Selection Approach to Identify the Optimal Biomarker Using Quantitative Muscle MRI and Functional Assessments in Becker Muscular Dystrophy

This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Neurology® Published Ahead of Print articles have been peer reviewed and accepted for publication. This manuscript will be published in its final form after copyediting, page composition, and review of proofs. Errors that could affect the content may be corrected during these processes.
Author(s):
Nienke M. van de Velde, MD\textsuperscript{1,6}; Melissa T. Hooijmans, PhD\textsuperscript{2}; Aashley S.D. Sardjoe Mishre, MSc\textsuperscript{2}; Kevin R. Keene, MD\textsuperscript{1,2}; Zaida Koeks, MD\textsuperscript{1}; Thom T.J. Veeger, MSc\textsuperscript{2}; Iris Alleman\textsuperscript{3}; Erik W. van Zwet, PhD\textsuperscript{1}; Jan-Willem M. Beenakker, PhD\textsuperscript{2,5}; Jan J.G.M. Verschuuren, MD PhD\textsuperscript{1,6}; Hermien E. Kan, PhD\textsuperscript{2,6}; Erik H. Niks, MD PhD\textsuperscript{1,6}

Corresponding Author:
Erik H. Niks
e.h.niks@lumc.nl

Affiliation Information for All Authors: 1. Department of Neurology, Leiden University Medical Center, the Netherlands; 2. C.J. Gorter Center for High Field MRI, Department of Radiology, Leiden University Medical Center, the Netherlands; 3. Department of Orthopaedics, Rehabilitation and Physical Therapy, Leiden University Medical Center, the Netherlands; 4. Department of Biomedical Data Sciences, Leiden University Medical Center, the Netherlands; 5. Department of Ophthalmology, Leiden University Medical Center, the Netherlands; 6. Duchenne Center Netherlands, the Netherlands

Contributions:
Nienke M. van de Velde: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Melissa T. Hooijmans: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Analysis or interpretation of data
Aashley S.D. Sardjoe Mishre: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Kevin R. Keene: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Zaida Koeks: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Thom T.J. Veeger: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Iris Alleman: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Erik W. van Zwet: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Jan-Willem M. Beenakker: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Jan J.G.M. Verschuuren: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design
Hermien E. Kan: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design; Analysis or interpretation of data
Erik H. Niks: Drafting/revision of the manuscript for content, including medical writing for content;
Study concept or design; Analysis or interpretation of data

Number of characters in title: 94

Abstract Word count: 250

Word count of main text: 3866

References: 41

Figures: 5

Tables: 2

Supplemental: Strobe checklist

Statistical Analysis performed by: Nienke M. van de Velde, MD, department of Neurology  Erik W. van Zwet, PhD, department of Biomedical Data Sciences

Search Terms: [ 120 ] MRI, [ 185 ] Muscle disease

Acknowledgements: Several of the authors are members of the European Reference Network for Rare Neuromuscular Diseases [ERN EURO-NMD].

Study Funding: this study was funded by the Netherlands Organization for Health Research and Development (113302001).

Disclosures: N. M. van de Velde reports no relevant disclosures; M. T. Hooijmans reports grants from Netherlands Organization for Health Research and Development, grant 113302001 during the conduct of the study; A.S.D. Sardjoe Mishre reports no relevant disclosures; K.R. Keene reports no relevant disclosures; Z. Koeks reports grants from Netherlands Organization for Health Research and Development, grant 113302001 during the conduct of the study; T.T.J. Veeger reports no relevant disclosures; I. Alleman reports no relevant disclosures; J. J.G.M. Verschuuren receives financial support Princes Beatrix Spierfonds; H.E. Kan reports grants from Netherlands Organization for Health Research and Development, grant 113302001 during the conduct of the study, consultancy for PTC therapeutics and Esperare, and trial support from ImagingDMD-UF outside the submitted work; E.H. Niks reports ad hoc consultancies for BioMarin, Summit, WAVE, Roche, NS Pharma, Reveragen and Sarepta outside the submitted work. All reimbursements from received by the LUMC. No personal financial benefits were received.

Abstract
**Objective:** To identify the best quantitative fat-water MRI biomarker for disease progression of leg muscles in Becker muscular dystrophy (BMD) by applying a stepwise approach based on standardized response mean (SRM) over 24 months, correlations with baseline ambulatory tests and reproducibility.

**Methods:** Dixon fat-water imaging was performed at baseline (n=24) and 24 months (n=20). Fat fractions (FF) were calculated for three center slices and the whole muscles for 19 muscles and six muscle groups. Contractile cross sectional area (cCSA) was obtained from the center slice. Functional assessments included knee extension and flexion force, and three ambulatory tests (North Star Ambulatory Assessment (NSAA), 10-meter run, six-minute walking test). MR parameters were selected using SRM ($\geq 0.8$) and correlation with all ambulatory tests ($\rho \leq -0.8$). Parameters were evaluated based on intraclass correlation coefficient (ICC) and standard deviation (SD) of the difference. Sample sizes (SS) were calculated assuming 50% reduction in disease progression over 24 months in a clinical trial with 1:1 randomization.

**Results:** Median whole muscle FF increased between 0.2-2.6% without consistent cCSA changes. High SRMs and strong functional correlations were found for eight FF but no cCSA parameters. All parameters showed excellent ICC ($\geq 0.999$) and similar SD of the inter-rater difference. Whole thigh three center slices FF was the best biomarker (SRM=1.04, correlations $\rho \leq -0.81$, ICC=1.00, SD=0.23%, SS=59) based on low SD and acquisition and analysis time.

**Conclusion:** In BMD, median FF of all muscles increased over 24 months. Whole thigh three center slices FF reduced the SS by approximately 40% compared to NSAA.
**Introduction**

Becker muscular dystrophy (BMD) is characterized by progressive muscle weakness due to the reduced production of a truncated dystrophin protein\(^1\). Disease progression is usually slow and highly variable\(^2\). This complicates design of clinical trials and emphasizes the development of objective and sensitive biomarkers.

Muscle fat fraction (FF) can be quantified reliably using Dixon chemical-shift magnetic resonance imaging (MRI) or MR spectroscopy (MRS), and is considered a promising biomarker in muscular dystrophies\(^3, 4\). In Duchenne muscular dystrophy (DMD), FF increases within 12 months\(^5, 6\), predicts change in function and clinical milestones\(^7, 8\), and reduces sample sizes needed to detect a treatment effect compared to functional outcome measures\(^5, 9, 10\). Quantitative muscle MRI also proved more sensitive to detect changes than functional measures in other, slowly progressive, neuromuscular diseases like FKRP related limb girdle muscular dystrophy R9, GNE myopathy, and Charcot Marie Tooth 1A\(^11-13\).

In BMD, longitudinal MRI data are limited. In a recent one year follow-up study in 16 patients no significant increase in FF was found\(^14\). Although FF correlated to functional outcomes in cross-sectional studies\(^15-18\), there is a large variability in FF between muscles. Quantitative MRI yields a large number of acquisition and analysis possibilities, including for instance acquisition location, size of the field of view, and number of included muscles. In this study, we aimed to assessed which muscle MRI parameter (FF or contractile cross sectional area (cCSA)), and which muscle or muscle group, was the most suitable biomarker to reflect disease progression in BMD.
Methods

Study participants

BMD patients were recruited from the Dutch Dystrophinopathy Database in a prospective longitudinal observational study conducted at the Leiden University Medical Center(19). Inclusion criteria were male, age ≥ 18 years, DNA confirmed diagnosis (in-frame mutation) of BMD and/or supported by clinical phenotype (ambulant >16 years without steroid treatment), and no contra-indications to MRI examination. Sex and age matched healthy controls (HCs) were recruited using flyers and advertisements.

Standard Protocol Approvals, Registrations, and Patient Consents.

The study was approved by the local medical ethical committee. All patients signed written informed consent.

Study design and functional tests

MR imaging and clinical assessments were performed at baseline and two years follow-up in BMD patients and only at baseline in HC. Clinical assessments consisted of the North Star Ambulatory Assessment (NSAA, amount of points), 10-meter run/walk test velocity (TMRv, in meters/second), Six minute walking test (6MWT) and quantitative muscle assessment (QMA) based strength of knee extensors (KE, in kilograms (kg)) and knee flexors (KF, in kg) on both sides within one week after the MRI by trained evaluators as described previously(20-22).
**MRI acquisition**

Thigh and lower leg muscles were examined at 3T (Philips Ingenia, Best, the Netherlands) using a 16-element receive coil placed on the anterior section of the leg and the twelve-element coil located inside the patient table. Position of the subjects was supine, feet first. A 3-point Dixon sequence (23 transverse slices, voxel size 1x1x10mm, slice gap 5mm, repetition time/echo time/echo time shift 210/4.41/0.76ms, flip angle 8°, number of signal averages 2) was performed. Field of view (FOV) in the lower leg was 180x180mm. FOV in the thigh ranged between 180x180mm and 220x220mm depending on the size of the thigh of the participant. Slice stacks were oriented perpendicular to the femur in the thigh and perpendicular to the tibia in the lower leg and centered at mid-thigh level and at the thickest part of the calf.

**MRI analysis**

MRI data were analyzed as described before (23). We created FF maps by dividing the signal intensity of the fat image by the signal intensity of the fat plus the water image. A single observer (NMV) manually drew Regions of interest (ROIs) on the boundaries of the muscles for a total of twelve thigh (Figure 1) and seven lower leg muscles on the water images of each slice using Medical Image Processing Analysis and Visualization (MIPAV) software (https://mipav.cit.nih.gov/). Images with clear movement or reconstruction artefacts were excluded. To test reproducibility, a second observer (KRK) drew ROIs in a subset of patients. These included three randomly selected participants with low FF (<20% fat), mid (20-65%), and high FF (>65%) based on the average FF of all thigh muscles. We determined reproducibility of the ROI analysis using intraclass correlation coefficient (ICC) with a 2-way
random effects model and absolute agreement, and using the standard deviation (SD) of the inter-rater difference between the two observers.

Quantification of muscle parameters

We used the ROIs to obtain three parameters: FF per slice and per muscle, and cCSA. FFs were calculated per slice and per muscle by transposing the ROIs onto the FF map. An in-house developed Matlab script (MathWorks, Natrick, MA) was used to erode the ROIs by two reconstruction voxels, and to correct for the chemical shift displacement to avoid contamination with subcutaneous fat. cCSA was calculated in the center slice as a measure of the fat-free area using the following formula: \( cCSA = CSA \times (1 - FF) \). This center slice was determined in consensus by three observers (NMV, KRK and HEK) and positioned using internal landmarks. The center slice was located at the slice containing the biceps femoris short head (BFS) insertion in the thigh and at the slice containing the flexor digitorum longus insertion in the lower leg(24). The three center slices were defined by the center slice plus one proximal and distal slice. We normalized the FFs for muscle area to obtain weighted FFs over two regions; three center slices and the whole depicted muscle. The same number of slices was used for the whole muscle analysis at baseline and follow-up. Finally, we combined outcomes of individual muscles into muscle groups to obtain a further set of MR parameters: whole thigh and whole lower leg, quadriceps (vastus lateralis VL, intermedius VI, medialis VM and rectus femoris RF), hamstrings (BFS, BFL, semitendinosus ST and semimembranosus SM), adductors (adductor longus AL and magnus AM, sartorius SAR and gracilis GR) and triceps surae (soleus SOL, gastrocnemius medialis GCM and lateralis GCL). This resulted in a total of 71 MR parameters: three outcomes per muscle (FF of the whole muscle, FF of three center slices, and cCSA) times 19 individual muscles and six muscle...
groups, minus cCSA and FF of three center slices for AL and BFS. We did not use these two muscles were for the analyses due to the choice of the landmark location at the BFS insertion. For readability, FF are expressed as percentage.

*Stepwise biomarker selection*

We used the acquired quantitative MRI parameters to identify the most optimal MR biomarker for disease progression using a stepwise approach. The first step was based on the responsiveness to disease progression of each parameter, measured by the standardized response mean (SRM), calculated as the mean change/SD of the change (Figure 2). We only included parameters with a SRM of \( \geq 0.8 \) in the next step(25). We calculated corresponding sample sizes for a potential clinical trial for each parameter as primary outcome measure using Lehr’s formula according to the method described by Morrow et al(11). In this calculation, we assumed a 50% reduction in disease progression over 24 months, a power of 80% and an \( \alpha<0.05 \) in a 1:1 randomization. In the second step, only those parameters which had a strong negative correlation of \( \leq -0.8 \) between baseline FF and all baseline functional tests (NSAA, TMRv and 6MWT) were included. In the third and final step, we ranked the parameters using the ICC and the SD of the difference between the two observers. The parameter with the highest reproducibility in this final step was considered the best biomarker.

We assessed the effect of a different center slice selection on measuring disease progression by shifting the center slice of the follow-up MRI one slice (corresponding to 1.5 centimeters) proximal or distal compared to the original center slice.
Statistical analysis

Data are presented as median (range) unless otherwise stated. Wilcoxon signed-rank tests were used to assess changes in functional tests and quantitative muscle parameters between baseline and follow-up. Spearman’s correlation coefficient was used to correlate functional and MRI parameters at baseline and change over two years. We only correlated the delta FF of the three muscle parameters with the lowest SD of the inter-rater difference in the stepwise biomarker selection to change in function. The current study was complementary to a second BMD longitudinal observational study that was performed in parallel at our institution(26), and some patients participated in both studies. As in a few patients the functional tests could not be performed within one week of the MRI scan at baseline, multiple imputation (five times) was performed for these functional tests using the available data from the parallel study (3-5 time points per patient, one per year). We also calculated SRMs and SS for the functional tests as described above. All statistical tests were performed with IBM SPSS Statistics version 25. Statistical significance was set at p≤0.05.

Data availability

Anonymized data can be made available to qualified investigators on request.

Results

Cohort characteristics

24 BMD patients (median age 41.3 years (18.8–66.3) and 13 HCs (median age 43.3 years (21.3–63.6)) were included at baseline. All patients were ambulant (defined as being able to walk 10 meters with support of a cane). After imputation (TMRv n=4, NSAA n=5, 6MWT n=2), baseline functional assessments were available for 23 out of 24 patients. Functional
testing could not be performed in one patient due to known muscle cramps and pain following exercise.

Two patients did not complete follow up (one deceased and one no show). Of the 22 patients who participated in the follow-up visit after a mean of 1.95±0.23 years, one patient did not perform functional tests again due to muscle cramps. In two patients, MRI data were not available at follow up due to a contraindication to MR-scanning (ICD implantation) and a protocol deviation in which the opposite leg was scanned. Thus, longitudinal functional and MRI measurements were available for 21 and 20 patients respectively.

Functional assessments

At baseline, all functional tests showed large variability between patients (Table 1). One patient lost ambulation during the study. Both the NSAA (-2.5 points, p=0.002) and TMRv (-0.22 m/s, p=0.014) declined significantly in two years, whereas the 6MWT did not. The SRM of the NSAA was higher (-0.81) than those of TMRv (-0.68) and 6MWT (-0.46). This resulted in a sample size of 98 for the NSAA, 138 for the TMRv and 310 for the 6MWT. Of the QMA measures, only the strength of the knee flexors decreased significantly in 24 months (median -1.4kg, p=0.04, SRM=-0.71, SS 126).

Fat fraction and cCSA at baseline and follow-up

Visual inspection showed that fat replacement was non-uniform over the proximodistal axis in the majority of individual muscles, and that the pattern differed between muscles. For example, the FF of the VL seemed higher near the proximal and distal end of the muscle while in the ST the FF appeared higher near the distal end and was lower towards the
proximal muscle part (Figure 1). Interestingly, some muscles (e.g. the quadriceps muscles) also showed heterogeneous fat replacement in the axial plane in some patients (Figure 1b).

Baseline median whole thigh FF was 52.0% (4.4–72.9) and median whole lower leg FF was 25.5% (4.4–56.5). Values of individual muscles whole muscle median FFs were highly variable between muscles and between patients, starting from median FF 7.9% (3.8–14.2) in the TP to 68.1% (3.2–87.0) in the BFL.

Whole thigh and lower leg FF changes from baseline to follow-up were low in patients with baseline whole leg fat <20% (median FF change in the thigh of -0.06% (-0.7–1.8), and in the lower leg -0.07% (-1.0–0.8)). In patients with baseline 40-60% fat, and even in patients with >60% fat, the increase was relatively high (up to 5.4% in the thigh and 6.4% in the lower leg), see Figure 3. Overall, whole muscle FF increased significantly over 24 months in the thigh (median +1.9%, (-0.7–5.4), p=0.01) and lower leg (median +0.7%, (-1.0–6.4), p=0.02).

FF increases varied somewhat between individual muscles (Figure 4a, table e-1 [https://git.lumc.nl/neuroscience/2021_VeldeNMvande_Neurology_BMD_MRI]), ranging from a median increase of +0.2% (-1.9–10.7) in the GCL to + 2.6% (-2.3–2.6) in the BFS.

cCSA showed a large variability between individual muscles and patients at baseline (table e-1). In the thigh, the VL had the highest median cCSA (692.4mm$^2$ (246.7–3116.2), versus median cCSA 2991.5mm$^2$ (2390–3540) in controls), and in the lower leg the median cCSA of the SOL was highest (2078.8mm$^2$, (1354.6–3621.6), versus a median of 2266.3mm$^2$ (1343.4–3005.7) in controls). The cCSA change over 24 months in BMD patients was highly variable in and between muscles, ranging from a median decrease in the GCM of -95.5mm$^2$ (-516.4–181.6) to a median increase in the TP of 16.3mm$^2$ (-58.4–159.3) (Figure 4b and table e-1 [https://git.lumc.nl/neuroscience/2021_VeldeNMvande_Neurology_BMD_MRI]).
Selection of qMRI biomarkers

We included a total of 71 muscle MRI parameters in the selection procedure. For the first step, i.e. selection based on SRM, nine parameters had a SRM >= 0.8 (Figure 2). Only FF of individual muscles or muscle groups passed this criterion, but none of the muscles or muscle groups cCSA. The SRMs of three center slices FF of the whole thigh and quadriceps were highest (1.04, corresponding sample size of 59). In the second step, we found a correlation stronger than -0.8 between baseline FF and baseline tests of ambulatory function (NSAA, TMRv and 6MWT) in eight out of nine parameters. For individual muscles, these included FF of three center slices of the VL and the VI, and FF of the whole muscle of the VL and the RF. For muscle groups, this applied to the FF of both three center slices and whole muscle of the whole thigh and the quadriceps. The FF of three center slices of the VM showed a weaker correlation than -0.8 to the ambulatory tests. In the third and final step, all eight parameters showed an excellent ICC (Table 2). The FF of whole thigh three center slices had the lowest SD of the difference between the two observers (0.23%). SD of the difference of whole thigh whole muscle was about the same (0.24%). The FF of VI three center slices had the highest SD (1.55%). The values of all parameters of each step are given in table e-1. Definite ranking in the third step was not meaningful to the high ICCs and relatively small variation between SDs of the difference of the parameters.

Relation between change in function and fat fraction

Changes in NSAA correlated weakly to changes in FF of three center slices and whole muscle of the whole thigh (rho=-0.244, p=0.314 and rho=-0.222, p=0.361) and weakly to quadriceps whole muscle (rho=-0.341, p=0.152). None of these longitudinal correlations were significant. Correlations between changes in the TMRv and 6MWT versus changes FF
of the three muscle groups were all weak (-0.133 ≤ \( \rho \) ≤ -0.332) and not statistically significant (Figure 5).

Discussion

This prospective study describes longitudinal functional and quantitative muscle MRI data of BMD patients. Quantitative water-fat imaging demonstrated an increase in median fat of all individual muscles over 24 months, irrespective of the FF at baseline. After a stepwise biomarker selection approach using SRMs, correlation to baseline function and reproducibility, eight thigh FF parameters were selected as potential biomarkers for disease progression.

The conduction of clinical trials in muscular dystrophies has been hampered by the lack of objective biomarkers that are sensitive to disease progression, that are reproducible, and that show a relation with functional assessments. Trial development and conduction is even more complicated in BMD due to the low prevalence and the slow and heterogeneous disease severity and progression. FF has been suggested as a sensitive biomarker for disease progression in several other muscular dystrophies and can be quantified by MRS or fat-water imaging(5, 27, 28). Dixon fat-water MR imaging allows to determine FFs and cCSAs of not only a wide range of different muscles, but also parts of muscles or muscle groups. For example, studies in several neuromuscular diseases have used single slice, multi-slice (ranging from 3 up to 15 slices), and different sets of individual muscles or muscle groups(10-13, 29, 30) in their analyses. It is therefore not immediately clear which (part of) muscle or muscle group is most sensitive to detect disease progression. In our study, only thigh FF parameters had a sufficiently high SRM to be included in the second step of the biomarker selection, and no cCSA parameters. The variability of cCSA between patients is
apparently larger than the change that can be detected in this slowly progressive disease which is in line with findings in other neuromuscular diseases (11, 12). In contrast to FF of the thigh, none of the lower leg FF parameters had high SRMs. This may be explained by the BMD muscle pattern of involvement, where thigh muscles show weakness in early disease states, while lower leg muscles are spared until later (16, 18).

Our data from the second step in the biomarker selection confirmed the high correlation previously observed between muscle FF and clinical outcome measures at one time point in BMD (15-17), as eight out of nine thigh FF parameters passed this criterion. These eight MRI parameters all had higher SRM than the functional tests, including the NSAA. Translated to sample sizes for a hypothetical clinical trial, this would result in a reduction of 39 (compared to the NSAA) to 251 (compared to the 6MWT) patients per group to detect a 50% reduction of disease progression over 24 months for the two MRI parameters with the highest SRM in our study, i.e. the three center slices FF of whole thigh and quadriceps.

The excellent ICC in the third step is in line with previous findings that quantification of FF by qMRI is accurate and reproducible (8, 27, 31). We assessed the reproducibility of the FF parameters in even more detail by calculating the SD of the difference between two observers. Although low SD values did not allow a final ranking based on statistically significant differences, even small variabilities in SDs of the difference may obscure the relatively small increases in FF over time present in slowly progressive diseases, and can therefore influence the sensitivity to detect treatment effect in clinical trials. We found that muscle groups had lower between observer variability, compared to individual muscles, and hence showed more promise as biomarkers. This is probably due to easier determination of borders of muscle groups in highly fatty replaced muscles compared to delineating individual muscles (12). A recent study described higher SRMs for segmentation of whole muscle groups compared to individual muscles in several other chronic and progressive
neuromuscular diseases(32), supporting the results of our study. Using MRI parameters which are based on only a small part of the muscle has significant advantages over using a whole muscle. First, imaging whole muscles may lead to difficulties in obtaining sufficient scan quality at the edges of the field of view due to B0 and B1 artifacts. This could be solved by imaging the thigh in two rather than one slice stack, but leads to a duplication of scanning time. Second, as currently no software exists that can accurately segment (partly) fatty replaced muscles, all analyses are must be performed manually, which is very time consuming. Based on the small interrater differences and the advantages of imaging a small part of the muscle, whole thigh three center slices FF was the best biomarker in our cohort.

A potential disadvantage of using only three center slices could be the non-uniform fat replacement that we observed in the majority of thigh muscles. This variability has also been described in other muscular disorders(33-35), and could influence the repeatability when the center slice is not determined consistently. Indeed, our previous study in DMD demonstrated that shifting a slice stack of four center slices one slice proximal or distal resulted in a significant mean difference of 1-2% in FF, ranging up to 12%(33). In the present study, the difference between change in whole thigh three center slices FF of the original center slice versus one slice shift proximal and distal ranged between a minimum of -2.4% to a maximum 2.7% within the patients. This highlights that the center slice should be selected precisely and consistently in longitudinal follow-up by trained technicians, especially in multicenter studies.

We assessed reproducibility in detail for analysis of ROIs. Final implementation of a biomarker in clinical trials would also require assessment of other sources of variability, such as intra-site and inter-site variability in a multi-center setting. This has been addressed in healthy controls for a single vendor(36) and in patients with DMD across vendors(37). High measures of reproducibility were achieved in both studies. Thus, we believe that a
standardized image acquisition and analysis protocol, including instructions for center slice selection, could support the use of FF as biomarker in multicenter clinical trials. In such trials, it also remains important to assess system stability and variability between sites using a phantom, as was recently shown for liver fat fraction(38).

The results of our study confirm the insensitivity of functional tests as biomarker over shorter periods of time, as shown in several neuromuscular diseases(11, 27, 39). Although this substantiates the need for biomarker development, implementation of such biomarkers is only possible upon proving a clear association with a clinical meaningful endpoint. This interdependence has been recently shown in DMD(7, 8), but should still be demonstrated with longer follow-up in BMD. Furthermore, other imaging biomarkers, such as the global T2 relaxation time, have also been shown to be sensitive to disease progression and predictive of function in DMD(7, 40). However, this global T2 relaxation time (or MRI T2) is dominated by the fat signal, and therefore largely reflects the changes in fat replacement that we assessed here using the Dixon technique. By contrast, the water T2 relaxation time is thought to reflect edema or inflammation, and has been shown to be elevated early in the disease process in DMD(41). We recently showed no differences in water T2 between BMD and healthy controls(23) and have therefore not included this in the current analysis.

This study has limitations. Some functional tests (TMRv n=4, NSAA n=5, 6MWT n=2) were performed more than one week after baseline MR examination. These values were therefore imputed to assess the decline in functional tests over two years, and to enable the direct comparison of SRMs of functional tests and MR parameters. The cohort was also relatively small, and definite ranking based on reproducibility was not possible. Our results should therefore be confirmed in another preferably larger cohort. Finally, only ambulant patients were included in our study, although this was not an inclusion criterium. Results may therefore not be extrapolated to BMD patients in later disease stages.
In conclusion, FF of whole thigh three center slices was the most optimal biomarker based on high SRM, low SD of the difference, and practical considerations in this ambulant and adult BMD population. This biomarker potentially lowered the sample size of a clinical trial over 24 months by 39 patients compared to the NSAA, which is equivalent to a reduction of approximately 40%. The results support the use of FF quantified by qMRI as biomarker in clinical trials in slowly progressive and heterogeneous diseases like BMD.
## Appendix 1. Authors

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nienke M. van de Velde, MD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Analyzed the data; conducted the statistical analysis; interpreted the data; drafted the manuscript for intellectual content</td>
</tr>
<tr>
<td>Melissa T. Hooijmans, PhD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Acquisition of the data, analyzed the data; revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>Aashley S.D. Sardjoe Mishre, MSc</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Analyzed the data; revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>Kevin R. Keene, MD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Analyzed the data; revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>Name</td>
<td>Institution</td>
<td>Role in acquisition of data; revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Zaïda Koeks, MD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td></td>
</tr>
<tr>
<td>Thom T.J. Veeger, MSc</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Analyzed the data; revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>Iris Alleman</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Role in acquisition of data</td>
</tr>
<tr>
<td>Erik W. van Zwet, PhD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Conducted the statistical analysis, revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>Jan-Willem</td>
<td>Leiden</td>
<td>Interpreted the data; revised the</td>
</tr>
<tr>
<td>Name</td>
<td>Institution</td>
<td>Contribution</td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------------------------------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>M. Beenakker, PhD</td>
<td>University Medical Centre, The Netherlands</td>
<td>manuscript for intellectual content</td>
</tr>
<tr>
<td>Jan J.G.M. Verschuuren, MD PhD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Designed and conceptualized study; interpreted the data; revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>Hermien E. Kan, PhD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Designed and conceptualized study; interpreted the data; revised the manuscript for intellectual content</td>
</tr>
<tr>
<td>Erik H. Niks, MD PhD</td>
<td>Leiden University Medical Centre, The Netherlands</td>
<td>Designed and conceptualized study; interpreted the data; revised the manuscript for intellectual content</td>
</tr>
</tbody>
</table>

References


**Figure 1: Heterogenous fat distribution in vastus lateralis and semitendinosus.** Example of regions of interest (B) and of heterogenous fat distribution from distal (A) to proximal (C) in the vastus lateralis (A-D) and semitendinosus (A-C, E). Figure B also shows an example of heterogenous fat distribution in the axial plane of the quadriceps. To enhance visualization of the fat differences along the proximodistal axis of the muscles, patients are grouped according to their weighted fat of the whole muscle (light gray: <20% fat, dark grey: 20-65% fat, black: >65% fat). The muscles are aligned based on the insertion of the BFS (marked as ‘center’) in figure D and E. Distal and proximal muscle parts are depicted in the left and right part of the figure respectively. Dashed lines: healthy controls, continuous lines: Becker muscular dystrophy patients. Each line represents one individual subject. AM/AL=adductor magnus/longus; BFS/BFL=biceps femoris short/long head; EDL=extensor digitorum longus; G=gracilis; GL/GM=gastrocnemius lateralis/medialis; PR=peroneus; RF=rectus femoris; SR=sartorius; SM=semitendinosus; SOL=soleus ST=semitendinosus; TA/TP=tibialis anterior/posterior; VL/VI/VM=vastus lateralis/intermedius/medialis
Figure 2: Flow chart of the stepwise process for selection of MR parameters. FF: fat fraction, 3C slices: three center slices, WM: whole muscle, cCSA: contractile cross sectional area, SRM: Standardized response mean, *Correlation with baseline North Star Ambulatory Assessment, ten meter run test and 6 Minute Walk Test.

Individual muscles (n = 19) 
Muscle groups (n = 6) 
FF of 3 center slices 
FF whole muscle 
cCSA

Potential parameters (n = 71)

SRM ≥ 0.8

No: Parameters (n = 62)
FF 3C slices: 18 out of 23 parameters
FF WM: 21 out of 25 parameters

cCSA: 23 out of 23 parameters

Yes: Parameters (n = 9)

Correlation with baseline function ≤ 0.8*

No: Parameter (n = 1) (FF vastus medialis 3 center slices)

Yes: Parameters (n = 8)

Parameters selected for ranking:
- FF whole thigh 3 center slices
- FF whole thigh WM
- FF quadriceps WM
- FF quadriceps 3 center slices
- FF vastus lateralis 3 center slices
- FF vastus lateralis WM
- FF rectus femoris WM
- FF vastus intermedius 3 center slices
Figure 3. Mean change in fat in thigh and lower leg whole muscles per patient. Each patient is represented by a black and gray bar stitched together. Black and gray bars indicate mean change over thigh and lower leg muscles respectively. *indicates a mean change of zero. Patients are subdivided in four groups based on their average weighted fat over all thigh and lower leg muscles (<20%, 20-40%, 40-60% and >60% fat).
Figure 4. Change in fat (A) and contractile cross sectional area (B) over 24 months.

Boxes show median (IQR). The lines represent the 10-90\textsuperscript{th} percentile with diamonds indicating values outside this window. Median baseline values are given at the left border.

Figure 5. Relation between change in function tests and fat fraction. Change in whole thigh FF of three center slices versus NSAA (a) and TMRv (b). Closed arrows: patients with baseline and follow-up measurements. Open triangles are patients with only one measurement (baseline). Horizontal arrows indicate functionally stable patients but with a change in FF. NSAA: North Star Ambulatory Assessment, TMRv: ten meter run test velocity.
Table 1. Change in functional assessments between baseline and after 24 months

<table>
<thead>
<tr>
<th>Test</th>
<th>Median at Baseline</th>
<th>Median change follow-up versus baseline (range)</th>
<th>P value</th>
<th>SRM</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSAA (points)</td>
<td>18 (5-34)</td>
<td>-2.5 (-12.0 – 1.0)</td>
<td>0.002</td>
<td>-0.81</td>
<td>98</td>
</tr>
<tr>
<td>TMRv (m/s)</td>
<td>1.45 (0.26-4.17)</td>
<td>-0.22 (-1.4 – 0.25)</td>
<td>0.014</td>
<td>-0.68</td>
<td>138</td>
</tr>
<tr>
<td>6MWT (m)</td>
<td>385 (0-650)</td>
<td>-12.6 (-151.9 – 33.0)</td>
<td>0.063</td>
<td>-0.46</td>
<td>310</td>
</tr>
<tr>
<td>KE (kg)</td>
<td>8.56 (2.9-54.5)</td>
<td>-1.3 (-11.1 – 3.8)</td>
<td>0.114</td>
<td>-0.49</td>
<td>264</td>
</tr>
<tr>
<td>KF (kg)</td>
<td>8.19 (2.4-29.7)</td>
<td>-1.4 (-7.1 – 2.8)</td>
<td>0.040</td>
<td>-0.71</td>
<td>126</td>
</tr>
</tbody>
</table>

Table 2. qMRI parameters in final step of the flowchart

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SRM</th>
<th>SS</th>
<th>Correlation to baseline function</th>
<th>Reproducibility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>NSAA</td>
<td>TMRv</td>
</tr>
<tr>
<td>Whole thigh 3CS</td>
<td>1.04</td>
<td>59</td>
<td>-0.888</td>
<td>-0.865</td>
</tr>
<tr>
<td>Whole thigh WM</td>
<td>1.01</td>
<td>64</td>
<td>-0.924</td>
<td>-0.891</td>
</tr>
<tr>
<td>Quadriceps WM</td>
<td>0.99</td>
<td>65</td>
<td>-0.878</td>
<td>-0.842</td>
</tr>
<tr>
<td>Quadriceps 3CS</td>
<td>1.04</td>
<td>59</td>
<td>-0.878</td>
<td>-0.842</td>
</tr>
<tr>
<td>Vastus lateralis 3CS</td>
<td>0.83</td>
<td>94</td>
<td>-0.866</td>
<td>-0.832</td>
</tr>
<tr>
<td>Vastus lateralis WM</td>
<td>0.92</td>
<td>76</td>
<td>-0.858</td>
<td>-0.818</td>
</tr>
<tr>
<td>Rectus femoris WM</td>
<td>0.84</td>
<td>92</td>
<td>-0.896</td>
<td>-0.877</td>
</tr>
<tr>
<td>Vastus intermedius 3CS</td>
<td>0.85</td>
<td>90</td>
<td>-0.874</td>
<td>-0.849</td>
</tr>
</tbody>
</table>

Selection Approach to Identify the Optimal Biomarker Using Quantitative Muscle MRI and Functional Assessments in Becker Muscular Dystrophy
Nienke M. van de Velde, Melissa T. Hooijmans, Aashley S.D. Sardjoe Mishre, et al.
Neurology published online June 23, 2021
DOI 10.1212/WNL.0000000000012233

This information is current as of June 23, 2021