which you can access for free at:
http://n.neurology.org/content/99/9/e877.full#ref-list-1

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Class II
http://n.neurology.org/cgi/collection/class_ii
MRI
http://n.neurology.org/cgi/collection/mri

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
http://www.neurology.org/about/about_the_journal#permissions

Reprints
Information about ordering reprints can be found online:
http://n.neurology.org/subscribers/advertise