Outcome measures in Duchenne muscular dystrophy

Naarding, K.J.

Karin J. Naarding

Outcome measures in Duchenne muscular dystrophy

Towards quantifying clinically relevant differences

Karin J. Naarding

Outcome measures in Duchenne muscular dystrophy

Towards quantifying clinically relevant differences

ISBN: 978-94-6469-207-5

Cover design & lay-out: Wendy Schoneveld || www.wenziD.nl
Printed by: Proefschriftmaken || Proefschriftmaken.nl

The research described in this thesis was performed at Leiden University Medical Center and was mostly supported by a grant from Stichting Spieren for Spieren (grant number SvS15). Other supporting grants were from the Netherlands Organization for Health Research and Development (ZonMw; grant 113302001), Prosensa Therapeutics B.V., FP7-HEALTH-2013-INNOVATION-1, Duchenne Parent Project Netherlands (DPP NL) and the Gratama Foundation (grant number 10.13).

Outcome measures in Duchenne muscular dystrophy
Towards quantifying clinically relevant differences

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 16 maart 2023
klokke 10.00 uur

door
Karin Joanna Naarding

geboren te ‘s-Gravenzande
in 1990

Promotor
Prof. dr. J.J.G.M. Verschuuren

Copromotores
Dr. E.H. Niks
Dr. H.E. Kan

Promotiecommissie
Prof. dr. J.G. van Dijk
Prof. dr. N. Geijsen
Dr. I.J.M. de Groot
(Radboud UMC)
Prof. dr. ir. G.J. Strijkers
(Amsterdam UMC)

Opgedragen aan Erik Vrielink,
buurman en inspirator

Contents

Chapter 1

General introduction

Duchenne muscular dystrophy

Clinical course

Duchenne muscular dystrophy (DMD) is a fatal and rare muscle wasting disease with an incidence of approximately 1 in 5000 new-born boys.[1, 2] The most frequently reported symptoms before diagnosis are gross motor delay, muscle weakness, difficulty walking, running and stair climbing, and frequent falls, whereas a proportion of patients show delay in cognitive and language development.[3] Most boys without a family history are diagnosed before five years of age.[4] Children with DMD suffer from a proximal to distal gradient of muscle weakness. The ambulatory phase of the disease can be divided into four clinical stages: early ambulatory, late ambulatory, early non-ambulatory and late-non-ambulatory.[4] Clinical signs in the early ambulatory stage that can be observed are: delays in achieving developmental milestones, difficulties with running and jumping, a Gowers' sign, frequent falls and keeping up with peers regarding gross motor functions. The late ambulatory stage is characterized by a significantly reduced walking speed, fatigue and pain after walking long distances, the increased use of a wheelchair, and difficulties with rising from floor and stair climbing. Loss of independent ambulation defines the transition to the early non-ambulatory stage. Median age at loss of independent ambulation shifted from ten to 13 years of age due to the use of glucocorticoids, although the age range remains wide.[5, 6] Transition to late non-ambulatory stage is less clearly defined. Patients in the late non-ambulatory stage progressively require assistive devices to function independently, such as remote control units to operate electronic devices including televisions, computers and lights. For upper extremity function, arm muscle strength already decreases in the ambulant phase.[7] In the early non-ambulatory stage patients increasingly experience difficulties raising the arms due to loss of shoulder strength. Upper arm function, such as the ability to move the hand to the mouth, is preserved until the mid-teens.[5] In the late non-ambulatory phase patients have limited arm and hand function left. Hand function is preserved into the twenties, although hand strength of DMD patients has been found to be lower than that of healthy peers as early as five years of age.[5, 8] Preservation of minimum function of hand muscles can significantly improve participation in daily life for patients in this stage, because it could allow them to use electronic devices such as an electric wheelchair, smartphone, tablet, computer or a game console.[9]

Not only motor functioning is affected by the absence of dystrophin, but clinical manifestations can also be observed in for instance the heart and the brain. Examples of the cognitive manifestations are a higher prevalence of learning and behavioral disabilities in DMD compared to the general population.[10]

Pathophysiology

DMD is an X-linked inherited neuromuscular disorder caused by mutations in the *DMD* gene located at the short arm of the X chromosome (Xp21 locus). The *DMD* gene consists of 79 exons that together encode the dystrophin protein. The mutations causing DMD lead to

premature termination of dystrophin production and thereby nearly complete absence of the full-length dystrophin protein.[11]

Becker muscular dystrophy (BMD) is also caused by mutations in the *DMD* gene, but in general these mutations do not lead to absence of dystrophin, but to a partly functional dystrophin protein with an altered molecular weight. BMD patients have a more variable and generally milder disease course.[12, 13]

The full-length dystrophin protein is expressed in skeletal muscle, where it stabilizes the muscle fiber membrane and protects it from contraction induced damage.[14] It also has a function as signaling complex in skeletal muscle by providing binding sites for signaling proteins such as nitric oxide synthase.[14] In DMD, disruption of these functions is assumed to cause muscle fibers to be easily damaged, which leads to fiber degeneration and regeneration, inflammation and finally muscle wasting with irreversible replacement of muscle fibers with fat and fibrotic tissue (Figure 1).[15]

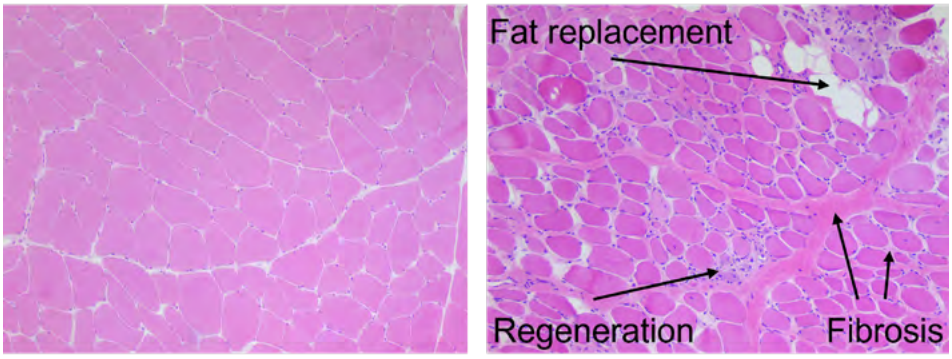


Figure 1. Biopsy of dystrophic muscle
Muscle biopsy of an anonymous healthy control (left) and DMD patient (right) from the Leiden University Medical Center. The dystrophic biopsy shows variation in size and shape of the muscle fibers with regeneration, some necrosis, inflammation, fibrosis and fatty replacement. *Kindly provided by Dr. S.G. van Duinen.*

Dystrophin proteins with different lengths (i.e. isoforms) are encoded by the *DMD* gene. The full-length dystrophin protein is the primary isoform in skeletal muscle, but at least three dystrophin isoforms are also expressed in the brain: full-length Dp427, and the shorter Dp140 and Dp71 (Figure 2).[16] The location of the mutation within the *DMD* gene influences the number of dystrophin isoforms that are lacking. In Figure 2, a mutation at location A (exon 1 until 44) will only lead to absence of Dp427. In case of a mutation at location B (exon 51 until 62), Dp140 will be lacking in addition to Dp427. Finally, a mutation at location C (exon 63 until 79) will lead to absence of all three dystrophin isoforms. DMD patients lacking the Dp140 isoform have been demonstrated to on average perform more poorly on neuropsychological tests and have a higher incidence of learning and behavioral disabilities.[17, 18] These seem even more pronounced in patients lacking all three brain isoforms.[19, 20]

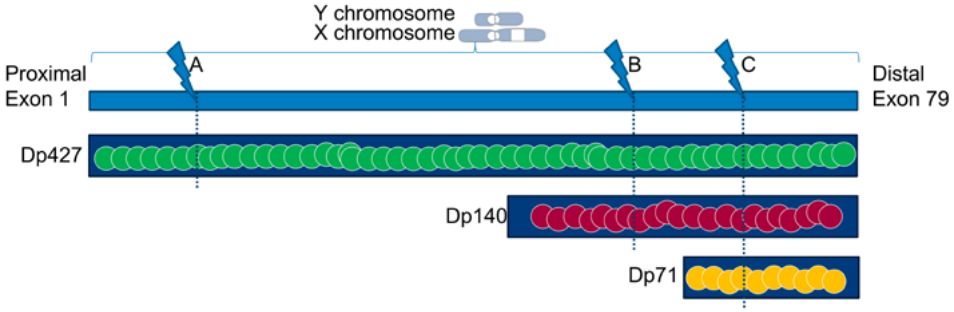


Figure 2. Mutation locations of dystrophin isoforms
Image representing mutation locations in the *DMD* gene and their effect on the different dystrophin isoform expression. The full-length dystrophin Dp427 and the shorter Dp140 and Dp71 isoforms are present in the healthy brain. A mutation at location A, in exon 1 until 44, will only lead to absence of Dp427. Dp140 will be lacking in addition to Dp427 with a mutation at location B, in exon 51 until 62. A mutation at location C, which is in exon 63 until 79, will lead to absence of all three dystrophin isoforms. *Modified with permission from Dr. N. Doorenweerd.*

Clinical trials in DMD and the importance of outcome measures

The International standards of care for DMD have been published in 2010 and updated in 2018. Guidelines on the implementation of these standards of care in the Dutch healthcare system can be found on the website of the Duchenne Center Netherlands: www.duchenneexpertisecentrum.nl/passende-zorg/zorgverleners/duchenne-richtlijn/.[4, 21, 22] Treatment with glucocorticoids is recommended and has been shown to delay loss of ambulation and upper limb disease progression, to reduce the need for scoliosis surgery, to improve pulmonary function, and to delay cardiomyopathy onset.[5, 23] Life expectancy has shifted to the late twenties and thirties due to the combination of improved cardiac care, orthopedic interventions, the use of glucocorticoids and, most importantly, advancements in respiratory care such as air stacking, use of cough assist devices, and assisted ventilation.[24, 25] Nonetheless, more severely affected patients can still die in the late teens or early twenties. A fully approved cure for DMD is currently lacking, but three dystrophin-restoring drugs have received conditional approval, two by the U.S. Food and Drug Administration (FDA) and another one by the European Medicines Agency (EMA).[26, 27] Unfortunately, these drugs only seem to have a limited effect on disease progression.[26]

Over the past decade, many clinical trials in DMD have been conducted in primarily ambulant patients to study effectiveness of therapies aiming at dystrophin restoration or improvement of muscle quality, and many trials are still being conducted.[26] Assessment of the safety and efficacy of a new therapy follows a standardized order. After preclinical studies, a new drug is studied in humans in four sequential clinical trial phases. In phase I, the optimal dosage is determined based on safety in healthy volunteers. In phase II, efficacy and side effects are assessed in a small group of patients. In DMD, phase I and II are often combined and

called phase II, because of potential harm in healthy volunteers when RNA modifying drugs are used, such as antisense oligonucleotides designed to alter RNA splicing. Phase II is sometimes split in IIa and IIb, where phase IIa is specifically designed to assess the required dose and phase IIb to study efficacy in a small group of patients. Phase III is used to assess efficacy, effectiveness and safety in a larger patient population, that is predefined using a power calculation based on the effect in the phase II trials, compared to a placebo. In such randomized controlled trials (RCT), patients are randomized into the study drug arm or comparative arm, and both patients and study personnel, including clinicians and clinical evaluators, are blinded for the received treatment. It is the last phase before approval is requested from the regulators. Phase IV consists of post-marketing surveillance.

Phase III trials have a number of important characteristics. The main purpose is to demonstrate the effect of the intervention on one outcome measure, the primary endpoint. Secondary endpoints may contribute supportive information about the effectiveness of a drug, but they are less important for drug approval, because the sample size was usually not determined to find an effect on secondary endpoints. Approval of the regulatory agencies to use an outcome measure as primary endpoint depends on the natural history of the disease, the disease phase that is studied, the availability of outcome measures, and the expected effect and mechanism of action of the intervention.[28]

When the first phase IIb RCT in DMD was conducted using ataluren as stop codon readthrough in 2008, the six minute walking test (6MWT) was available from pulmonary and metabolic disease studies and was acceptable for regulators as primary endpoint in DMD.[29-31] However, no natural history data were available at that time, and it was not until the placebo arm of this trial was analyzed that it became clear that the total distance walked in the 6MWT by DMD patients first improved with age due to maturation, then stabilized and finally declined due to disease progression until loss of ambulation (Figure 3).[32] Many natural history studies have since described that 6MWT distances vary largely between and within patients over time due the non-linear progression, different progression rates, and other factors such as interobserver variation and motivational issues (Figure 3).[32-35] The ataluren RCT demonstrated the hurdles that arise in a maturing population of ambulant DMD patients with a heterogeneous disease course, an outcome measure that has a motivational component and with drugs that potentially limit the progression rather than improve muscle strength, and thus have a limited effect. This initiated research into the natural history and the development of new outcome measures.[28] As a consequence, following trials narrowed inclusion criteria to select a more homogenous study population, which complicated study inclusion in this rare disease. Furthermore, different primary endpoints were selected that potentially had a linear progression and were more sensitive to change, such as the North Star Ambulatory Assessment (NSAA), the time to climb four stairs and the time to rise from the floor (examples of trials using these as primary endpoint: NCT04281485, NCT04632940, NCT02851797, NCT04587908 and NCT03439670). [28, 36, 37] The NSAA is an assessment of ambulatory motor performance that consists of 17 items and yields a maximum score of 34 points.[37]

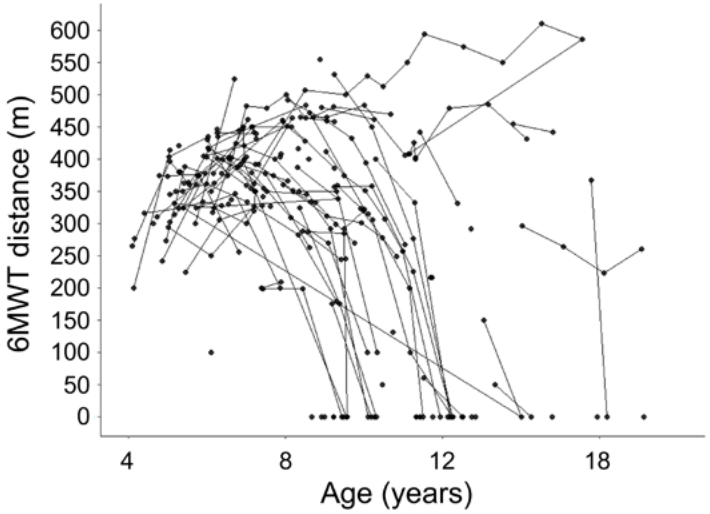


Figure 3. Six minute walking test (6MWT) change over time data
Variability in 6MWT distance results of DMD patients from the Leiden University Medical Center outpatient clinic cohort. Age at loss of ambulation is plotted as 0 meters.

In the past, the FDA and EMA have ruled differently in the conditional approval phase of new treatments for DMD. The FDA ruled in favor of treatments that demonstrated an increase of dystrophin levels, which has a causal relation with the symptoms in DMD and therefore were anticipated to have a potential clinical benefit upon continued treatment.[38] The FDA conditionally approved eteplirsen based on an open label study that showed a minor increase in dystrophin levels of up to 0.9% after 180 weeks of treatment.[26, 39] The EMA ruled in favor of treatments that demonstrated a clinically relevant effect for patients, and expected this to originate from the intended effect of a therapy.[40] The EMA therefore conditionally approved ataluren based on a subgroup analysis that showed a 6MWT difference of 68.2 meters and a favorable side-effect profile, although the study did not meet the primary endpoint for the whole group.[26, 29] These differences in the FDA's and EMA's rulings thus led to conditional approval of one drug by the FDA that the EMA did not approve, and vice versa.[26]

There are several types of outcome measures: clinical outcome measures that require specific patient related tasks, such as the 6MWT or Performance of the Upper Limb, patient reported outcome measures, such as the DMD Upper Limb Patient Reported Outcome Measure (PROM), and biomarkers that for example reflect tissue characteristics, such as fat fraction measured using muscle magnetic resonance imaging (MRI), or dystrophin levels using muscle biopsies, or circulating muscle related micro RNA or proteins in blood or urine samples. Clinical relevance could be assumed for clinical outcome measures and patient

reported outcome measures that are relevant for a specific disease and disease stage, although the amount of change that is clinically relevant is topic of debate. For biomarkers clinical relevance is less clear. One method to study clinical relevance is by using the biomarker to predict loss of an important disease milestone, such as the ability to ambulate or to bring a glass to the mouth. Another method is to determine the minimally clinically important difference, for instance via the Delphi method by using a panel of experts, or via an anchor-based approach where the biomarker is linked to an independent measure with clinical relevance to patients, such as a global rating of change.[41, 42]
All outcome measures should demonstrate sufficient reliability, construct validity, concurrent validity, longitudinal change and accessibility. A measure is reliable, when repeated tests lead to similar results. Construct validity is the extent to which the test measures what it is intended to measure. Concurrent validity is the extent to which the studied outcome measure correlates with an established outcome measure. An outcome measures should be sensitive to longitudinal change in DMD, even more so because therapies so far have been expected to reach a stabilizing effect in this progressive disease. The sensitiveness to change of the outcome measure and the expected effect of a therapy together determine the number of participants that is required per clinical trial. Due to its variability between patients and within patients over time, the 6MWT required relatively large sample sizes.[33, 34] A more sensitive outcome measure or more effective drugs would obviously lead to a smaller required sample size and thus increase the likelihood of such a trial to be completed successfully.[43] Accessibility of outcome measures is important in DMD, because trials are usually conducted worldwide over a period of years due to the rarity of the disease and use of stringent in- and exclusion criteria. Outcome measures should therefore preferably be easily operatable and accessible to use over a long period of time. It can be considered to improve existing outcome measures by changing the included items, however, the disadvantage is that new natural history data is required, such as with the improvement of the Performance of the Upper Limb from version 1.2 to 2.0.[44, 45]

Another important characteristic of phase III trials are the in- and exclusion criteria, which influence the generalizability of the results. Inclusion criteria often contain a statement that patients have to be able to follow study instructions, which could be harder for patients with learning and behavioral disabilities. This could result in inclusion of less patients with a distal mutation in exon 63-79, because of a higher prevalence of these symptoms in this population. Whether specific patient characteristics influence the likeliness to participate in studies could be studied by comparing baseline characteristics of participating patients to patients who were eligible but decided together with their parents to refrain from participation.

Outcome measures of the upper extremities

Outcome measures of lower extremity function, such as the 6MWT, can only be used in ambulant patients. Drugs that restore dystrophin or improve the muscle quality need sufficient muscle tissue to target and the progressive replacement of muscle by fat and fibrosis, currently considered to be irreversible, limits the amount of muscle that can be targeted. Therefore, drugs that are proven effective in ambulant patients need to be studied separately and with different outcome measures in non-ambulant patients before the EMA and FDA will approve use in these populations as well.[38, 40] Most studies in non-ambulant patients have focused on the early non-ambulant phase. The available outcome measures often have a floor and ceiling effect showing little to no longitudinal change in patients who have limited function left or have little functional impairment. An example is the stable phase in the 6MWT and a maximum PUL score in many ambulant patients.[34, 46] In addition, there is a lack of outcome measures that are suitable for the more advanced stages of the disease. For known outcome measures there is a need for longitudinal natural history data in non-ambulant patients. Examples are strength tests, the Performance of the Upper Limb (PUL) motor function measure and the PROM. To overcome issues of the described outcome measures, innovative outcome measures and their characteristics could be explored, such as Kinect and Leap Motion outcome measures and quantitative muscle MRI.

Upper extremity muscle strength tests

Muscle strength is considered an important outcome measure, because it can quantitatively assess the clinical effect of the underlying muscle pathology in DMD.[40] However, it is not approved as primary endpoint, because FDA and EMA ruled that clinical relevance has to be established.[40] Muscle strength can be measured reliably using a hand-held dynamometer,[47] but this method cannot be used to assess distal muscle function in very weak patients. For this the MyoGrip and MyoPinch have been developed to assess isometric grip and pinch strength respectively (Figure 4).[48] The smallest change that can be measured is 0.01 kg for the MyoGrip and 0.001 kg for the MyoPinch. Both devices have been shown to be reliable and to show differences between DMD patients and controls.[48] Disadvantages are the need for a trained assessor, and vulnerable accessibility with a single manufacturer that is also needed for maintenance. Furthermore, the ability of MyoGrip and MyoPinch to predict important disease milestones has not yet been demonstrated.

Performance of the Upper Limb

The only approved outcome measure in non-ambulant patients is the Performance of the Upper Limb (PUL).[44, 45] It has been used as primary endpoint in two recent clinical trials in non-ambulant patients (NCT03406780 and NCT04371666). The first widely used version of the PUL was version 1.2 consisting of 20 items with a maximum total score of 74 points. Using a Rasch analysis and input from clinicians, the scoring per item was simplified to options of 0, 1, or 2 points and some items were removed because of redundancy and some

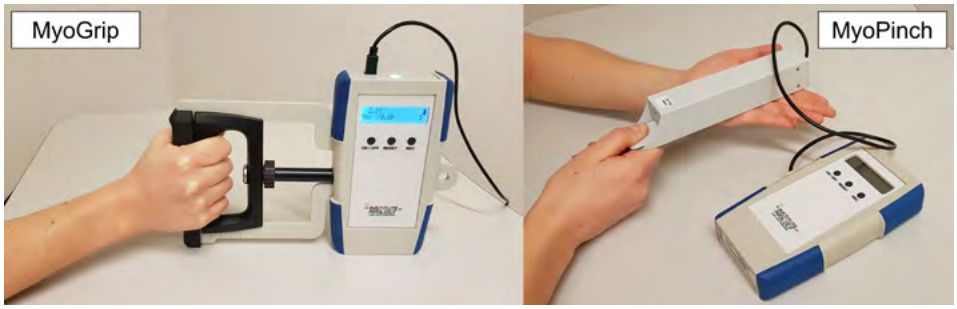


Figure 4. MyoGrip and MyoPinch
MyoGrip (left) and MyoPinch (right) devices are shown with the correct positions during strength measurements.

were separated or added.[45] The resulting PUL 2.0 consists of 22 motor items that are assessed by an observer and it yields a maximum total score of 42 points.[45] The items can be divided into three dimensions with a maximum score of 12 for the shoulder dimension, 17 for the elbow dimension and 13 for the distal wrist/hand dimension. During the course of this thesis, the first PUL longitudinal natural history data in DMD has been published. For a cohort of 177 DMD patients the annual change was found to be -1.5 points, where non-ambulant patients lost on average 1.1 points more than ambulant patients.[49] Limitations of the PUL are a ceiling effect in the upper and lower regions of the score and observer-dependence. Its prediction of important disease milestones has not yet been demonstrated, nor is a minimal clinically important difference available.

Patient reported outcome measures for the upper extremity

For the upper extremity in DMD there is one patient reported outcome measure: the DMD Upper Limb Patient Reported Outcome Measure (PROM). The most recent version of the PROM questionnaire contains 32 daily-life activity items which are scored on a three-level scale ('cannot do'; 'with difficulty'; 'easy') with a maximum total score of 64 points.[50] It describes self-reported meaningful upper extremity daily-life activities that could not otherwise be observed in a clinical or research setting (e.g. feeding, washing, and leisure activities). Internal consistency and test-retest reliability have been demonstrated.[50] Advantages are the fact that an assessor is not required and that data can be acquired off-site. Disadvantages are the subjective nature of the questionnaire and the current lack of change over time data and a minimal clinically important difference, reflecting relevant change for a patient. The FDA and EMA recommend the use of Patient Reported Outcome Measures as secondary outcome measure, because these can aid in determining clinical meaningfulness of objective findings of small magnitude, contribute to the assessment of benefit and risk, and assess the effect of a therapy on daily life activities.[38, 40]

Innovative outcome measures using Leap Motion and Microsoft Kinect

The Leap Motion controller and Microsoft Kinect v2 sensor are two innovative, low-cost marker-less motion capture systems that were developed by the gaming industry. Outcome measures of upper extremity motor function using these devices provide a continuous outcome parameter without a maximum score. This mitigates disadvantages of current outcome measures, such as a floor and ceiling effect. The Leap Motion controller uses infrared cameras to estimate the location of wrist and hand joints. This information can be used to calculate active ranges of motion (aROM) of these joints, and could have potential as outcome measure in advanced stages of DMD.[51] The Kinect obtains depth data with an infrared laser transmitter and an infrared camera. Using this data, real-time 3D-coordinates of body points, including the head, shoulders, elbows and wrists, are provided to estimate human posture. So far, two methods have used Kinect to assess the reaching ability of the arms in DMD: the 'reachable workspace envelope relative surface area' (RSA) and the 'Ability Captured Through Interactive Video Evaluation' (ACTIVE).[52, 53] Both measures could be used to objectively quantify upper extremity motor function without ceiling effect or observer-dependence, and might therefore be more sensitive to disease progression. Potential drawbacks could be the lack of insight in constraints of the software and hardware due to intellectual property, and possible software updates and hardware discontinuation which could jeopardize their use in clinical trials.

Quantitative muscle MR methods

Quantitative muscle magnetic resonance (qMR) methods such as MRI and MR spectroscopy (MRS) can be used to assess muscle pathology in muscular dystrophies including DMD. qMR is considered to be a promising biomarker, because it has the potential to accurately assess different aspects of muscle wasting in this muscle wasting disease.[54-59] MRI is non-invasive and safe, and most patients aged five years and older tolerate scanning up to one hour well without anesthesia.[60-62] However, lying still for prolonged periods of time in a specific position can be strenuous for patients. qMR methods of the lower extremity have been shown to accurately describe fat replacement and tissue edema in DMD, and to have excellent test-retest reproducibility within and across centers.[55, 57-59, 63-66]

At the start of the projects described in this thesis, muscle MRI of the upper extremities had only been studied in small cohorts of primarily ambulant DMD patients and mainly of the forearm muscles.[58, 67, 68] Muscle MRI of the upper extremities could pose extra hurdles compared to the lower extremities, because of the smaller muscle mass and position of the arms at the side of the body and therefore not in the center of the MR scanner, which decreases the image quality and frequently causes artefacts. A position where the participants are lying on their side could be chosen to overcome the off-center position of the upper extremity, but this could be difficult for patients to maintain.

Different qMR techniques can be used to measure fat replacement and tissue edema. Muscle fat fraction (FF) is the most studied qMR parameter and can be assessed using water-fat MRI scans or MRS. MRS is the gold standard for FF determination with a high accuracy and reproducibility.[65] A disadvantage of MRS is that results are usually obtained for one pre-specified region of interest, resulting in information for only a specific part of a single muscle or muscle group. Therefore, information about spatial variation within or between muscles cannot be obtained. Tissue edema or inflammation can be studied via the T2 relaxation time of water in the muscle ($T2_{water}$), which can be determined using multi-echo spin-echo (MSE) imaging or spectroscopy sequences.[69-71]

Fat-water imaging

Chemical shift based fat-water imaging is based on the difference in precession speed between protons in water and fat.[65] In a technique originally shown by Dixon,[72] images are acquired at different echo times, or phases, and from these images a purely water and fat image can be obtained (Figure 5). The FF in a specific area, typically a single muscle or muscle group, can be calculated by dividing the signal on the fat image by the combined signal of the water and fat image in that area.

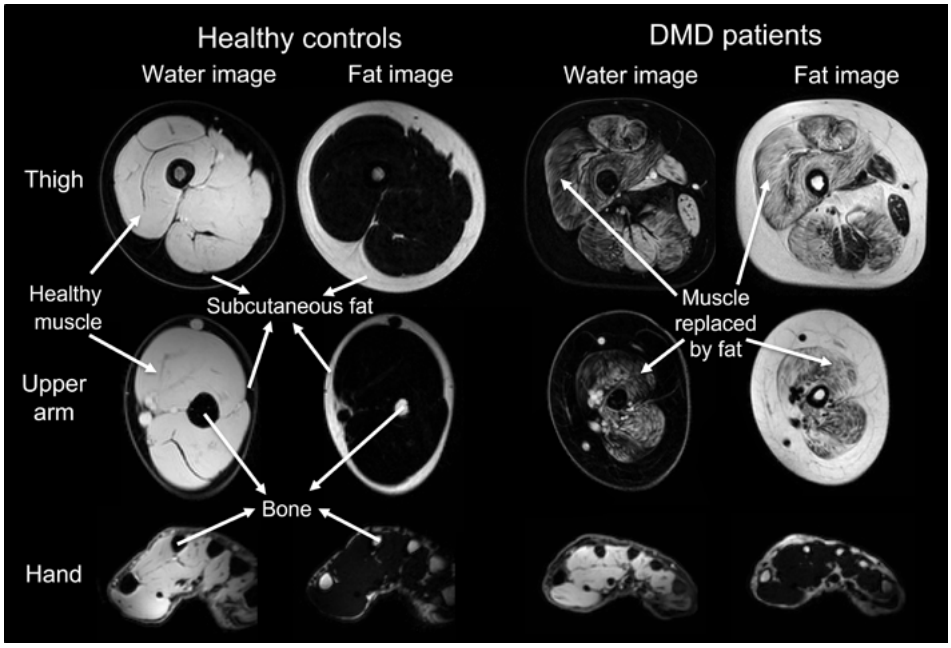


Figure 5. Dixon muscle MRI images of the extremities
Dixon water and fat images of thigh, upper arm, and hand muscles of healthy controls and DMD patients. In the figure, arrows point at the following tissues: healthy muscle, muscle replaced by fat, subcutaneous fat, and bone.

qMR muscle fat fraction

In the lower extremities, muscle FF has been shown to increase with age and to correlate with motor function tests and strength of specific muscle groups.[54, 57, 58, 73, 74] A study by Willcocks at al. illustrated that FF is more sensitive to disease progression than functional tests, by showing that FF increased significantly over 12 months, even in some boys in whom 6MWT distances increased due to growth.[58] Also in forearm muscles FF has been shown to increase over six or 12 months, and in upper arm muscles FF correlated with the PUL and grip and pinch strength.[67, 68, 75] In non-ambulant patients, limited data on qMR FF of upper arm muscles is available.[75] No studies have examined very distal muscles in DMD, which are thought to be relatively spared in advanced disease stages.

Furthermore, the extensively described correlations between muscle FF, strength, and function at the time of qMR do not prove causality.[54, 56, 57, 68, 74, 76-79] Even unrelated biological parameters that consistently increase or decrease with age will inherently correlate with functional parameters in a progressive disease. Therefore, such correlations alone are not sufficient to show clinical relevance of muscle FF. For this, qMR FF should be able to predict loss of an important disease milestone, such as the ability to ambulate or bring a glass to the mouth. Such a predictive ability has not yet been demonstrated, and would help to qualify muscle qMR FF as outcome measure that can be used as primary or secondary endpoint in trials.

T2 relaxation time of water in muscle

The T2 relaxation time is an MR parameter which is prolonged by increased water mobility, such as in inflammation, necrosis and fatty replacement.[70] The T2 relaxation time of water in muscle ($T2_{water}$) is relatively short. This parameter has been used to study tissue edema or inflammation, and thus disease activity, as these changes are thought to occur in muscles prior to fat replacement and fibrosis.[69-71] $T2_{water}$ can be determined accurately using MRS, as water and fat signals can be easily discriminated using this method. However, this method is commonly applied using a single voxel approach due to time constraints, and hence only data are acquired from a pre-specified region of interest.

$T2_{water}$ can also be determined with an imaging based method using an MSE sequence. In this way, spatial information is preserved and data can be obtained for a large field of view that covers different muscles or muscle groups. The T2 relaxation time of fat ($T2_{fat}$) is much longer (~ 130-170 ms at 3 Tesla) than the $T2_{water}$ in muscle (25-30ms), and the increase in $T2_{water}$ due to disease activity (~ 1-4 ms).[69-71] An MRI MSE sequence consist of multiple spin echoes from which the 'global' T2 relaxation time can be determined using a mono-exponential fit of the signal decay as a function of the echo time (Figure 6).[70] This global T2 consists of both the signal decay of water and fat. In DMD, due to progressive fat replacement, the signal of fat dominates the value of the global T2. Therefore, it is necessary to separate the signal decay into a water and fat component, to determine the $T2_{water}$ and $T2_{fat}$ (Figure 6).[71]

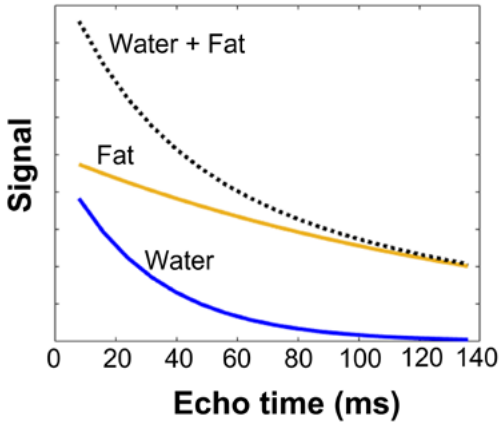


Figure 6. T2 relaxation time
MRI signal as a function of the echo time, where the signal of water and fat combined is shown with a dotted line which decays due to 'global' T2 relaxation. The water (blue line) signal experiences a different T2 relaxation ($T2_{water}$) than the fat (yellow) signal ($T2_{fat}$). This global T2 relaxation can be separated in the T2 relaxation time of fat (yellow) and water (blue). *Kindly provided by K.R. Keene.*

$T2_{water}$ has been studied as a qMR marker for disease activity,[69, 70] and in young DMD patients, lower extremity muscles with limited fat replacement were shown to have an elevated $T2_{water}$.[69] Some data on $T2_{water}$ in forearm muscles is available, but $T2_{water}$ is more difficult to interpret in fat replaced muscles and therefore more difficult to study in advanced disease stages.[68] No studies have examined $T2_{water}$ in very distal muscles, which might show interesting results in advanced disease stages in case of limited fat replacement.

Objectives of this thesis

DMD is a rare and fatal muscle wasting disorder. Beside chronic use of glucocorticoids there is currently no fully approved therapy available. Due to improved care and glucocorticoids, patients have a longer life-expectancy and therefore go longer through life in the non-ambulant phase. The importance of outcome measures and natural history data was demonstrated when the 6MWT was used as primary endpoint in the first clinical trials in DMD. Outcome measures should demonstrate sufficient reliability, construct validity, concurrent validity, longitudinal change and accessibility. Clinical relevance of outcome measures should also be demonstrated, for instance via their relation with important disease milestones. Due to a progressive reduction in muscle tissue to be targeted by drugs, separate clinical trials need to be performed in non-ambulant patients, therefore specific outcome measures are required for this disease stage.
The overall aim of this thesis was to identify outcome measures in DMD, specifically for non-ambulant patients, that detect an effect of a new therapy that is clinically relevant for

patients in a relatively short period of time according to what is feasible in clinical trials. The use of such outcome measures could lead to a reduction in the sample sizes and potentially to a lower burden for patients.

Outline of this thesis

Chapter 2 describes reasons why DMD and BMD patients and parents declined participation in observational studies and discusses the presence of selection bias by comparing characteristics of participants and non-participants in these studies. These results can be used to optimize patient participation in studies on this rare disease and to stimulate representative observational research.

Chapter 3 describes the relation between quantitative MRI (qMRI) FF of a lower extremity muscle and loss of ambulation on top of the effect of age in ambulant patients with DMD. The aim was to predict the loss of this important milestone and thus establish the clinical relevance of muscle qMRI FF to support its use as an outcome measure in clinical trials.

Chapter 4 describes the relation between qMRI FF of an upper extremity muscle and loss of hand-to-mouth movement on top of the effect of age in non-ambulant patients with DMD. The aim was to establish clinical relevance by predicting the loss of this important upper extremity disease milestone and thus to support the use of muscle qMRI FF in clinical trials in non-ambulant patients.

Chapter 5 describes qMRI results of the thenar muscles and hand function over one year to establish the value of the thenar muscles for monitoring treatment effects in non-ambulant DMD patients. Preservation of these muscles and measurable disease progression in these muscles would indicate that the thenar muscles are a valuable and quantifiable biomarker and target for systemic or local therapy in the later non-ambulant stages of the disease.

Chapter 6 illustrates the potential and constraints of using sensors from the gaming industry to develop upper extremity outcome measures for non-ambulant patients with DMD. The results of these novel outcome measures are compared to the currently used PUL and PROM outcome measures.

References

1. Mendell JR, Shilling C, Leslie ND, et al. Evidence-based path to newborn screening for Duchenne muscular dystrophy. Ann Neurol 2012;71:304-313. doi:10.1002/ana.23528
2. Moat SJ, Bradley DM, Salmon R, et al. Newborn bloodspot screening for Duchenne muscular dystrophy: 21 years experience in Wales (UK). Eur J Hum Genet 2013;21:1049-1053. doi:10.1038/ejhg.2012.301
3. Ciafaloni E, Fox DJ, Pandya S, et al. Delayed Diagnosis in Duchenne Muscular Dystrophy: Data from the Muscular Dystrophy Surveillance, Tracking, and Research Network (MD STARnet). J Pediatr-Us 2009;155:380-385. doi:10.1016/j.jpeds.2009.02.007
4. Birnkrant DJ, Bushby K, Bann CM, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and neuromuscular, rehabilitation, endocrine, and gastrointestinal and nutritional management. Lancet Neurol 2018;17:251-267. doi:10.1016/S1474-4422(18)30024-3
5. McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. Lancet 2018;391:451-461. doi:10.1016/S0140-6736(17)32160-8
6. van den Bergen JC, Ginjaar HB, van Essen AJ, et al. Forty-Five Years of Duchenne Muscular Dystrophy in The Netherlands. J Neuromuscul Dis 2014;1:99-109
7. Janssen M, Harlaar J, Koopman B, de Groot IJM. Dynamic arm study: quantitative description of upper extremity function and activity of boys and men with duchenne muscular dystrophy. J Neuroeng Rehabil 2017;14:45. doi:10.1186/s12984-017-0259-5
8. Hogrel JY, Decostre V, Ledoux I, et al. Normalized grip strength is a sensitive outcome measure through all stages of Duchenne muscular dystrophy. Journal of neurology 2020;267:2022-2028. doi:10.1007/s00415-020-09800-9
9. Heutinck L, Kampen NV, Jansen M, Groot IJ. Physical Activity in Boys With Duchenne Muscular Dystrophy Is Lower and Less Demanding Compared to Healthy Boys. J Child Neurol 2017;32:450-457. doi:10.1177/0883073816685506
10. Banihani R, Smile S, Yoon G, et al. Cognitive and Neurobehavioral Profile in Boys With Duchenne Muscular Dystrophy. J Child Neurol 2015;30:1472-1482. doi:10.1177/0883073815570154
11. Monaco AP, Bertelson CJ, Liechti-Gallati S, et al. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics 1988;2:90-95
12. Bushby KM, Gardner-Medwin D. The clinical, genetic and dystrophin characteristics of Becker muscular dystrophy. I. Natural history. Journal of neurology 1993;240:98-104. doi:10.1007/BF00858725
13. Clemens PR, Niizawa G, Feng J, et al. The CINRG Becker Natural History Study: Baseline characteristics. Muscle Nerve 2020;62:369-376. doi:10.1002/mus.27011
14. Allen DG, Whitehead NP, Froehner SC. Absence of Dystrophin Disrupts Skeletal Muscle Signaling: Roles of Ca2+, Reactive Oxygen Species, and Nitric Oxide in the Development of Muscular Dystrophy. Physiol Rev 2016;96:253-305. doi:10.1152/physrev.00007.2015
15. Cros D, Harnden P, Pellissier JF, Serratrice G. Muscle hypertrophy in Duchenne muscular dystrophy. A pathological and morphometric study. Journal of neurology 1989;236:43-47. doi:10.1007/bf00314217
16. Doorenweerd N, Mahfouz A, van Putten M, et al. Timing and localization of human dystrophin isoform expression provide insights into the cognitive phenotype of Duchenne muscular dystrophy. Sci Rep 2017;7:12575. doi:10.1038/s41598-017-12981-5
17. Chamova T, Guergueltcheva V, Raycheva M, et al. ASSOCIATION BETWEEN LOSS OF Dp140 AND COGNITIVE IMPAIRMENT IN DUCHENNE AND BECKER DYSTROPHIES. Balk J Med Genet 2013;16:21-29. doi:10.2478/bjmg-2013-0014
18. Ricotti V, Mandy WPL, Scoto M, et al. Neurodevelopmental, emotional, and behavioural problems in Duchenne muscular dystrophy in relation to underlying dystrophin gene mutations. Developmental Medicine and Child Neurology 2016;58:77-84. doi:10.1111/dmcn.12922
19. Taylor PJ, Betts GA, Maroulis S, et al. Dystrophin gene mutation location and the risk of cognitive impairment in Duchenne muscular dystrophy. PLoS One 2010;5:e8803. doi:10.1371/journal.pone.0008803
20. Pane M, Lombardo ME, Alfieri P, et al. Attention deficit hyperactivity disorder and cognitive function in Duchenne muscular dystrophy: phenotype-genotype correlation. J Pediatr 2012;161:705-709 e701. doi:10.1016/j.jpeds.2012.03.020
21. Birnkrant DJ, Bushby K, Bann CM, et al. Diagnosis and management of Duchenne muscular dystrophy, part 2: respiratory, cardiac, bone health, and orthopaedic management. Lancet Neurol 2018;17:347-361. doi:10.1016/S1474-4422(18)30025-5
22. Birnkrant DJ, Bushby K, Bann CM, et al. Diagnosis and management of Duchenne muscular dystrophy, part 3: primary care, emergency management, psychosocial care, and transitions of care across the lifespan. Lancet Neurol 2018;17:445-455. doi:10.1016/S1474-4422(18)30026-7

23. Gloss D, Moxley RT, Ashwal S, Oskoui M. Practice guideline update summary: Corticosteroid treatment of Duchenne muscular dystrophy Report of the Guideline Development Subcommittee of the American Academy of Neurology. Neurology 2016;86:465-472. doi:10.1212/Wnl.0000000000002337
24. Landfeldt E, Thompson R, Sejersen T, et al. Life expectancy at birth in Duchenne muscular dystrophy: a systematic review and meta-analysis. Eur J Epidemiol 2020;35:643-653. doi:10.1007/s10654-020-00613-8
25. Bushby K, Finkel R, Birnkrant DJ, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. Lancet Neurol 2010;9:77-93. doi:10.1016/S1474-4422(09)70271-6
26. Verhaart IEC, Aartsma-Rus A. Therapeutic developments for Duchenne muscular dystrophy. Nature reviews Neurology 2019;15:373-386. doi:10.1038/s41582-019-0203-3
27. Roy B, Griggs R. Advances in Treatments in Muscular Dystrophies and Motor Neuron Disorders. Neurol Clin 2021;39:87-112. doi:10.1016/j.ncl.2020.09.005
28. Straub V, Mercuri E, Grp DOMS. Report on the workshop: Meaningful outcome measures for Duchenne muscular dystrophy, London, UK, 30-31 January 2017. Neuromuscular Disorders 2018;28:690-701. doi:10.1016/j.nmd.2018.05.013
29. Bushby K, Finkel R, Wong B, et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. Muscle Nerve 2014;50:477-487. doi:10.1002/mus.24332
30. Mendell JR, Goemans N, Lowes LP, et al. Longitudinal effect of eteplirsen versus historical control on ambulation in Duchenne muscular dystrophy. Ann Neurol 2016;79:257-271. doi:10.1002/ana.24555
31. McDonald CM, Henricson EK, Han JJ, et al. The 6-Minute Walk Test as a New Outcome Measure in Duchenne Muscular Dystrophy. Muscle & Nerve 2010;41:500-510. doi:10.1002/mus.21544
32. McDonald CM, Henricson EK, Abresch RT, et al. The 6-minute walk test and other endpoints in Duchenne muscular dystrophy: longitudinal natural history observations over 48 weeks from a multicenter study. Muscle Nerve 2013;48:343-356. doi:10.1002/mus.23902
33. Goemans N, Vanden Hauwe M, Signorovitch J, et al. Individualized Prediction of Changes in 6-Minute Walk Distance for Patients with Duchenne Muscular Dystrophy. PLoS One 2016;11:e0164684. doi:10.1371/journal.pone.0164684
34. Mercuri E, Signorovitch JE, Swallow E, et al. Categorizing natural history trajectories of ambulatory function measured by the 6-minute walk distance in patients with Duchenne muscular dystrophy. Neuromuscul Disord 2016;26:576-583. doi:10.1016/j.nmd.2016.05.016
35. Arora H, Willcocks RJ, Lott DJ, et al. Longitudinal timed function tests in Duchenne muscular dystrophy: ImagingDMD cohort natural history. Muscle & Nerve 2018;58:631-638. doi:10.1002/mus.26161
36. McDonald CM, Henricson EK, Abresch RT, et al. The 6-minute walk test and other clinical endpoints in duchenne muscular dystrophy: reliability, concurrent validity, and minimal clinically important differences from a multicenter study. Muscle Nerve 2013;48:357-368. doi:10.1002/mus.23905
37. Mazzone E, Martinelli D, Berardinelli A, et al. North Star Ambulatory Assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. Neuromuscul Disord 2010;20:712-716. doi:10.1016/j.nmd.2010.06.014
38. CDER, CBER. Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment. Guidance for Industry. https://www.fda.gov/media/92233/download. Accessed on May 1, 2019. US Food & Drug Administration (FDA) 2018
39. Charleston JS, Schnell FJ, Dworzak J, et al. Eteplirsen treatment for Duchenne muscular dystrophy: Exon skipping and dystrophin production. Neurology 2018;90:e2146-e2154. doi:10.1212/WNL.0000000000005680
40. CHMP. Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. In European Medicines Agency (EMA). Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf. Accessed December 18, 2020.
41. Johnston BC, Ebrahim S, Carrasco-Labra A, et al. Minimally important difference estimates and methods: a protocol. Bmj Open 2015;5. doi:ARTN e007953
10.1136/bmjopen-2015-007953
42. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Rev Pharm Out 2011;11:171-184. doi:10.1586/Erp.11.9
43. Straub V, Balabanov P, Bushby K, et al. Stakeholder cooperation to overcome challenges in orphan medicine development: the example of Duchenne muscular dystrophy. Lancet Neurol 2016;15:882-890. doi:10.1016/S1474-4422(16)30035-7
44. Mayhew A, Mazzone ES, Eagle M, et al. Development of the Performance of the Upper Limb module for Duchenne muscular dystrophy. Dev Med Child Neurol 2013;55:1038-1045. doi:10.1111/dmcn.12213
45. Mayhew AG, Coratti G, Mazzone ES, et al. Performance of Upper Limb module for Duchenne muscular dystrophy. Dev Med Child Neurol 2019. doi:10.1111/dmcn.14361

46. Pane M., Mazzone E.S., Sivo S., et al. The 6 minute walk test and performance of upper limb in ambulant duchenne muscular dystrophy boys. PLoS Curr 2014. doi:10.1371/currents.md.a93d9904d57dcb08936f2ea89bca6fe6
47. Mendell JR, Florence J. Manual muscle testing. Muscle Nerve 1990;13 Suppl:S16-20. doi:10.1002/mus.880131307
48. Servais L, Deconinck N, Moraux A, et al. Innovative methods to assess upper limb strength and function in non-ambulant Duchenne patients. Neuromuscul Disord 2013;23:139-148. doi:10.1016/j.nmd.2012.10.022
49. Pane M, Coratti G, Brogna C, et al. Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal data. PLoS One 2018;13:e0199223. doi:10.1371/journal.pone.0199223
50. Klingels K, Mayhew AG, Mazzone ES, et al. Development of a patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy: DMD Upper Limb PROM. Dev Med Child Neurol 2017;59:224-231. doi:10.1111/dmcn.13277
51. Nizamis K, Rijken NHM, Mendes A, et al. A Novel Setup and Protocol to Measure the Range of Motion of the Wrist and the Hand. Sensors (Basel) 2018;18. doi:10.3390/s18103230
52. Han JJ, de Bie E, Nicorici A, et al. Reachable workspace and performance of upper limb (PUL) in duchenne muscular dystrophy. Muscle Nerve 2016;53:545-554. doi:10.1002/mus.24894
53. Lowes LP, Alfano LN, Crawfis R, et al. Reliability and validity of active-seated: An outcome in dystrophinopathy. Muscle Nerve 2015;52:356-362. doi:10.1002/mus.24557
54. Akima H, Lott D, Senesac C, et al. Relationships of thigh muscle contractile and non-contractile tissue with function, strength, and age in boys with Duchenne muscular dystrophy. Neuromuscul Disord 2012;22:16-25. doi:10.1016/j.nmd.2011.06.750
55. Azzabou N, Loureiro de Sousa P, Caldas E, Carlier PG. Validation of a generic approach to muscle water T2 determination at 3T in fat-infiltrated skeletal muscle. J Magn Reson Imaging 2015;41:645-653. doi:10.1002/jmri.24613
56. Barnard AM, Willcocks RJ, Finanger EL, et al. Skeletal muscle magnetic resonance biomarkers correlate with function and sentinel events in Duchenne muscular dystrophy. PLoS One 2018;13:e0194283. doi:10.1371/journal.pone.0194283
57. Bonati U, Hafner P, Schadelin S, et al. Quantitative muscle MRI: A powerful surrogate outcome measure in Duchenne muscular dystrophy. Neuromuscul Disord 2015;25:679-685. doi:10.1016/j.nmd.2015.05.006
58. Willcocks RJ, Rooney WD, Triplett WT, et al. Multicenter prospective longitudinal study of magnetic resonance biomarkers in a large duchenne muscular dystrophy cohort. Ann Neurol 2016;79:535-547. doi:10.1002/ana.24599
59. Strijkers GJ, Araujo ECA, Azzabou N, et al. Exploration of New Contrasts, Targets, and MR Imaging and Spectroscopy Techniques for Neuromuscular Disease - A Workshop Report of Working Group 3 of the Biomedicine and Molecular Biosciences COST Action BM1304 MYO-MRI. Journal of Neuromuscular Diseases 2019;6:1-30. doi:10.3233/Jnd-180333
60. Chou IJ, Tench CR, Gowland P, et al. Subjective discomfort in children receiving 3 T MRI and experienced adults' perspective on children's tolerability of 7 T: a cross-sectional questionnaire survey. BMJ Open 2014;4:e006094. doi:10.1136/bmjopen-2014-006094
61. Mercuri E, Pichiecchio A, Counsell S, et al. A short protocol for muscle MRI in children with muscular dystrophies. Eur J Paediatr Neurol 2002;6:305-307
62. Tornqvist E, Mansson A, Hallstrom I. Children having magnetic resonance imaging: A preparatory storybook and audio/visual media are preferable to anesthesia or deep sedation. J Child Health Care 2015;19:359-369. doi:10.1177/1367493513518374
63. Forbes SC, Walter GA, Rooney WD, et al. Skeletal muscles of ambulant children with Duchenne muscular dystrophy: validation of multicenter study of evaluation with MR imaging and MR spectroscopy. Radiology 2013;269:198-207. doi:10.1148/radiol.13121948
64. Wokke BH, Bos C, Reijnierse M, et al. Comparison of dixon and T1-weighted MR methods to assess the degree of fat infiltration in duchenne muscular dystrophy patients. Journal of Magnetic Resonance Imaging 2013;38:619-624. doi:10.1002/jmri.23998
65. Burakiewicz J, Sinclair CDJ, Fischer D, et al. Quantifying fat replacement of muscle by quantitative MRI in muscular dystrophy. Journal of neurology 2017. doi:10.1007/s00415-017-8547-3
66. Hooijmans MT, Niks EH, Burakiewicz J, et al. Non-uniform muscle fat replacement along the proximodistal axis in Duchenne muscular dystrophy. Neuromuscul Disord 2017;27:458-464. doi:10.1016/j.nmd.2017.02.009
67. Ricotti V, Evans MR, Sinclair CD, et al. Upper Limb Evaluation in Duchenne Muscular Dystrophy: Fat-Water Quantification by MRI, Muscle Force and Function Define Endpoints for Clinical Trials. PLoS One 2016;11:e0162542. doi:10.1371/journal.pone.0162542
68. Hogrel JY, Wary C, Moraux A, et al. Longitudinal functional and NMR assessment of upper limbs in Duchenne muscular dystrophy. Neurology 2016;86:1022-1030. doi:10.1212/WNL.0000000000002464

69. Arpan I, Willcocks RJ, Forbes SC, et al. Examination of effects of corticosteroids on skeletal muscles of boys with DMD using MRI and MRS. Neurology 2014;83:974-980. doi:10.1212/WNL.0000000000000775
70. Carlier PG. Global T2 versus water T2 in NMR imaging of fatty infiltrated muscles: different methodology, different information and different implications. Neuromuscul Disord 2014;24:390-392. doi:10.1016/j.nmd.2014.02.009
71. Keene KR, Beenakker JM, Hooijmans MT, et al. T2 relaxation-time mapping in healthy and diseased skeletal muscle using extended phase graph algorithms. Magn Reson Med 2020;84:2656-2670. doi:10.1002/mrm.28290
72. Dixon WT. Simple Proton Spectroscopic Imaging. Radiology 1984;153:189-194. doi:DOI 10.1148/radiology.153.1.6089263
73. Godi C, Ambrosi A, Nicastro F, et al. Longitudinal MRI quantification of muscle degeneration in Duchenne muscular dystrophy. Ann Clin Transl Neurol 2016;3:607-622. doi:10.1002/acn3.319
74. Wokke BH, van den Bergen JC, Versluis MJ, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. Neuromuscul Disord 2014;24:409-416. doi:10.1016/j.nmd.2014.01.015
75. Forbes SC, Arora H, Willcocks RJ, et al. Upper and Lower Extremities in Duchenne Muscular Dystrophy Evaluated with Quantitative MRI and Proton MR Spectroscopy in a Multicenter Cohort. Radiology 2020;295:616-625. doi:10.1148/radiol.2020192210
76. Fischmann A, Hafner P, Gloor M, et al. Quantitative MRI and loss of free ambulation in Duchenne muscular dystrophy. Journal of neurology 2013;260:969-974. doi:10.1007/s00415-012-6733-x
77. Wary C, Azzabou N, Giraudeau C, et al. Quantitative NMRI and NMRS identify augmented disease progression after loss of ambulation in forearms of boys with Duchenne muscular dystrophy. NMR Biomed 2015;28:1150-1162. doi:10.1002/nbm.3352
78. Willcocks RJ, Triplett WT, Forbes SC, et al. Magnetic resonance imaging of the proximal upper extremity musculature in boys with Duchenne muscular dystrophy. Journal of neurology 2017;264:64-71. doi:10.1007/s00415-016-8311-0
79. Gaeta M, Messina S, Mileto A, et al. Muscle fat-fraction and mapping in Duchenne muscular dystrophy: evaluation of disease distribution and correlation with clinical assessments. Preliminary experience. Skeletal Radiol 2012;41:955-961. doi:10.1007/s00256-011-1301-5

Chapter 2

Decision-making and selection bias in four observational studies on Duchenne and Becker muscular dystrophy

Published in: J Neuromuscul Dis
Sept 2020; 7(4); 433-442; DOI:10.3233/JND-200541

Karin J. Naarding | Nathalie Doorenweerd | Zaïda Koeks | Ruben G.F. Hendriksen
Kinita A. Chotkan | Yvonne D. Krom | Imelda J.M. de Groot | Chiara S. Straathof
Erik H. Niks | Hermien E. Kan

Abstract

Background

Natural history data are essential for trial design in Duchenne (DMD) and Becker muscular dystrophy (BMD), but recruitment for observational studies can be challenging.

Objective

We reviewed reasons why patients or caregivers declined participation, and compared characteristics of participants and non-participants to assess possible selection bias in four observational studies, three on DMD and one on BMD.

Methods

Three pediatric DMD studies focused on cross-sectional cognitive function and brain MRI (DMDbrain, n=35 and DMDperfusion, n=12), and on longitudinal upper extremity function and muscle MRI (DMDarm, n=22). One adult BMD study assessed longitudinal functioning (n=36). Considerations for non-participation were retrospectively reviewed from screening logs. Age, travel-time, *DMD* gene mutations and age at loss of ambulation (DMDarm and BMD study only), of participants and non-participants were derived from the Dutch Dystrophinopathy Database and compared using nonparametric tests (p<0.05).

Results

The perceived burden of the protocol (38.2%), use of MRI (30.4%), and travel-time to the study site (19.1%) were the most frequently reported considerations for non-participation. Only few patients reported lack of personal gain (0.0-5.9%). Overall, participating patients were representative for the studied sub-populations, except for a younger age of DMDarm study participants and a complete lack of participants with a mutation beyond exon 63.

Conclusion

Optimizing patient involvement in protocol design, improving MRI experiences, and integrating research into clinics are important factors to decrease burden and facilitate participation. Nationwide registries are essential to compare participants and non-participants and ensure representative observational research. Specific effort is needed to include patients with distal mutations in cognitive studies.

Introduction

Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) are caused by mutations in the *DMD* gene.[1] This leads to absence of dystrophin in DMD, and to a truncated and partly functional protein in BMD muscles. These neuromuscular diseases form a spectrum in which DMD patients lose ambulation around their early teens, while BMD patients have a milder but more variable disease course.[2, 3] In both diseases, a higher prevalence of learning and behavioral disabilities has been reported.[4-6] In DMD, this is associated with absence of different dystrophin isoforms in the brain.[7, 8]

Although the first drugs in DMD have now received regulatory approval, there is no cure yet.[9] Currently, many studies worldwide are recruiting DMD patients simultaneously: 21 interventional clinical trials and 15 observational studies (ClinicalTrials.gov accessed on February 2nd 2020). BMD patients are being recruited for three interventional clinical trials and two observational studies worldwide (ClinicalTrials.gov accessed on February 2nd 2020). These clinical trials are challenging because of the rarity of the diseases and a variable rate of progression.[9] which stresses the importance of detailed knowledge of the natural history.[10] The possibility to use historical controls reduces the required number of participants per study,[11] but even further highlights the need for high quality natural history data. In observational studies however, direct benefit to patients is lacking, while the added burden of research on top of the disease and clinical care could be perceived as high. Knowledge of factors that influence the decision-making process for participation can be used when designing study protocols in order to increase the participation rate and avoid selection bias. Such detailed and high quality natural history data would enable their use for placebo arms, and for determination of primary and secondary outcome measures in interventional trials. While considerations for not participating have been described for interventional trials,[12-14] only one observational study reported on this topic.[5]

In the present study, we reviewed the decision-making considerations reported by eligible patients and compared patient characteristics of participants and non-participants in three DMD and one BMD observational studies conducted at our institute.

Methods

DMD and BMD patients were recruited in the following observational studies at the Leiden University Medical Center (LUMC): 'Non-invasive assessment of brain involvement in DMD' (DMDbrain; ABR number NL23184.058.09; onset of recruitment in 2010), 'The background of the reduced cerebral blood flow in DMD' (DMDperfusion; ABR number NL58182.058.16; onset of recruitment in 2017), 'Upper extremity outcome measures in non-ambulant DMD patients' (DMDarm; ABR number NL63133.058.17; onset of recruitment in 2018), and 'The natural history study of BMD' (BMD; ABR number NL50171.058.14**;** onset of recruitment in 2014). All studies are registered at ToetsingOnline (www.toetsingonline.nl). For recruitment, the Dutch

Dystrophinopathy Database (DDD) was used ('Epidemiology, natural course and registration of dystrophinopathies in the Netherlands'; ABR number NL21411.058.08) [3]. This nationwide registry, initiated in 2008, provided the opportunity for all Dutch DMD and BMD patients to list their names and contact details together with details on comorbidities, medication use, disease history and current functional status. The local ethics committee approved all studies and the registry in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. Written informed consent had been obtained from all patients and from legal representatives for patients under 16 years of age.

DMDbrain – Non-invasive assessment of brain involvement in DMD

Thirty-five DMD patients were recruited from the DDD and through the Duchenne Parent Project Netherlands (DPP NL) newsletter. Inclusion criteria were: male, genetically confirmed DMD patients ≥8 years old. Exclusion criteria were: MRI contra-indications such as scoliosis surgery, daytime artificial ventilation or the inability to lie supine for 45 minutes. Patients were included in the study from March 2010 until October 2012. The cross-sectional study design consisted of a single visit and included a one-hour neuropsychological assessment and two 30 minute MRI scans (at 3 Tesla and at 7 Tesla) of the brain. Results have previously been reported.[15-18]

DMDperfusion – The background of the reduced cerebral blood flow in DMD

Thirteen DMD patients were recruited from the DDD, the LUMC outpatient clinic, and through a poster at the DPP NL annual conference. Inclusion criteria were: ambulant male, genetically confirmed DMD patients ≥10 years old. Exclusion criteria were MRI contraindications, a medical history of cardiovascular disease, diabetes mellitus, neurological disease (other than DMD), recurrent syncope, and joint contractures preventing the use of the tilting table. Patients were included in the study beginning January 2017 and recruitment is ongoing. The cross-sectional protocol includes a one-and-a half hour tilting table experiment with transcranial doppler and blood pressure measurements, 30 minute (task-based) MRI of the brain and cerebral vasculature, brief neuropsychological assessment (20 minutes), and cardiac ultrasound if this was not available from a recent clinical care visit.

DMDarm – Upper extremity outcome measures in non-ambulant DMD patients

Twenty-two DMD patients were recruited from the DDD, via Dutch neurologists and rehabilitation specialists, and through the Spierziekten Nederland (SN) website, the DPP NL website and Facebook page, and a poster at the DPP NL and SN annual conferences. Inclusion criteria were male, non-ambulant genetically confirmed DMD patients ≥8 years old. Exclusion criteria were: MRI contra-indications, exposure to an investigational drug ≤6 months prior to participation and recent (≤6 months) upper extremity surgery or trauma. Patients were included in the study from April 2018 until June 2019. This ongoing longitudinal study consist of three half-day visits at 0-12-18 months and the protocol includes functional upper extremity outcome measures and a 45 minute MRI scan of the upper extremity.

BMD – The natural history study of BMD

Thirty-six BMD patients were recruited from the DDD and the LUMC outpatient clinic. Inclusion criteria were male BMD patients ≥18 years old. BMD was defined as follows: an in-frame mutation in the dystrophin gene, or a reduced amount of dystrophin in a muscle biopsy, or an out-of-frame mutation with a mild disease course (>16 years old at loss of ambulation). Patients were included in the study from November 2014 until June 2016. The longitudinal study required four half-day to full-day visits at 0-12-24-36 months including functional tests, cardiac ultrasound and pulmonary function tests, and a single neuropsychological assessment. Optionally, patients could also participate in the following sub studies: 1) yearly blood sample collection for biomarker studies, 2) muscle biopsies at one time-point, and 3) lower extremity muscle MRI at two time-points. Last follow-up visit took place in August 2019.

Data collection

Review of considerations for non-participation

All considerations for non-participation were obtained during the telephone calls used for the inclusion. For the DMDbrain study, the study information letter was sent first, and potential participants or their legal representatives were called within a few weeks to discuss inclusion. If patients or their legal representatives decided not to participate, they were not actively asked for reasons for non-participation as this could be perceived as pressure to participate. When they volunteered a reason, this was recorded. For the other three studies, the decision not to participate could either be made at the first telephone call, before the study information letter was sent, or at the second telephone call after reading the study information letter. At the time of these studies, more thorough implementation of Good Clinical Practice (GCP) guidelines led to more in depth logging of the screening and enrollment process. Therefore, information on considerations for non-participation in these three studies was actively requested, although patients were always allowed to not answer this question.

All considerations for non-participation that had been recorded in the screening and enrollment logs of all studies were gathered retrospectively by one observer (KJN) and checked by a second observer (ND). Patients or parents could provide one or more considerations, and these considerations were divided in the following groups: 'Burden of protocol', 'Travel-time', 'Burden of clinical care', 'Other research', 'No advantage', 'Not interested', 'MRI'. Definitions and examples of these considerations can be found in Table 1.

Assessment of patient characteristics

Age for both participants and non-participants was defined as the age at which study information was received. For the DMDbrain study this exact date was unavailable for 30 subjects, resulting in a maximal uncertainty of nine months. Travel-time to the LUMC was derived with registered postal codes from the DDD, using 'www.google.nl/maps' and setting the date and time at a Monday in June 2019 outside rush hour.

Table 1. Definitions and examples of considerations for non-participation

Consideration	Definition	Examples
Burden of protocol	All characteristics of the protocol, except the MRI, that could lead to a burden for patients or parents.	*Amount of time participation would cost patients or parents, needing to take time off from school or work. *Amount of physical energy the study would cost. *Psychological stress patients and parents already had due to the disease to which participation would add. *Behavioral difficulties that would lead to strain of patients and parents when taking part in the study. *Stress that neuropsychological testing might cause due to the potential diagnosis of a cognitive impairment.
Travel-time	The time needed for traveling.	*Travel-time needed when for instance private wheelchair transportation was used.
Burden of clinical care	Number of tests and visits already required for clinical care to which the study would add.	*Number of tests, cardiac MRIs and visits already required for clinical care which made some patients not want to undergo an extra MRI or go to the hospital for an extra visit for the study. *In the Netherlands patients visit the hospital one average once a year, and upon publication of the revised standards of care in 2018, a cardiac MRI after the age of ten has been added to this.
Other research	Patients already having participated in previous or participating in current interventional trials or observational studies.	*Previous and current interventional trials. *Previous observational studies including MRI studies.
No advantage	Patient or parent did not want to participate due to a lack of potential personal benefit for the patient.	
Not interested	Patient was not interested to participate in research, but gave no further specification.	
MRI	Patient not wanting to undergo an MRI or a predicted difficulty for the patient to conform to the MRI protocol.	*Predicted difficulty for the patient to lie still or maintain the supine (DMDbrain and DMDperfusion) or lateral MRI position (DMDarm). *If patients gave 'MRI' as consideration because they did not want to undergo an MRI, it was asked whether patients had previously undergone an MRI.

All DMD gene mutations were derived from the DDD. For DMD, the mutation locations within the DMD gene predicted the absence of the following dystrophin isoforms in the brain: absence of only Dp427 (mutation in exon 1-44), absence of Dp427 and possibly Dp140 (mutation in exon 45-50), absence of Dp427 and Dp140 (mutation in exon 51-62), and absence of Dp427, Dp140 and Dp71 (mutation in exon 63-79).[7] The same locations were used to group mutations in the BMD patients although a similar prediction of isoform expression cannot be made.

At the time of registration in the DDD, patients or their caregivers had received a general questionnaire about their disease, including a question concerning comorbidities. This self-reported neurological and psychiatric comorbidity was used in the analysis for the DMDbrain and BMD studies only, as these studies started less than five years after most patients registered in the DDD. Age at loss of ambulation was derived from the DDD and included in the analysis for the DMDarm and BMD studies only, as these were the studies assessing motor performance.

Statistical analysis

For each study, the considerations for non-participation were summed per consideration group and adjusted for the total number of participants from whom a consideration was recorded per study. To compare the incidences of the different considerations for non-participation over all studies, percentages from the different studies were averaged to get an overall percentage. The three considerations with the highest overall percentage are reported here.

Age, travel-time and age at loss of ambulation (for DMDarm and BMD), were compared between participants and non-participants per study using the Mann-Whitney *U* test. Presence of a distal mutation upstream of exon 51 or 63 was compared between participants and non-participants per study using the Fisher's exact test. Statistical significance was set at $p<0.05$. Bonferroni-Holm correction was used to correct for the multiple comparisons of presence of a distal mutation upstream of exon 51 or 63 within each observational study.

Results

Participation

After pre-screening for age and diagnosis using the DDD, the patients who were reached by phone were registered per study as the first quantifiable step in the inclusion process (Table 2). The participation rate for the different studies was 35.4% (n=35) for DMDbrain, 40.0% (n=12) for DMDperfusion, 21.6% (n=22) for DMDarm, and 48.6% (n=36) for the BMD study. Characteristics of participants and non-participants for each study are shown in Table 3.

Table 2. Participation in the four observational studies

	Eligible and reached by phone	Did not meet inclusion criteria	Inclusion criteria met	Did not want to participate	Participated	% of patients willing to participate
DMDbrain	116	17	99	64	35	35.4%
DMDperfusion	44	14	30	18	12	40.0%
DMDarm	122	20	102	80	22	21.6%
BMD	92	19	74	38	36	48.6%

Table 3. Characteristics of participants and non-participants of the four observational studies

	DMDbrain (n=35/n=64)	DMDperfusion (n=12/n=18)	DMDarm (n=22/n=80)	BMD (n=36/n=38)
Age at study information, years				
participants	12 (10-15)	10.6 (10.1-11.8)	13.2 (12.1-16.1)*	42.3 (31.5-52.4)
non-participants	13 (11-15)	11.3 (10.4-12.7)	16.1 (13.2-20.4)*	42.5 (33.7-54.4)
Travel-time, minutes				
participants	65 (35-85)	35 (22-76)*	73 (34-86)	48 (30-80)
non-participants	63 (40-87)	73 (40-110)*	65 (35-85) n=75	65 (45-90)
Age at loss of ambulation, years				
participants	Not recorded	Not applicable	11.5 (10.1-13.1)	32 (22-41) n=6
non-participants			11.0 (9.1-12.5) n=72	13 (11-42) n=13

Data are median (1st quartile; 3rd quartile). After the study acronym the total number of participants and non-participants is shown as follows: (participants/non-participants). In case of missing data the number of patients for whom the data was available was presented after the result with n=number. *= p-value <0.05 for difference between participants and non-participants per study.

Considerations for non-participation

Considerations for not participating in the four observational studies are shown in Figure 1 and underlying data are given in Supplementary table 1, 2 and 3. The number of patients for whom a consideration for non-participation was recorded differed per study and was 44 (68.8%) for DMDbrain, 17 (94.4%) for DMDperfusion, 80 (100.0%) for DMDarm, and 29 (76.3%) for the BMD study. The consideration 'Burden of protocol' was provided most often by a mean of 37.9% (range 11.8-55.2%) of responders. This was followed by 'MRI' reported by 30.0% (range 23.8-41.2%) averaged over the three MRI studies, and 'Travel-time' reported by 19.0% (range 9.1-26.3%). Regarding the three MRI studies, 40.5% gave 'MRI' as a reason to decline participation because of a predicted difficulty for the participant to conform to the MRI protocol, while 59.5% gave this consideration because they did not want to undergo an MRI. Of this last group, 62.5% noted that this was due to a previous MRI experience, which had either been scary, unpleasant or long. The consideration 'No advantage' was not provided in the DMDbrain and BMD studies, and only by 2.5% of the non-participants in the DMDarm and 5.9% in the DMDperfusion studies.

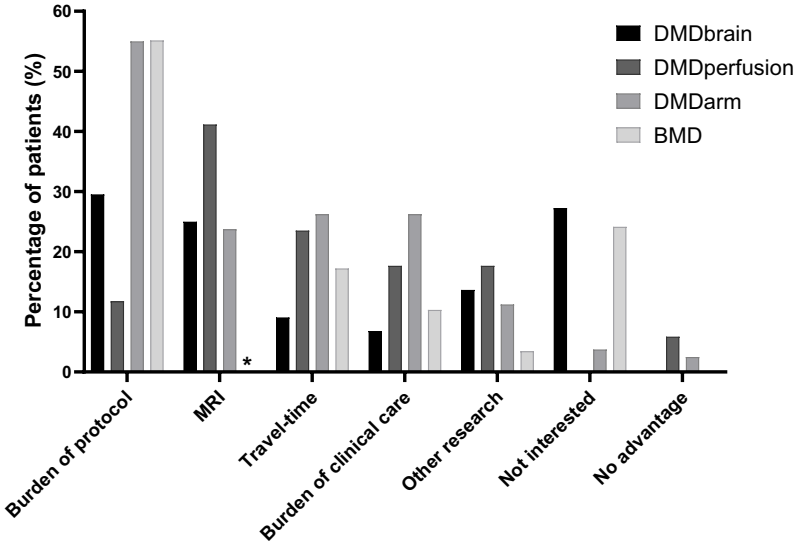


Figure 1. Considerations for not participating in the four observational studies
The presented percentage of non-participants who provided a consideration is adjusted for the percentage of patients for whom a consideration for non-participation was recorded: 68.8% for DMDbrain (black), 94.4% for DMDperfusion (dark gray), 100.0% for DMDarm (gray), and 76.3% for the BMD study (light gray). * MRI was an optional part of the BMD study.

The number of participants who provided more than one consideration differed per study and was 3 for DMDbrain (6.8%), 3 for DMDperfusion (17.7%), 34 for DMDarm (42.5%) and 3 for the BMD study (10.3%). The most often occurring combinations of two considerations were: 'Burden of the protocol' and 'Travel-time' (n=19), 'Burden of the protocol' and 'Burden of clinical care' (n=10), and 'Burden of the protocol' and 'MRI' (n=6).

Patient characteristics

Age at receipt of study information was comparable between participants and non-participants for all studies except the DMDarm study, where non-participants were 2.9 years older (p=0.012; Table 3).
Travel-time only differed between participants and non-participants in the DMDperfusion study (p=0.016; Table 3), where it was 38 minutes longer for non-participants. There was missing travel-time data for five DMDarm non-participants who had not registered in the DDD.

The presence of the different mutations in participants and non-participants of all studies is presented in Figure 2. While there were some patients with exon 63-79 mutations in the non-participant group of all studies, no patients with this mutation participated in any of the studies. However, the proportion of exon 63-79 mutations or exon 51-79 mutations did not differ significantly between participants and non-participants for any of the studies.

There were missing mutation data for some non-participants: one (5.6%) from DMDperfusion whose mutation was not registered in the DDD, five (6.3%) from DMDarm who had not registered in the DDD, and one (2.6%) from the BMD study in whom the diagnosis had been based on a muscle biopsy only.

Neurological and psychiatric comorbidity was self-reported by two participants (5.7%) or their caregivers in the DMDbrain study. For one patient this was autism spectrum disorder (ASD), intellectual disability, and oppositional defiant disorder (ODD), and for the other patient this was attention deficit hyperactivity disorder (ADHD). In the DMDbrain non-participant group, such comorbidity was self-reported by five patients (7.8%) or their caregivers: for one patient this was ASD, for two patients ASD and intellectual disability, for one patient ASD and attention deficit disorder (ADD), and for one patient ADHD. In the BMD

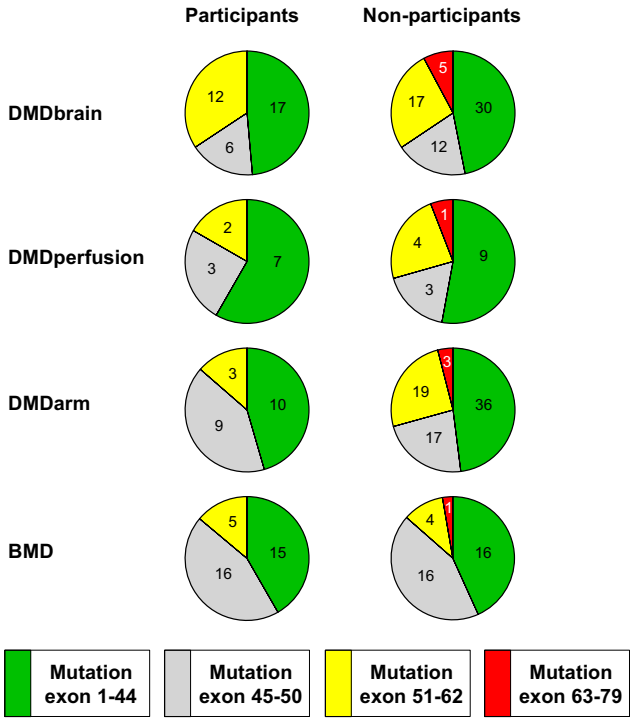


Figure 2. Position of the mutation within DMD gene in participants and non-participants of the four observational studies
In DMD, this predicts the following isoforms to be absent: exon 1-44 mutations affect Dp427, exon 45-50 mutations affect Dp427 and possibly Dp140, exon 51-62 mutations affect Dp427 and Dp140, and exon 63-79 mutations affect Dp427, Dp140 and Dp71. The same locations were used to group mutations in the BMD patients although a similar prediction of isoform expression cannot be made. The numbers in the pie charts represent the number of participants and non-participants in the different studies that have a certain mutation.

study, similar comorbidity was self-reported by one participant (2.8%) as ADHD, while in the non-participants group five patients (13.2%) reported this: two reported ASD, two ADHD, and one dyslexia.

In the DMDarm study, age at loss of ambulation was comparable between participants and non-participants (p=0.231; Table 3). For the BMD study, loss of ambulation had occurred in six (16.7%) participants and 14 (36.8%) non-participants. Here too, age at loss of ambulation was comparable between participants and non-participants (p=0.333). Loss of ambulation data was missing for one non-participant from the BMD study and eight non-participants from the DMDarm study, of whom five had not registered in the DDD and three had not yet lost ambulation at the time of the DDD questionnaire.

Discussion

In this study, we assessed which considerations played a role in decision-making for participation in four observational studies in DMD and BMD. We found that the following three considerations for not participating were most often provided: 'Burden of protocol', 'MRI', and 'Travel-time'. Additionally, we compared patient characteristics between participants and non-participants, and showed that the included cohorts were representative for the currently studied variables, except for a younger age of participants in the DMDarm study and a lack of patients with the most distal mutations.

An important example of protocol burden was the amount of time that participation would cost patients or parents, needing to take time off from school or work. This was also underscored by a previous observational study in BMD patients that reported lack of time as main reason for non-participation.[5] The four studies at our institute mainly took place on week days, mostly outside of school hours. Many participating patients, their parents, and the local ethical committee also expressed that they preferred holidays and weekends, while other patients and parents actually preferred schooldays. Other examples of protocol burden in our studies were the amount of physical energy the study costs, extra psychological stress due to participation, and behavioral difficulties of patients that lead to extra strain when taking part. Both examples of study burden and the timing of study days could be ameliorated by obtaining the patient and parent perspective on the study design and its feasibility early-on in the protocol development. This is increasingly being advocated by many stakeholders in the field,[19-22] and was recently summarized in a report by the European Neuromuscular Centre (ENMC).[23]

A previous MRI experience was mentioned by a relatively large proportion of patients as the reason to decline participation. Examples were "a painful cardiac MRI", "MRI was too long and very noisy", "stress before the MRI due to ASD", and "scared to fall off the MRI

table", highlighting that both the duration and the patient experience play a role. The duration of the MRI protocol can be reduced by scan acceleration techniques and combining different scan contrasts into one acquisition. While 30-45 minutes of MRI was sometimes considered as long by our non-participants, another large MRI study showed that yearly MRI sessions of 75-90 minutes in pediatric DMD patients are possible.[24] Understanding differences between these MRI studies and implementing corresponding adjustments could improve the MRI experience. To this end, more international collaboration and exchange of not only protocols, but also personal experience is needed. MRI vendors are developing methods to improve patient comfort as well, such as up to 99% sound reduction, calming visual themes projected on the MRI, using wider bores, and more comfortable coils.[25, 26] In research, stress reduction by showing videos, having a parent present in the MRI room during the scan, and using a mock scanner beforehand is already used. In regular clinical practice, this is more difficult due to time and budget constraints, potentially leading to negative subjective experiences. It is thus essential to dedicate specific attention to these vulnerable and rare patient categories when performing assessments in clinic.

'Travel-time' was also an important consideration for non-participation. While we reported travel-times outside of the rush hour for consistency and these may seem low, participants were often unable to avoid the busy rush hour of the Netherlands and this could cause the actual travel-times to be twice as high. For the DMDperfusion study, travel-times were also longer for non-participants than participants, which supports that travel-times influenced the decision to participate. In some observational studies the travel-time and time cost of participation has been minimized by performing these studies during regular visits as part of the outpatient clinical care.[27-30] This can both reduce the burden of research and limit the number of visits to the hospital, and should therefore be explored for all future studies.

'Other research' was mentioned more often as consideration for non-participation in our DMD studies (13.6% in DMDbrain, 17.7% in DMDperfusion, 11.3% in DMDarm) compared to our BMD study (3.5%). At the time of inclusion a much larger number of studies had been performed or were ongoing in DMD compared to BMD, which is likely the reason for this difference. This also could have influenced the lower participation rates in the DMD studies compared to BMD.
Interestingly, the lack of personal gain was only rarely reported in our studies (2.5% in DMDarm, 5.9% in DMDperfusion). This is in contrast to the study by Peay et al. where an online survey was used to assess barriers for parents to have their child with DMD or BMD for the first time participate in an interventional trial.[12] Since most parents wanted their child to participate, only the barrier "my child could receive placebo" was deemed more true than untrue on a Likert-type scale. Personal gain therefore seems more important in the decision to participate in an interventional trial than an observational study.

Representativeness of included cohorts is of utmost importance in any study, both interventional and observational. Details on this are often lacking because most studies extensively describe the included, but not the excluded patients in primary tables. Although screening logs are getting more comprehensive due to GCP regulations, data are hardly ever analyzed or published. In our studies, assessment of selection bias was possible because of the DDD registry,[3] as this contained characteristics of the non-participants. While participants in the DMDarm study were 2.9 years younger than the non-participants, age at loss of ambulation, as a proxy for rate of disease progression, was comparable. Therefore, participants were probably less progressed at the time of inclusion. While age and disease progression were taken into account in the design and analyses of this study, this bias could be problematic in a study targeting older and more severely affected patients. We also found that a few non-participants in all studies had an exon 63-79 mutation, while no patients with this rare mutation participated in any of the studies. This was not statistically significant, which could have been caused by the rarity of this very distal mutation. DMD patients with exon 63-79 mutations have absent Dp427, Dp140 and Dp71 in the brain and a higher occurrence of learning and behavioral disabilities.[7, 8] This could cause them to opt out of research in general or be unable to follow study specific instructions especially regarding the MRI protocols. To prevent selection bias, future observational studies in DMD and BMD should aim to include patients with exon 63-79 mutations, especially when assessing cognitive and behavioral aspects of the diseases.

A more detailed study of selection bias would be possible if registries contained more extensive information than contact details and clinical information that is provided on inclusion. We are currently improving the DDD with an optional yearly update of important clinical items via a short questionnaire, and a formal collaboration with patient organizations to ensure nationwide participation. Future Dutch studies can use this extensive data to study selection bias in more detail. Furthermore, every observational study and interventional trial should use a similar registry to analyze selection bias in their cohort and publish the results.

There are limitations to our study. Due to the retrospective study design, considerations for not participating were not equally available for all studies. Furthermore, only limited data was available on non-participants via the previous version of the DDD registry. Finally, we reported a much lower prevalence of neurologic and psychiatric comorbidity (2.8%-13.2%) compared to literature (up to 67% of BMD and 90% of DMD patients).[4, 6] As our results were self-reported at the time of registering in the DDD, any diagnoses made after that have not been automatically recorded. This underestimation supports more consistent screening and assessment of cognitive diagnoses, as well as the need for regular updates of registries.

In summary, we reviewed the considerations provided for not taking part in four DMD and BMD observational studies and found that 'Burden of protocol', 'MRI', and 'Travel-time' were most frequently reported. Participating patients were overall representative of the studied

sub-populations, except for age in the DMDarm study which may point to the challenge of studying more advanced stages of these conditions and the lack of distal mutations upstream of exon 63.
Optimizing the involvement of patients while designing protocols, improving the MRI experience, and integrating observational research and clinical care are all factors that need to be addressed to facilitate and increase patient participation. Nationwide registries that enable the recruitment of patients are essential for the comparison of participants and non-participants to ensure that observational research is representative.

Study funding
The DMDbrain study was supported by Duchenne Parent Project Netherlands (DPP NL) and the Gratama Foundation (grant number 10.13). The DMDperfusion study was funded by DPP NL. The DMDarm study was supported by Stichting Spieren for Spieren (grant number SvS15). The BMD study was funded by the Netherlands Organization for Health Research and Development (ZonMw) (grant number 113302001).

References

1. Hoffman EP, Brown RH, Jr., Kunkel LM. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell 1987;51:919-928. doi:10.1016/0092-8674(87)90579-4
2. Bushby KM, Gardner-Medwin D. The clinical, genetic and dystrophin characteristics of Becker muscular dystrophy. I. Natural history. Journal of neurology 1993;240:98-104. doi:10.1007/BF00858725
3. van den Bergen JC, Ginjaar HB, van Essen AJ, et al. Forty-Five Years of Duchenne Muscular Dystrophy in The Netherlands. J Neuromuscul Dis 2014;1:99-109
4. Banihani R, Smile S, Yoon G, et al. Cognitive and Neurobehavioral Profile in Boys With Duchenne Muscular Dystrophy. J Child Neurol 2015;30:1472-1482. doi:10.1177/0883073815570154
5. Mori-Yoshimura M, Mizuno Y, Yoshida S, et al. Psychiatric and neurodevelopmental aspects of Becker muscular dystrophy. Neuromuscul Disord 2019;29:930-939. doi:10.1016/j.nmd.2019.09.006
6. Young HK, Barton BA, Waisbren S, et al. Cognitive and psychological profile of males with Becker muscular dystrophy. J Child Neurol 2008;23:155-162. doi:10.1177/0883073807307975
7. Taylor PJ, Betts GA, Maroulis S, et al. Dystrophin gene mutation location and the risk of cognitive impairment in Duchenne muscular dystrophy. PLoS One 2010;5:e8803. doi:10.1371/journal.pone.0008803
8. Pane M, Lombardo ME, Alfieri P, et al. Attention deficit hyperactivity disorder and cognitive function in Duchenne muscular dystrophy: phenotype-genotype correlation. J Pediatr 2012;161:705-709 e701. doi:10.1016/j.jpeds.2012.03.020
9. Verhaart IEC, Aartsma-Rus A. Therapeutic developments for Duchenne muscular dystrophy. Nature reviews Neurology 2019;15:373-386. doi:10.1038/s41582-019-0203-3
10. Straub V, Balabanov P, Bushby K, et al. Stakeholder cooperation to overcome challenges in orphan medicine development: the example of Duchenne muscular dystrophy. Lancet Neurol 2016;15:882-890. doi:10.1016/S1474-4422(16)30035-7
11. CDER, CBER. Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment. Guidance for Industry. https://www.fda.gov/media/92233/download. Accessed on May 1, 2019. US Food & Drug Administration (FDA) 2018
12. Peay HL, Biesecker BB, Wilfond BS, et al. Barriers and facilitators to clinical trial participation among parents of children with pediatric neuromuscular disorders. Clin Trials 2018;15:139-148. doi:10.1177/1740774517751118
13. Tromp K, Zwaan CM, van de Vathorst S. Motivations of children and their parents to participate in drug research: a systematic review. Eur J Pediatr 2016;175:599-612. doi:10.1007/s00431-016-2715-9
14. Akmatov MK, Jentsch L, Riese P, et al. Motivations for (non)participation in population-based health studies among the elderly - comparison of participants and nonparticipants of a prospective study on influenza vaccination. BMC Med Res Methodol 2017;17:18. doi:10.1186/s12874-017-0302-z
15. Doorenweerd N, Dumas EM, Ghariq E, et al. Decreased cerebral perfusion in Duchenne muscular dystrophy patients. Neuromuscul Disord 2017;27:29-37. doi:10.1016/j.nmd.2016.10.005
16. Doorenweerd N, Hooijmans M, Schubert SA, et al. Proton Magnetic Resonance Spectroscopy Indicates Preserved Cerebral Biochemical Composition in Duchenne Muscular Dystrophy Patients. J Neuromuscul Dis 2017;4:53-58. doi:10.3233/JND-160201
17. Doorenweerd N, Straathof CS, Dumas EM, et al. Reduced cerebral gray matter and altered white matter in boys with Duchenne muscular dystrophy. Ann Neurol 2014;76:403-411. doi:10.1002/ana.24222
18. Straathof CS, Doorenweerd N, Wokke BH, et al. Temporalis muscle hypertrophy and reduced skull eccentricity in Duchenne muscular dystrophy. J Child Neurol 2014;29:1344-1348. doi:10.1177/0883073813518106
19. Lochmuller H, Torrent IFJ, Le Cam Y, et al. The International Rare Diseases Research Consortium: Policies and Guidelines to maximize impact. Eur J Hum Genet 2017;25:1293-1302. doi:10.1038/s41431-017-0008-z
20. EURORDIS Open Academy. https://openacademy.eurordis.org/. Accessed on March 26, 2020. [online].
21. Witteman HO, Chipenda Dansokho S, Colquhoun H, et al. Twelve Lessons Learned for Effective Research Partnerships Between Patients, Caregivers, Clinicians, Academic Researchers, and Other Stakeholders. J Gen Intern Med 2018;33:558-562. doi:10.1007/s11606-017-4269-6
22. 'Patiëntenparticipatie' (patient participation). https://subsidieportaal.spierfonds.nl/nl/soorten%20subsidies. Accessed on March 26, 2020. [online].
23. Lochmuller H, Ambrosini A, van Engelen B, et al. The Position of Neuromuscular Patients in Shared Decision Making. Report from the 235th ENMC Workshop: Milan, Italy, January 19-20, 2018. J Neuromuscul Dis 2019;6:161-172. doi:10.3233/JND-180368

24. Forbes SC, Walter GA, Rooney WD, et al. Skeletal muscles of ambulant children with Duchenne muscular dystrophy: validation of multicenter study of evaluation with MR imaging and MR spectroscopy. Radiology 2013;269:198-207. doi:10.1148/radiol.13121948
25. MRI in-bore experience. https://www.philips.nl/healthcare/educatie/technologies/mri/mri-in-bore-experience. Accessed on March 26, 2020. [online].
26. MRI patient experience. https://www.siemens-healthineers.com/magnetic-resonance-imaging/patient-experience. Accessed on April 7, 2020. [online].
27. Pane M, Mazzone ES, Sivo S, et al. Long term natural history data in ambulant boys with Duchenne muscular dystrophy: 36-month changes. PLoS One 2014;9:e108205. doi:10.1371/journal.pone.0108205
28. Goemans N, Vanden Hauwe M, Signorovitch J, et al. Individualized Prediction of Changes in 6-Minute Walk Distance for Patients with Duchenne Muscular Dystrophy. PLoS One 2016;11:e0164684. doi:10.1371/journal.pone.0164684
29. Wong BL, Rybalsky I, Shellenbarger KC, et al. Long-Term Outcome of Interdisciplinary Management of Patients with Duchenne Muscular Dystrophy Receiving Daily Glucocorticoid Treatment. J Pediatr 2017;182:296-303 e291. doi:10.1016/j.jpeds.2016.11.078
30. Muntoni F, Domingos J, Manzur AY, et al. Categorising trajectories and individual item changes of the North Star Ambulatory Assessment in patients with Duchenne muscular dystrophy. PLoS One 2019;14:e0221097. doi:10.1371/journal.pone.0221097

Supplementary material

All data on considerations for not participating in the four observational studies

Supplementary table 1. Number of non-participants for whom a particular consideration was recorded

	Total number of non-participants for whom a consideration was recorded	Burden of protocol	MRI	Travel-time	Burden of clinical care	Other research	Not interested	No advantage
DMDbrain	44	13	11	4	3	6	12	0
DMDperfusion	17	2	7	4	3	3	0	1
DMDarm	80	44	19	21	21	9	3	2
BMD	29	16		5	3	1	7	0
Total	170	75	37	34	30	19	22	3

Supplementary table 2. Percentage of non-participants for whom a particular consideration was recorded, corrected for the percentage for whom a consideration was recorded in that study

	% of non-participants for whom a considerations was recorded	Burden of protocol	MRI	Travel-time	Burden of clinical care	Other research	Not interested	No advantage
DMDbrain	68.8%	**29.6%**	25.0%	9.1%	6.8%	13.6%	27.3%	0.0%
DMDperfusion	94.4%	**11.8%**	41.2%	23.5%	17.7%	17.7%	0.0%	5.9%
DMDarm	100.0%	**55.0%**	23.8%	26.3%	26.3%	11.3%	3.8%	2.5%
BMD	80.6%	**55.2%**		17.2%	10.3%	3.5%	24.1%	0.0%
Corrected percentages of considerations for non-participation averaged over the four studies	85.9%	**37.9%**	**22.5%**	**19.0%**	15.3%	11.5%	13.8%	2.1%
Corrected percentage of non-participants for whom 'MRI' was recorded averaged over the three MRI studies			**30%**					

Supplementary table 3 Number and percentage of non-participants for whom the consideration 'MRI' was recorded studied in more detail

	Total number of non-participants for whom a consideration was recorded	MRI	Predicted difficulty for the patient to conform to the MRI protocol	% that is predicted to have difficulty to conform to the MRI protocol	Doesn't want MRI	% that doesn't want MRI	Doesn't want MRI due to previous MRI	% that doesn't want MRI due to previous MRI
DMDbrain	44	11	7	63.6%	4	36.4%	1	25.0%
DMDperfusion	17	7	0	0.0%	7	100.0%	7	100.0%
DMDarm	80	19	11	57.9%	8	42.1%	5	62.5%
Total or average %	141	37	18	40.5%	19	59.5%	13	62.5%

Chapter 3

MRI vastus lateralis fat fraction predicts loss of ambulation in Duchenne muscular dystrophy

Published in: Neurology
Mar 2020; 94; e1386-e1394; DOI: 10.1212/WNL.0000000000008939

Karin J. Naarding | Harmen Reyngoudt | Erik W. van Zwet | Melissa T. Hooijmans
Cuixia Tian | Irina Rybalsky | Karen C. Shellenbarger | Julien Le Louër | Brenda L. Wong
Pierre G. Carlier | Hermien E. Kan, | Erik H. Niks

Abstract

Objective

We studied the potential of quantitative MRI (qMRI) as a surrogate endpoint in Duchenne muscular dystrophy by assessing the additive predictive value of vastus lateralis (VL) fat fraction (FF) to age on loss of ambulation (LoA).

Methods

VL FFs were determined on longitudinal Dixon MRI scans from two natural history studies in Leiden University Medical Center (LUMC) and Cincinnati Children's Hospital Medical Center (CCHMC). CCHMC included ambulant patients, while LUMC included a mixed ambulant and non-ambulant population. We fitted longitudinal VL FF values to a sigmoidal curve using a mixed model with random slope to predict individual trajectories. The additive value of VL FF over age to predict LoA was calculated from a Cox model, yielding a hazard ratio.

Results

Eighty-nine MRIs of 19 LUMC and 15 CCHMC patients were included. At similar age, 6-minute walking test distances were smaller and VL FFs were correspondingly higher in LUMC compared to CCHMC patients. Hazard ratio of a percent-point increase in VL FF for the time to LoA was 1.15 for LUMC (95% confidence interval [CI] 1.05–1.26; $p = 0.003$) and 0.96 for CCHMC (95% CI 0.84–1.10; $p = 0.569$).

Conclusions

The hazard ratio of 1.15 corresponds to a 4.11-fold increase of the instantaneous risk of LoA in patients with a 10% higher VL FF at any age. Although results should be confirmed in a larger cohort with prospective determination of the clinical endpoint, this added predictive value of VL FF to age on LoA supports the use of qMRI FF as an endpoint or stratification tool in clinical trials.

Introduction

Duchenne muscular dystrophy (DMD) is characterized by progressive replacement of muscle tissue with fat and fibrosis due to absence of full-length dystrophin.[1] Although the first drugs have now received regulatory approval, there is still a medical need with many ongoing and planned clinical trials.[2] Such trials are challenging in the pediatric population because of the rarity of the disease, a different rate of progression in different patients, and clinical endpoints that can be influenced by patient motivation.[3] Objectively quantified and predictive surrogate endpoints could overcome these limitations, but only if a clear interdependence between the outcome measure and clinically meaningful milestones is demonstrated.

Both scientists and regulatory agencies consider muscle fat fraction (FF), measured by quantitative MRI (qMRI) or magnetic resonance spectroscopy (MRS), as a potential surrogate endpoint in trials.[4-7] qMRI of the lower extremity can non-invasively, objectively and accurately assess muscle FF in DMD, and is reproducible.[8-12] Longitudinal DMD studies demonstrated a sigmoidal increase in FF in leg muscles.[13, 14] The FF of the vastus lateralis (VL) muscle has been shown to have a large effect size in detecting 1-year change compared to other leg muscles.[8]

Cross-sectional correlations between strength, function and muscle FF have been described.[4, 5, 9, 15-20] However, because FF increases with age, it will always correlate with the declining functional parameters in DMD. Therefore, a simple correlation alone will not suffice. In this study, we show that VL FF has additive predictive value to age on loss of ambulation (LoA).

Methods

Participants and study design

Patients with DMD participated in natural history studies at Leiden University Medical Center (LUMC), the Netherlands, or at Cincinnati Children's Hospital Medical Center (CCHMC), Ohio, USA. The LUMC patients were recruited from the Dutch Dystrophinopathy Database.[21] Selected MRI results of this study have been previously reported.[13, 22, 23] The CCHMC participants were recruited from the international Prospective Natural History Study of Progression of Subjects With Duchenne Muscular Dystrophy (PRO-DMD-01) that started in 2012. For this work, LUMC, CCHMC, and the Institute of Myology collaborated as part of the BIOIMAGE-Neuromuscular Diseases (BIOIMAGE-NMD) consortium (project identifier 602485, funded under FP7-HEALTH). Inclusion criteria at both LUMC and CCHMC were a confirmed genetic mutation in the *DMD* gene, and being ≥5 years of age. In addition, at CCHMC, patients had to have a mutation that would be amenable to skipping of exon 44, 45, 51, 52, 53 and 55 and had to be able to walk at least 75 m unassisted in the 6-minute walking test (6MWT); these constraints were not applied in the selection of LUMC patients. Main exclusion criteria at both LUMC and CCHMC were the presence of contra-indications for MRI and participation

in a clinical study with an investigational medicinal product. MRI examinations took place at LUMC at baseline and 12, 24 and 30 months between August 2013 and December 2016; at CCHMC, examination took place at baseline and 6, 12 and 18 months between January 2015 and August 2016.

Standard protocol approvals, registrations, and patient consents

The local ethics committee at each site approved the study conducted at that site. Written informed consent was obtained from patients and parents. The PRO-DMD-01 natural history study was registered under the following clinical trial identifier number: NCT01753804.

Determination of clinical endpoint and cohort characteristics

We defined LoA as the patient being unable to walk 5 m without assistance or orthoses. If LoA was established during ongoing yearly follow-up, we used the month and year of LoA registered in clinical documentation. In addition, if there was recent clinical documentation that the patient was still ambulant, we used the date of that documentation as the last follow-up. When it was unknown whether patients were still ambulant or exact month and year of LoA had not been registered, detailed interviews with patients and parents were conducted by telephone between July 2017 and July 2018. For those still ambulant patients, we defined the last interview date as last follow-up. History of corticosteroid use was also established using clinical documentation or assessed during this interview. 6MWT data were derived from the natural history study visits. For LUMC patients, data from outpatient visits before and after the natural history study were added.

MRI acquisition

At both sites, position of the patients was feet first, supine. A 16-channel anterior array receive coil was used in combination with a 12-channel array receive coil that was located within the table. FFs were determined from 3-point gradient echo Dixon images of the thigh. At the LUMC, images of the right thigh were acquired on a 3T MRI scanner (Ingenia, Philips Healthcare, Best, The Netherlands), as described previously.[23] Dixon scans were acquired with 23 slices, a voxel size of 1 x 1 x 10 mm, and an interslice gap of 5 mm (repetition time [TR]/echo time [TE]/echo time shift [ΔTE] 210/4.41/0.76 milliseconds, flip angle 8°). At CCHMC, images of both thighs were acquired on a 1.5T MRI scanner (Ingenia). Scans consisted of 35 partially overlapping slices with a voxel size of 1 x 1 x 10 mm (TR/TE/ΔTE 11.25/2.4/2.3 milliseconds, flip angle 3°). At both sites, sedation was not necessary, and while the entire scan protocol, including other scan types and lower leg scans, took up to 1 hour, we could acquire the Dixon scan within ten minutes, including planning and positioning.

MRI analysis

Data were reconstructed using the manufacturer's software assuming a single peak in the lipid spectrum and without T2* relaxation correction. Recent studies have shown differences in fat replacement along the proximodistal axis of the muscle in DMD,[23, 24] which necessitated precise definition of the region of interest (ROI) along this axis for comparison of MRI data

from both sites. Two observers from LUMC (K.J.N.) and the Institute of Myology (H.R.) determined in consensus the most proximal slice where the biceps femoris short head was still visible.[17] This slice was defined as the center slice around which multiple slices covering 70 mm of the VL muscle were analyzed. The observers also determined in consensus which scans had to be excluded because of major movement artefacts, water/fat swaps, or other artefacts in the VL. The observers then independently drew ROIs of the VL muscle using Medical Image Processing, Analysis and Visualization software (www.mipav.cit.nih.gov). ROIs were drawn on five consecutive slices for LUMC and on every second slice for CCMHC for a total of seven slices around the center slice. The boundaries of the ROIs were drawn exactly on the muscle border. Next, we performed an inward erosion of 2 mm for every ROI to avoid contamination of ROIs with subcutaneous fat and fatty intermuscular septa (Figure 1). FF values were calculated per slice as signal intensity (SI) fat/(SI fat + SI water) × 100 from the reconstructed fat and water images. VL FFs were calculated as a weighted mean value based on the number of VL pixels per slice. To correct for differences in TR and flip angle, FF values were corrected for field strength-specific T1 partial saturation effects using literature values for muscle and fat tissue.[25]

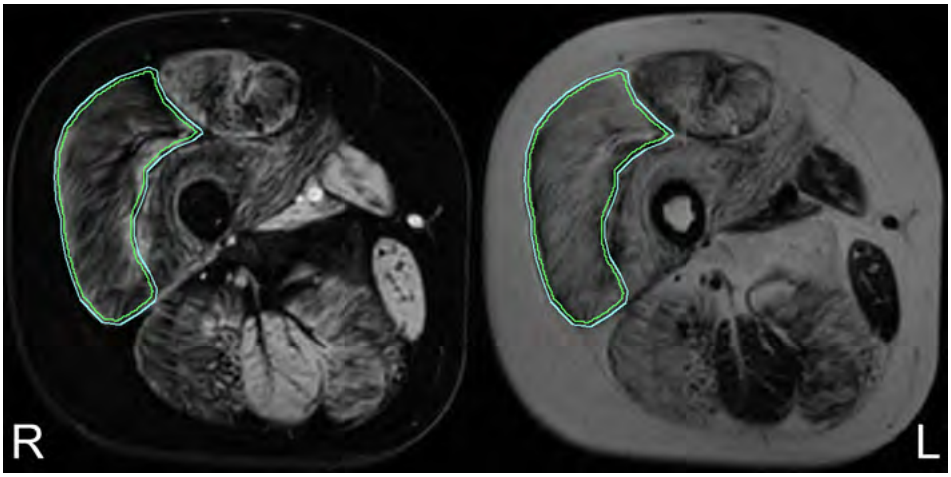


Figure 1. VL ROI
Example of a region of interest (ROI) drawn on the vastus lateralis (VL) (outer line) and the 2 mm inward erosion (inner line) on a water image (left) and corresponding fat image (right).

Statistical analysis

We assessed agreement between VL FFs from the two observers using an intraclass correlation coefficient (ICC) with a 2-way random model and absolute agreement.[26] Bland-Altman analysis was also performed to determine bias and limits of agreement between both observers. We modeled the VL FFs, calculated the hazard ratios, did a Spearman correlation, and generated the growth charts for the two cohorts separately because of the differences in inclusion criteria.

As a percentage, VL FF is bound between values of 0 and 100. Bounded outcomes often have nonstandard probability distributions that make their statistical analysis complex. However, when we apply the logit transformation to FF, the range of the new variable (Y) becomes unbounded, and we can use standard statistical methods that rely on the normal (gaussian) distribution:

$$\mathrm{Y} = \ln\left(\frac{\mathrm{FF}}{100 - \mathrm{FF}}\right)$$

To account for the correlation between the longitudinal FF measurements for each patient, we fitted a linear (mixed) model to our outcome Y with age as the only covariate and a random slope per individual. After transformation back to the original scale by application of a logistic transformation, VL FFs of all patients at any time had been predicted.

To calculate the additive predictive value of VL FF to age on LoA, we fitted a Cox proportional hazards model with predicted VL FF as a time-varying covariate. This model defines time intervals for the entire follow-up on the basis of the ages at which an event occurs, that is, either LoA or the end of follow-up. Per time interval, it links the predicted VL FF to the ambulant status at the end of that interval for each patient. An advantage of this method is that LoA events that took place before the first MRI had been acquired (i.e., in patients who were non-ambulant at baseline) and all patients with ≥1 VL FF data points can be included in the analysis. The hazard ratio from the model is then calculated as follows:

$$\text{Hazard ratio} = \mathrm{e}^{(loghazard * 1)}$$

For this, a value of $p < 0.05$ tested using the Wald test was considered significant. Next, the increase in instantaneous risk of LoA caused by the number of percent-points increase in VL FF (ΔFF) can be calculated as follows:

$$\text{Increase in instantaneous risk of LoA} = \mathrm{e}^{(loghazard * \Delta\mathrm{FF})}$$

We performed Spearman correlation analysis to assess the relationship between age at LoA and the nonnormally distributed individual slope of the predicted VL FF curve. On the basis of the predicted VL FF curves with a normally divided set of slopes, we generated a growth chart with differing slopes for patients at different percentiles of the disease spectrum. Using the hazard ratio, we transformed the predicted VL FF growth curves to survival curves for preserved ambulation, in which, for example, a patient on the third percentile in the VL FF growth chart is also on the third percentile in the survival chart.

Data availability

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

Results

Data inclusion and cohort characteristics

Twenty-two patients from LUMC and 16 from CCHMC participated in both natural history studies. For three patients from LUMC and one from CCHMC, no usable thigh MRI scan was available at any time point due to either an inability to complete the scan or movement artefacts. This led to the availability of 46 usable MRIs and therefore VL FF data points from one to four time points of 19 LUMC patients and 43 data points from one to four time points of 15 CCHMC patients. Figure 2 shows flowcharts of MRI data inclusion, and Table 1 presents characteristics for the LUMC and CCHMC cohorts. On average LUMC patients were younger and taller and had lower body mass index values than CCHMC patients. LUMC patients mostly used prednisone in an intermittent schedule and from a later age than CCHMC patients, who all used daily deflazacort. Furthermore, CCHMC patients were more often eligible for skipping of exon 44. Seven participants from LUMC with available VL FF data were non-ambulant at baseline, while all participants from CCHMC were ambulant as a result of the inclusion criteria for that study. During follow-up, LoA occurred in ten patients with available VL FF data, seven from LUMC and three from CCHMC. On average, LUMC patients lost ambulation at a younger age.

Table 1. Characteristics of both study cohorts

	LUMC (n=22)	CCHMC (n=16)
Characteristics at baseline		
Age, years	9.2 (7.4; 12.3)	11.2 (9.0; 12.5)
Mutation amenable to skipping of exon 44, n (%)	2 (9.1)	5 (31.3)
Corticosteroids, n (%)	18 (81.8)	16 (100.0)
Prednisone intermittent	17 (77.3)	0 (0.0)
Deflazacort daily	1 (4.5)	16 (100.0)
Age at start steroid use, years	5.9 (5.0; 7.7)[a]	4.3 (3.5; 6.2)
Height by age, SD[b]	-0.7 (-1.5; 0.3)[c]	-3.3 (-4.2; -2.7)[c]
Weight by height, SD[b]	1.3 (0.7; 2.4)[c]	2.6 (1.4; 3.9)[c]
Body mass index, kg/m^2	17.5 (16.4; 25.4)[c]	19.4 (17.3; 23.9)[c]
Non-ambulant, n (%)	9 (41.9)	0 (0.0)
Characteristics at follow-up		
Loss of ambulation during follow-up, n (%)	8 (36.4)	3 (18.8)
Age at loss of ambulation, years	10.8 (9.0; 12.1)	13.6; 14.5; 15.3
Age at last follow-up in still ambulant participants, years	12.2 (10.6; 13.7)	12.9 (11.5; 14.7)

Abbreviations: CCHMC = Cincinnati Children's Hospital Medical Center; LUMC = Leiden University Medical Center. Characteristics of all patients from the LUMC and CCHMC cohorts. For three LUMC and one CCHMC patient, there was no usable thigh MRI available at any time point. Values are median (first; third quartiles), number of patients (percent), or the actual values of all patients.

[a] One patient never used corticosteroids and therefore has no starting age.

[b] SDs calculated from the Dutch growth diagrams for height by age and weight by height for boys one to 21 years of age, which originate from the 2009 Fifth Dutch Growth Study.[43]

[c] Height of one patient from both LUMC and CCHMC could not be recovered.

Reliability of MRI parameters

The interobserver reliability for VL FF was excellent, with an ICC of 1.0 (95% confidence interval [CI] 1.0-1.0). Using the Bland-Altman analysis, we found a mean bias of 0.1% in VL FF with limits of agreement of -0.9 to 1.2%.

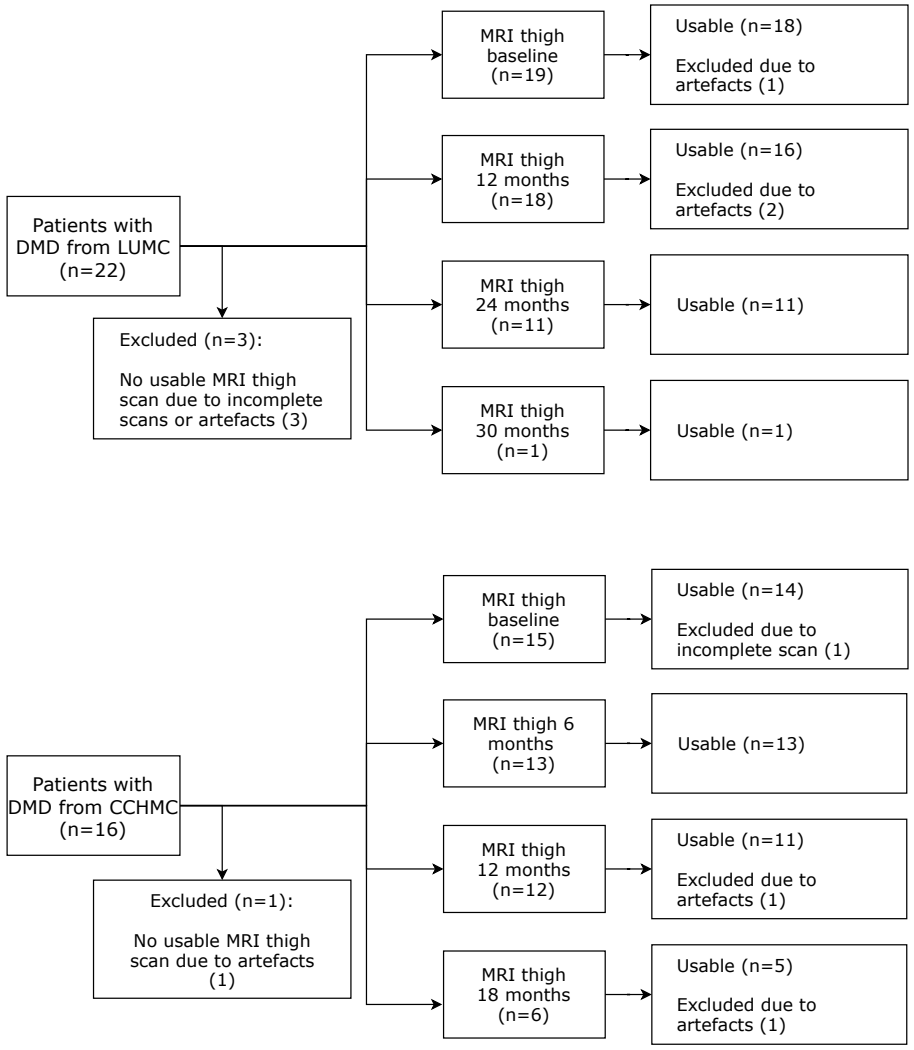


Figure 2. Flowchart of included thigh MRI datasets
Inclusion of patients with Duchenne muscular dystrophy (DMD) and thigh MRI scan data at Leiden University Medical Center (LUMC) and Cincinnati Children's Hospital Medical Center (CCHMC). Forty-six usable MRIs from one to four time points were available for 19 LUMC patients, and 43 usable MRIs from again one to four time points were available for 15 CCHMC patients.

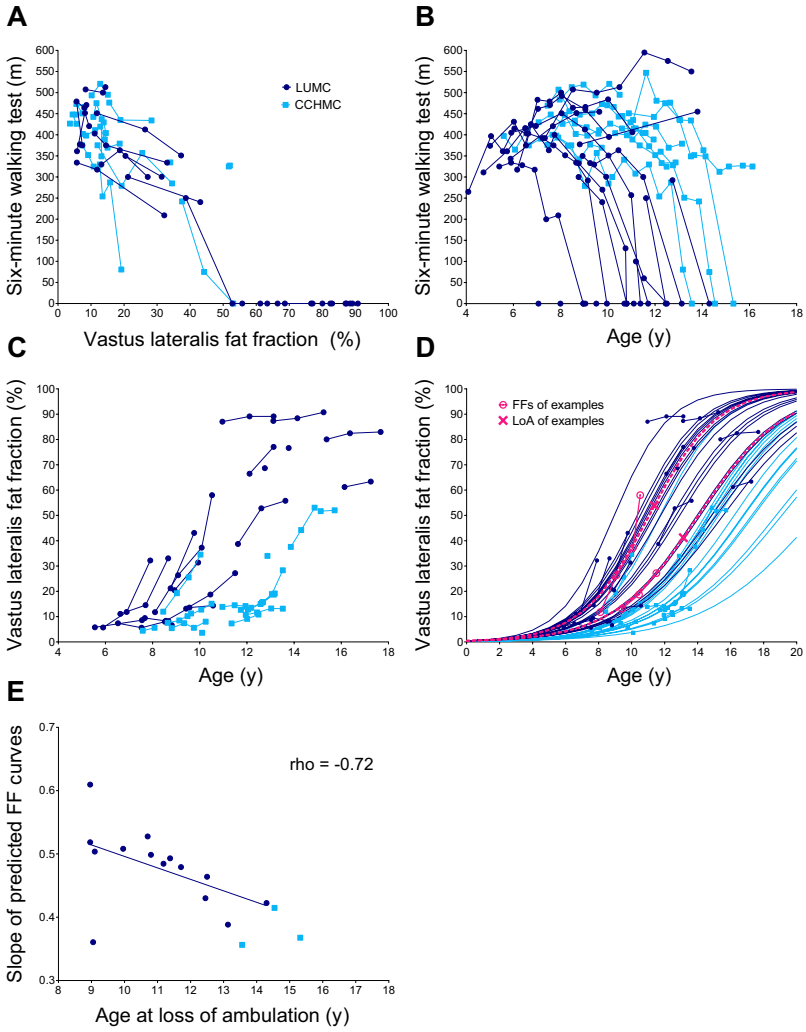


Figure 3. Longitudinal 6MWT and VL FF data
Longitudinal data of patients with Duchenne muscular dystrophy (DMD) from Leiden University Medical Center (LUMC) (dark circles) and Cincinnati Children's Hospital Medical Center (CCHMC) (lighter squares). (A) 6-Minute walking test (6MWT) results plotted vs vastus lateralis (VL) fat fraction (FF). LUMC and CCHMC patients with similar 6MWT results have similar VL FFs. FFs from non-ambulant patients are plotted as 0 m on the 6MWT. (B) 6MWT data plotted vs age. CCHMC patients on average walk longer distances at later ages than LUMC patients. Age at loss of ambulation (LoA) is plotted as 0 m. (C) Original VL FF results plotted vs age. On average VL FF results were higher and increased faster over time in LUMC patients compared to CCHMC patients. Data visually correspond to a sigmoid curve. (D) Original and predicted VL FF results plotted vs age. Patients with higher VL FFs at younger ages or faster FF increases had steeper predicted FF slopes. On average, LUMC patients showed steeper slopes than patients from CCHMC. Logistic curves from two LUMC patients at comparable ages are highlighted (pink lines) to illustrate the relationship with their LoA, depicted as an X, at 11.4 and 13.1 years of age. (E) Individual slope of the predicted FF curves plotted vs age at LoA. With the use of a Spearman correlation analysis on the LUMC cohort, there was a negative correlation between these variables (rho = -0.72, p = 0.001).

Visual relation between ambulation and VL FF

The relation between the absolute 6MWT distances and VL FF was similar in both cohorts (Figure 3A). However, CCHMC patients walked longer distances at a later age than LUMC patients although with considerable interindividual variation in the 6MWTs (Figure 3B). In parallel, VL FFs were higher and increased faster over time in LUMC patients compared to CCHMC patients, and VL FFs from patients from both centers corresponded visually to the sigmoid curves described previously (Figure 3C).[13, 14]

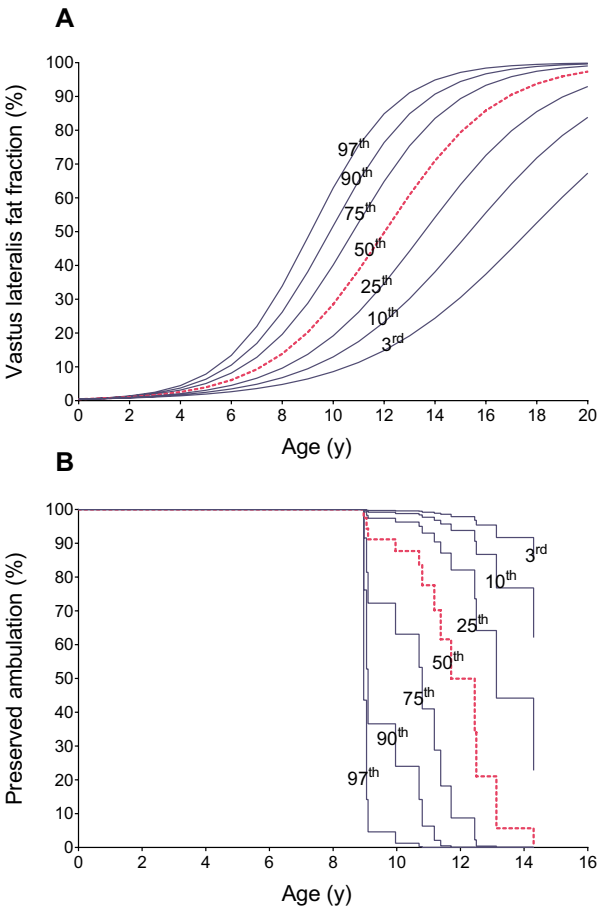


Figure 4. VL FF growth chart and survival chart of preserved ambulation for the LUMC cohort
Growth charts based on data of patients with Duchenne muscular dystrophy (DMD) from Leiden University Medical Center (LUMC) plotted vs age. (A) We generated a vastus lateralis (VL) fat fraction (FF) growth chart with a 3rd, 10th, 25th, 50th, 75th, 90th, and 97th percentile curve from the predicted LUMC VL FF data. (B) Using the resulting hazard ratio from the LUMC cohort, we transformed the predicted LUMC VL FF growth curves to survival curves for preserved ambulation. A patient on the third percentile in the VL FF growth chart is also on the third percentile in the survival chart.

Quantitative relation between ambulation and VL FF

Figure 3D shows original and predicted VL FF data. On average, LUMC patients showed steeper slopes for increasing FF than patients from CCHMC. The hazard ratio of a percent-point increase in VL FF for the time to LoA was 1.15 (log hazard ratio 0.14, 95% CI 1.05-1.26, Wald test p = 0.003) in the LUMC dataset, and 0.96 (log hazard ratio -0.04, 95% CI 0.84-1.10, Wald test p = 0.569) in the CCHMC dataset. The hazard ratio in the LUMC cohort of 1.15 corresponds to a 4.11-fold increase of the instantaneous risk of LoA in patients with a 10% higher VL FF at any age. Because of the limited number of events in the CCHMC cohort, we could assess the Spearman rank correlation and the survival chart only in the LUMC cohort. In LUMC patients, FF increased more rapidly (i.e., steeper slope) in those who lost ambulation at an earlier age (rho = -0.72, p = 0.001, Figure 3E). The LUMC VL FF growth chart (Figure 4A) and survival chart (Figure 4B) illustrate how the different VL FF curves relate to a range of LoA trajectories.

Discussion

This study aimed to relate quantitative muscle MRI to a clinically meaningful endpoint in DMD to substantiate the use of qMRI in clinical trials. The elevated hazard ratio in the LUMC cohort supports the use of qMRI as a surrogate outcome measure in clinical trials in DMD because it shows a direct relation between VL FF and losing ambulation.

Natural history studies and multiple placebo cohorts from clinical trials show that DMD progresses at different rates in different patients.[27-32] The most commonly used 6MWT and any functional test of individual ambulant patients with DMD are hallmarked by successive periods of increase, stabilization, and decline. This limits their applicability in clinical trials including both early- and late-ambulant boys, especially when studying drugs such as antisense oligonucleotides that aim to prevent deterioration through exon skipping rather than to improve muscle strength. In contrast to clinical parameters in this growing population, FF increases throughout life according to a sigmoidal curve.[13, 14] The use of qMRI-determined FFs as a surrogate outcome in clinical trials has so far been hampered by the absence of large natural history datasets[3] and a clear interrelation with functional and clinically meaningful outcomes. While cross-sectional correlations between muscle FF, strength, and function at the time of qMRI or magnetic resonance spectroscopy have been described extensively in DMD, a correlation does not prove causality.[4, 5, 9, 15-20] Even unrelated biological parameters that consistently increase or decrease with age will inherently correlate with functional parameters in a progressive disease. Therefore, such correlations alone are not sufficient to qualify qMRI FF as a surrogate outcome measure. In addition, many of these studies used outcomes of clinical assessments as a correlate rather than milestones such as LoA. This last variable is clearly clinically meaningful in the course of DMD, while, for example, absolute distances on the 6MWT are more difficult to interpret.[6,

[7] Finally, the survival analysis applied here enables the use of previously reached endpoints and longer follow-up without patients having to continue visiting the study site because it relates qMRI FF to a clinical endpoint rather than an assessment performed at the time of MRI.

The excellent interobserver reliability found is important in view of differences in fat replacement along the proximodistal axis of muscles in DMD.[23] In facioscapulohumeral muscular dystrophy, FF determination also yielded an excellent interobserver reliability (ICC 0.992).[33] Thus, ROIs can confidently be drawn at local sites in clinical trials across multiple institutions, while decisions on scan inclusion and which slices to use should be made with central reading. Furthermore, previous studies reported that FF determination is accurate and reproducible on 1.5T and 3.0T MRI scanners.[34] Together with the high intersite reliability and repeatability previously shown in DMD,[12] these results even further support the feasibility of the use of qMRI to determine FF in DMD in multicenter studies.

Increasingly, trial designs rely on categorizing patients according to their baseline walking distance[35] or the inclusion of patients in whom combined results of functional tests and age predict a certain trajectory (e.g., NCT02851797). Modeling the rate of disease progression by qMRI FF as applied in the current study could be used to determine a patient-specific sigmoidal FF curve that is consistent throughout the course of the disease and to facilitate the inclusion of a targeted population at baseline.[36] Similarly, the hazard ratio resulting from the survival analysis and VL FF growth chart could be used in trials to generate a survival chart. Every patient-specific sigmoidal FF curve then corresponds to a percentile curve on the VL FF growth chart and to a percentile curve on the survival chart for a clinical endpoint (Figure 4, A and B). Thus, the analysis can be used to stratify randomization to better account for prognostic factors in the small sample sizes that are available in this rare disease.

The two cohorts consisted of patients with clinically and genetically confirmed DMD, which was the main inclusion criterion for both natural history studies. Consistent with prior studies,[28] individual 6MWT trajectories at both sites showed a high interpatient variability. However, CCHMC patients were clearly at the better end of the disease spectrum. Possible explanations are the selection of only ambulant patients for the natural history study protocol at CCHMC, the earlier start of steroid treatment, the use of another steroid regimen, and the selection of skippable mutations. In previous studies, daily deflazacort and daily prednisone appear to have similar effects on strength and function tests in DMD,[37, 38] while in other studies, daily deflazacort and daily prednisone seem to have superior effects on ambulation compared to intermittent dosing of prednisone.[39, 40] Eligibility for skipping of exon 44 is also known to be associated with a milder disease course.[41, 42] Despite these differences in motor performance by age, the absolute distances walked on the 6MWTs showed a similar relationship with the VL FF in both cohorts, illustrating the tight association between VL FF and ambulatory ability.

Several limitations of the study need to be mentioned. First, the hazard ratio results could not be replicated in the CCHMC cohort. This could be due to the limited number of three LoA events that happened during a follow-up period for ambulant patients of 2.2 to 3.0 years. In contrast, seven of the 19 LUMC patients with VL FF data lost their ambulation before the start of the study, and another seven lost their ambulation during the follow-up period for ambulant patients of 4.5 to 5.0 years. Second, meticulous determination of LoA is important for the model, and the retrospective design of the current study could have influenced this. We suspect this influence to be minimal because patients and families considered LoA a life-changing event, and LoA was established during regular clinical follow-up for all three CCHMC patients and ten out of 14 LUMC patients with VL FF data. Finally, the small sample size and limited number of time points per patient did not allow us to model the intercept of the FF curves. Adding this parameter to the slopes would increase the flexibility of the model and could improve prediction of FF curves.[14]

We found VL FF to have added predictive value to age on LoA in the LUMC cohort that represented a more severe spectrum of the disease. By applying a well-defined anatomic landmark, we found an excellent interobserver reliability for VL FF determined by quantitative Dixon MRI, which supports the feasibility of multicenter muscle qMRI studies in DMD. In relatively small clinical trials, randomization can be stratified by the patient-specific modeling of a sigmoidal FF curve that corresponds to percentile curves on the VL FF growth chart and the survival chart. Although results should be confirmed in a larger cohort with prospective determination of the clinical endpoint, our results support the use of FF assessed with qMRI as a surrogate endpoint or stratification tool in clinical trials in DMD.

Acknowledgment

The authors thank Susan J. Ward, PhD, (founder and executive director of collaborative Trajectory Analysis Project), for her critical review of the manuscript.

Study funding

The LUMC natural history study was funded by the Netherlands Organization for Health Research and Development (ZonMw; grant 113302001). The PRO-DMD-01 natural history study (NCT01753804) was sponsored by Prosensa Therapeutics B.V. and BioMarin. The EU BIOIMAGE-NMD program with project ID 602485 was funded under FP7-HEALTH-2013-INNOVATION-1.

References

1. Cros D, Harnden P, Pellissier JF, Serratrice G. Muscle hypertrophy in Duchenne muscular dystrophy. A pathological and morphometric study. Journal of neurology 1989;236:43-47. doi:10.1007/BF00314217
2. Verhaart IEC, Aartsma-Rus A. Therapeutic developments for Duchenne muscular dystrophy. Nature reviews Neurology 2019;15:373-386. doi:10.1038/s41582-019-0203-3
3. Straub V, Balabanov P, Bushby K, et al. Stakeholder cooperation to overcome challenges in orphan medicine development: the example of Duchenne muscular dystrophy. Lancet Neurol 2016;15:882-890. doi:10.1016/S1474-4422(16)30035-7
4. Barnard AM, Willcocks RJ, Finanger EL, et al. Skeletal muscle magnetic resonance biomarkers correlate with function and sentinel events in Duchenne muscular dystrophy. PLoS One 2018;13:e0194283. doi:10.1371/journal.pone.0194283
5. Bonati U, Hafner P, Schadelin S, et al. Quantitative muscle MRI: A powerful surrogate outcome measure in Duchenne muscular dystrophy. Neuromuscul Disord 2015;25:679-685. doi:10.1016/j.nmd.2015.05.006
6. CDER, CBER. Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment. Guidance for Industry. https://www.fda.gov/media/92233/download. Accessed on May 1, 2019. US Food & Drug Administration (FDA) 2018
7. CHMP. Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. In European Medicines Agency (EMA). Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf. Accessed December 18, 2020.
8. Willcocks RJ, Rooney WD, Triplett WT, et al. Multicenter prospective longitudinal study of magnetic resonance biomarkers in a large duchenne muscular dystrophy cohort. Ann Neurol 2016;79:535-547. doi:10.1002/ana.24599
9. Wokke BH, van den Bergen JC, Versluis MJ, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. Neuromuscul Disord 2014;24:409-416. doi:10.1016/j.nmd.2014.01.015
10. Godi C, Ambrosi A, Nicastro F, et al. Longitudinal MRI quantification of muscle degeneration in Duchenne muscular dystrophy. Ann Clin Transl Neurol 2016;3:607-622. doi:10.1002/acn3.319
11. Azzabou N, Loureiro de Sousa P, Caldas E, Carlier PG. Validation of a generic approach to muscle water T2 determination at 3T in fat-infiltrated skeletal muscle. J Magn Reson Imaging 2015;41:645-653. doi:10.1002/jmri.24613
12. Forbes SC, Walter GA, Rooney WD, et al. Skeletal muscles of ambulant children with Duchenne muscular dystrophy: validation of multicenter study of evaluation with MR imaging and MR spectroscopy. Radiology 2013;269:198-207. doi:10.1148/radiol.13121948
13. Hooijmans MT, Doorenweerd N, Baligand C, et al. Spatially localized phosphorous metabolism of skeletal muscle in Duchenne muscular dystrophy patients: 24-month follow-up. PLoS One 2017;12:e0182086. doi:10.1371/journal.pone.0182086
14. Rooney WD, Berlow Y, Forbes SC, et al. Imaging in Neuromuscular Disease 2017: First International Conference on Imaging in Neuromuscular Disease, 19th - 21st November 2017, Berlin, Germany. Abstract 44: Statistical modeling of 5-year longitudinal data from a large cohort of Duchenne muscular dystrophy (DMD) subjects: an interim look at temporal characteristics of disease progression from the ImagingDMD study. J Neuromuscul Dis 2017;4:S1-s63
15. Fischmann A, Hafner P, Gloor M, et al. Quantitative MRI and loss of free ambulation in Duchenne muscular dystrophy. Journal of neurology 2013;260:969-974. doi:10.1007/s00415-012-6733-x
16. Wary C, Azzabou N, Giraudeau C, et al. Quantitative NMRI and NMRS identify augmented disease progression after loss of ambulation in forearms of boys with Duchenne muscular dystrophy. NMR Biomed 2015;28:1150-1162. doi:10.1002/nbm.3352
17. Akima H, Lott D, Senesac C, et al. Relationships of thigh muscle contractile and non-contractile tissue with function, strength, and age in boys with Duchenne muscular dystrophy. Neuromuscul Disord 2012;22:16-25. doi:10.1016/j.nmd.2011.06.750
18. Hogrel JY, Wary C, Moraux A, et al. Longitudinal functional and NMR assessment of upper limbs in Duchenne muscular dystrophy. Neurology 2016;86:1022-1030. doi:10.1212/WNL.0000000000002464
19. Willcocks RJ, Triplett WT, Forbes SC, et al. Magnetic resonance imaging of the proximal upper extremity musculature in boys with Duchenne muscular dystrophy. Journal of neurology 2017;264:64-71. doi:10.1007/s00415-016-8311-0
20. Gaeta M, Messina S, Mileto A, et al. Muscle fat-fraction and mapping in Duchenne muscular dystrophy: evaluation of disease distribution and correlation with clinical assessments. Preliminary experience. Skeletal Radiol 2012;41:955-961. doi:10.1007/s00256-011-1301-5

21. van den Bergen JC, Ginjaar HB, van Essen AJ, et al. Forty-Five Years of Duchenne Muscular Dystrophy in The Netherlands. J Neuromuscul Dis 2014;1:99-109
22. Hooijmans MT, Damon BM, Froeling M, et al. Evaluation of skeletal muscle DTI in patients with duchenne muscular dystrophy. NMR Biomed 2015;28:1589-1597. doi:10.1002/nbm.3427
23. Hooijmans MT, Niks EH, Burakiewicz J, et al. Non-uniform muscle fat replacement along the proximodistal axis in Duchenne muscular dystrophy. Neuromuscul Disord 2017;27:458-464. doi:10.1016/j.nmd.2017.02.009
24. Chrzanowski SM, Baligand C, Willcocks RJ, et al. Multi-slice MRI reveals heterogeneity in disease distribution along the length of muscle in Duchenne muscular dystrophy. Acta Myol 2017;36:151-162
25. Gold GE, Han E, Stainsby J, et al. Musculoskeletal MRI at 3.0 T: relaxation times and image contrast. AJR Am J Roentgenol 2004;183:343-351. doi:10.2214/ajr.183.2.1830343
26. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic 1981;86:127-137
27. Goemans N, Vanden Hauwe M, Signorovitch J, et al. Individualized Prediction of Changes in 6-Minute Walk Distance for Patients with Duchenne Muscular Dystrophy. PLoS One 2016;11:e0164684. doi:10.1371/journal.pone.0164684
28. Mercuri E, Signorovitch JE, Swallow E, et al. Categorizing natural history trajectories of ambulatory function measured by the 6-minute walk distance in patients with Duchenne muscular dystrophy. Neuromuscul Disord 2016;26:576-583. doi:10.1016/j.nmd.2016.05.016
29. McDonald CM, Henricson EK, Abresch RT, et al. The 6-minute walk test and other endpoints in Duchenne muscular dystrophy: longitudinal natural history observations over 48 weeks from a multicenter study. Muscle Nerve 2013;48:343-356. doi:10.1002/mus.23902
30. Pane M, Mazzone ES, Fanelli L, et al. Reliability of the Performance of Upper Limb assessment in Duchenne muscular dystrophy. Neuromuscul Disord 2014;24:201-206. doi:10.1016/j.nmd.2013.11.014
31. Bushby K, Finkel R, Wong B, et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. Muscle Nerve 2014;50:477-487. doi:10.1002/mus.24332
32. Goemans N, Mercuri E, Belousova E, et al. A randomized placebo-controlled phase 3 trial of an antisense oligonucleotide, drisapersen, in Duchenne muscular dystrophy. Neuromuscul Disord 2018;28:4-15. doi:10.1016/j.nmd.2017.10.004
33. Mul K, Vincenten SCC, Voermans NC, et al. Adding quantitative muscle MRI to the FSHD clinical trial toolbox. Neurology 2017;89:2057-2065. doi:10.1212/WNL.0000000000004647
34. Hernando D, Sharma SD, Aliyari Ghasabeh M, et al. Multisite, multivendor validation of the accuracy and reproducibility of proton-density fat-fraction quantification at 1.5T and 3T using a fat-water phantom. Magn Reson Med 2017;77:1516-1524. doi:10.1002/mrm.26228
35. McDonald CM, Campbell C, Torricelli RE, et al. Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet 2017;390:1489-1498. doi:10.1016/S0140-6736(17)31611-2
36. Pane M, Mazzone ES, Sivo S, et al. Long term natural history data in ambulant boys with Duchenne muscular dystrophy: 36-month changes. PLoS One 2014;9:e108205. doi:10.1371/journal.pone.0108205
37. Griggs RC, Miller JP, Greenberg CR, et al. Efficacy and safety of deflazacort vs prednisone and placebo for Duchenne muscular dystrophy. Neurology 2016;87:2123-2131. doi:10.1212/WNL.0000000000003217
38. Matthews E, Brassington R, Kuntzer T, et al. Corticosteroids for the treatment of Duchenne muscular dystrophy. Cochrane Database Syst Rev 2016:CD003725. doi:10.1002/14651858.CD003725.pub4
39. Ricotti V, Ridout DA, Scott E, et al. Long-term benefits and adverse effects of intermittent versus daily glucocorticoids in boys with Duchenne muscular dystrophy. J Neurol Neurosurg Psychiatry 2013;84:698-705. doi:10.1136/jnnp-2012-303902
40. McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. Lancet 2018;391:451-461. doi:10.1016/S0140-6736(17)32160-8
41. Pane M, Mazzone ES, Sormani MP, et al. 6 Minute walk test in Duchenne MD patients with different mutations: 12 month changes. PLoS One 2014;9:e83400. doi:10.1371/journal.pone.0083400
42. van den Bergen JC, Ginjaar HB, Niks EH, et al. Prolonged Ambulation in Duchenne Patients with a Mutation Amenable to Exon 44 Skipping. J Neuromuscul Dis 2014;1:91-94
43. Schonbeck Y, Talma H, van Dommelen P, et al. The world's tallest nation has stopped growing taller: the height of Dutch children from 1955 to 2009. Pediatr Res 2013;73:371-377. doi:10.1038/pr.2012.189

Chapter 4

Association of elbow flexor MRI fat fraction with loss of hand-to-mouth movement in patients with Duchenne muscular dystrophy

Published in: Neurology
Oct 2021; 97: e1737-e1742; DOI:10.1212/WNL.0000000000012724

Karin J. Naarding | Menno Van der Holst | Erik W. van Zwet | Nienke M. van de Velde
Imelda J.M. de Groot | Jan J.G.M. Verschuuren | Hermien E. Kan | Erik H. Niks

Abstract

Objective

To study the potential of quantitative MRI (qMRI) fat fraction (FF) as biomarker in non-ambulant Duchenne muscular dystrophy (DMD) patients, we assessed the additive predictive value of elbow flexor FF to age on loss of hand-to-mouth movement.

Methods

Non-ambulant DMD patients (≥8 years) were included. 4-point Dixon MRI scans of the right upper arm were performed at baseline and at 12, 18 or 24 months follow-up. Elbow flexor FFs were determined from five central slices. Loss of hand-to-mouth movement was determined at study visits and by phone-calls every four months. FFs were fitted to a sigmoidal curve using a mixed model with random slope to predict individual trajectories. The added predictive value of elbow flexor FF to age on loss of hand-to-mouth movement was calculated from a Cox model with the predicted FF as a time varying covariate, yielding a hazard ratio.

Results

Forty-eight MRIs of 20 DMD patients were included. The hazard ratio of a percent-point increase in elbow flexor FF for the time to loss of hand-to-mouth movement was 1.12 (95%-confidence interval 1.04-1.21; p=0.002). This corresponded to a 3.13-fold increase of the instantaneous risk of loss of hand-to-mouth movement in patients with a 10 percent-points higher elbow flexor FF at any age.

Conclusion

In this prospective study, elbow flexor FF predicted loss of hand-to-mouth movement independent of age. qMRI measured elbow flexor FF can be used as surrogate endpoint or stratification tool for clinical trials in non-ambulant DMD patients.

Classification of Evidence

This study provides Class II evidence that qMRI FF of elbow flexor muscles in patients with DMD predicts loss of hand-to-mouth movement independent of age.

Introduction

Duchenne muscular dystrophy (DMD) is characterized by muscle weakness in a proximal to distal gradient. Independent ambulation is generally lost in the early teens, and occurs years before loss of hand-to-mouth movement (LoHM).[1] While the first drugs for ambulant DMD have received conditional approval, results cannot be extrapolated to later stages of the disease due to progressive and irreversible replacement of muscle by fat and fibrosis causing a reduction in target tissue.[2, 3] Conducting clinical trials in DMD is challenging and may be facilitated by objective biomarkers that can be used for stratification or as surrogate endpoint. Quantitative MRI (qMRI) fat fraction (FF) of the vastus lateralis muscle has been shown to predict loss of ambulation in DMD.[4, 5] Importantly, this predictive value must be additive to age as any parameter that consistently changes over time will correlate to a decline in function in a progressive disease. Upper arm qMRI FF increases over time and correlates with function cross-sectionally.[6, 7] We studied the additive predictive value of elbow flexor FF (FF_{EF}) for LoHM to age in a prospective study in non-ambulant DMD patients.

Methods

We included male non-ambulant, genetically confirmed DMD patients aged ≥8 years between March 2018 and July 2019. Patients were recruited from the Dutch Dystrophinopathy Database[8], and via Dutch outpatient clinics and patient organizations. Exclusion criteria were MRI contra-indications (e.g. spinal fusion, daytime respiratory support or inability to lie still for 45 minutes), exposure to an investigational drug ≤6 months prior to participation, and recent (≤6 months) upper extremity surgery or trauma. 122 eligible patients were approached for participation, and details on this recruitment have been reported previously.[9] Patients visited the Leiden University Medical Center (LUMC) for a half-day of assessments at baseline, 12 and 18 months. Due to the COVID-19 pandemic, some follow-up visits were postponed from 12 to 18 months and from 18 to 24 months or missed. Telephone calls every 4 months were used to evaluate LoHM.

Standard Protocol Approvals, Registrations, and Patient Consents

The medical ethics committee of the LUMC approved the study, and we obtained written informed consent from patients and legal representatives. The study was registered on ToetsingOnline (NL63133.058.17, www.toetsingonline.nl).

MRI acquisition and analysis

4-point Dixon scans were acquired of the right upper arm on a 3T MRI scanner (Ingenia, Philips Healthcare, Best, The Netherlands) using two circular 15cm coils. Participants were positioned on the right side with the right shoulder and elbow in 90° flexion, because pilot experiments suggested this to be the most comfortable position and this position placed

the upper arm muscles more towards the center of the MRI scanner. If this was uncomfortable a supine position was chosen. Dixon scans were acquired with 33 slices and a voxel size of 1x1x10mm (repetition time 310ms, first echo 4.40ms, echo spacing 0.76ms, flip angle 20°), and aligned perpendicular to the humerus bone. Dixon water and fat images were generated using in-house developed software (Matlab 2016a, The Mathworks of Natick, Massachusetts, USA) with a 6-peak lipid spectrum, where B0 maps were calculated from the phase data of the first and last echoes. Regions of interest (ROI) were drawn on the muscle border of the elbow flexors (biceps and brachialis muscles) by one reviewer (K.J.N.) on 5 contiguous slices around a central slice (Figure 1A) using online software (mipav.cit.nih.gov). The central slice was located at 40% distance from the elbow based on the length of the humerus bone. The same reviewer performed quality control of elbow flexor ROIs, where scans with ROIs that contained artefacts or insufficient signal were excluded. ROIs from different time-points on similar slices in the same participant were also compared by this reviewer and adjusted in case of discrepancies. ROIs were eroded by two voxels, and FF_{EF} was calculated as a weighted mean value by averaging elbow flexor voxels of all eroded ROIs from the reconstructed fat and water images and correcting for partial saturation due to T1 effects by:

$$\mathrm{FF} = \frac{1.05 \times \mathrm{Fat}}{1.25 \times \mathrm{Water} + 1.05 \times \mathrm{Fat}}$$

Clinical assessments and endpoint

Performance of the Upper Limb (PUL) 2.0 was assessed for the right arm at all visits. LoHM was defined as the inability to move a filled glass independently to the mouth using the right hand and allowing support of the elbow on a table, similar to the PUL hand-to-mouth item where a 200gram weight is used. Age at LoHM was prospectively established to a month's precision. If LoHM had occurred before baseline, month and year were established retrospectively using a detailed interview and clinical documentation.

Statistical analysis

The difference in FF_{EF} between baseline and 12 months follow-up was assessed using Wilcoxon signed-rank test. The additive predictive value of FF_{EF} to age on LoHM was calculated using a Cox proportional hazards model as described previously.[4] For this, we applied a logit transformation to FF_{EF}, to allow use of standard statistical methods that rely on a gaussian distribution. A linear (mixed) model was fitted to the transformed data with age as the only covariate and a random slope per individual. The fitted lines were transformed back to the original scale using a logistic transformation, after which individual FF_{EF} trajectories were predicted at any time. The Cox proportional hazards model was fitted with the predicted FF_{EF} as a time-varying covariate, yielding a hazard ratio (Wald test; $p<0.05$). The primary research question of this study was: does FF_{EF} have additive predictive value to age on LoHM in non-ambulant DMD patients? The level of evidence was assigned as Class II during the review process.

Data availability

Anonymized data and analysis software can be made available to qualified investigators.

Results

Twenty-two DMD patients participated, but two patients refused the MRI. One patient switched to a medication trial after baseline, one patient discontinued after 12 months follow-up because of traveling distance, and eight visits were canceled due to COVID-19 restrictions. One 12 months follow-up scan was excluded due to insufficient signal. Forty-eight MRIs of 20 DMD patients were included, where 12 patients had three MRIs, four patients two and four patients one. All patients used glucocorticoids, but one patient had temporarily ceased treatment six weeks prior to baseline due to weight gain. Patient's characteristics at baseline, and FF_{EF} results and PUL scores at different time-points are presented in Table 1. LoHM had occurred before baseline in two patients, and occurred during the study in nine. Median decline in PUL total score over 12 months was 3 points (n=15; range -1 to 8). Median decline in PUL elbow domain score was 2 points (range 0 to 4; Figure 1B). There was a significant mean annual increase in FF_{FE} of 5.9% ± 5.4% ($p<0.01$).

Relation between hand-to-mouth movement and FF_{EF}

Acquired and predicted FF_{EF} data and predicted FF_{EF} at age of LoHM are shown in Figure 1C and 1D respectively. The hazard ratio of a percent-point increase in FF_{EF} for the time to LoHM was 1.12 (log hazard ratio 0.11; 95%-confidence interval 1.04-1.21; p=0.002). This hazard ratio corresponds to a 3.13-fold increase of the instantaneous risk of LoHM in patients with a 10 percent-points higher FF_{EF} at any age. An FF_{EF} growth chart (Figure 2A) and survival chart for preserved hand-to-mouth movement (Figure 2B) illustrate relationships between percentile FF_{EF} curves and LoHM trajectories.

Table 1. Patient characteristics

Characteristics at baseline	DMD patients n=20
Age, years	13.5 (12.5-16.4)
Righthanded	16 (80%)
Steroid use	
Prednisone intermittent	10 (50%)
Deflazacort intermittent	9 (45%)
Deflazacort daily	1 (5%)
Height, m	1.55 (1.46-1.66)
Weight, kg	65.1 (51.4-82.2)
Body mass index, kg/m^2	27.6 (23.3-31.1)

Table 1. Continued

Characteristics at baseline	DMD patients n=20
Age at start steroid use, years	5.6 (4.4-7.9)
Age at loss of ambulation, years	11.5 (10.0-12.9)
Loss of hand-to-mouth movement before inclusion	2 (10%)
PUL 2.0 total score baseline, points	21 (19-34)
Elbow flexor fat fraction baseline, %	50.9 (42.4-72.4)
Characteristics at follow-up	
PUL 2.0 total score 12 months, points	20 (17-30) [n=15]
PUL 2.0 total score 18 months, points	24 (16-29) [n=11]
PUL 2.0 total score 24 months, points	17; 20; 37 [n=3]
Elbow flexor fat fraction 12 months, %	60.2 (44.4-78.1) [n=14]
Elbow flexor fat fraction 18 months, %	67.9 (53.8-84.4) [n=11]
Elbow flexor fat fraction 24 months, %	23.3; 38.6; 86.1 [n=3]
Loss of hand-to-mouth movement during follow-up	9 (45%)
Age at loss of hand-to-mouth movement, years	15.3 (10.4-18.2) [n=11]
No loss of hand-to-mouth movement during follow-up	9 (45%)
Age last follow-up for patients with preserved hand-to-mouth movement, years	15.1 (14.4-16.3) [n=9]

Abbreviations: DMD = Duchenne muscular dystrophy. PUL = Performance of the Upper Limb.
Values are median (first; third quartiles), number of patients (%), or the actual values of all patients. If a certain value was not available for all patients, the number of patients for whom the data was available was presented after the result with *n* = number.

Discussion

In this prospective study, we show that FF_{EF} predicts LoHM in non-ambulant DMD patients on top of age. This added predictive value is essential, because parameters that consistently change over time will always correlate with functional tests in a progressive disease. Previous studies demonstrated that qMRI muscle FF increases over time and correlates with function cross-sectionally in DMD.[6, 7, 10, 11] However, any outcome measure that consistently changes over time will correlate to declining measures of function in a progressive disease. Two previous studies demonstrated the added predictive value of vastus lateralis FF on top of age on the clinical outcome loss of ambulation, and thus showed for the first time that muscle FF adds to the assessment of disease severity.[4, 5]

FF_{EF} increased according to a sigmoidal curve, similar to the vastus lateralis FF in ambulant patients.[4, 5] The hazard ratio of 1.12 was comparable to that of the vastus lateralis FF for the time to loss of ambulation.[4] These data thus support the use of qMRI FF as objective

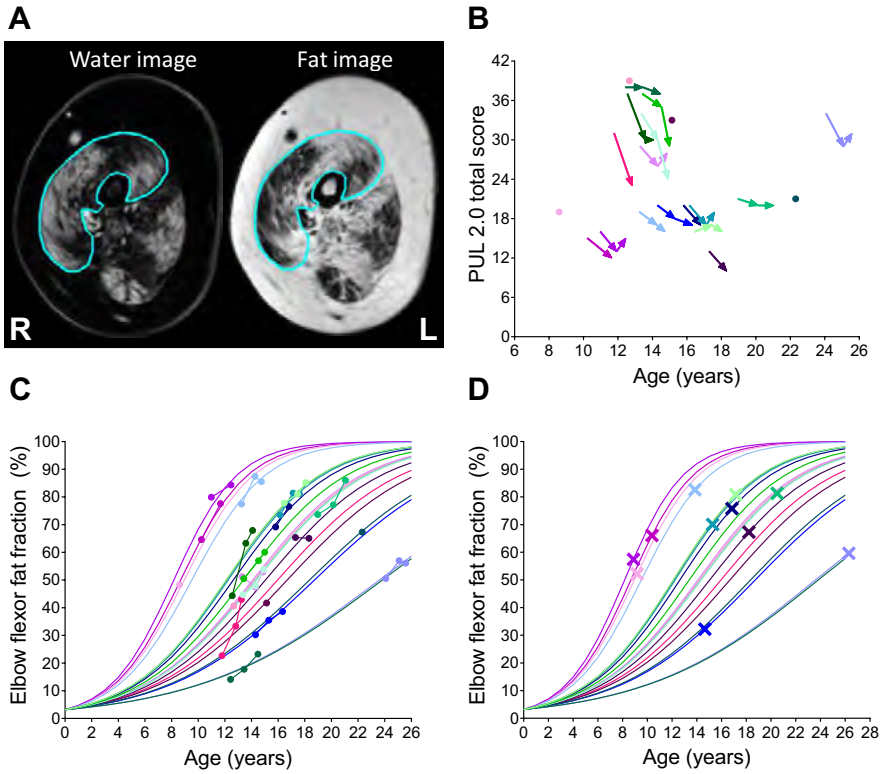


Figure 1. Longitudinal clinical and elbow flexor fat fraction (FF_{EF}) data
In (A) an example of a region of interest (ROI) drawn on the elbow flexor muscles (line) is shown on a water image (left) and corresponding fat image (right). In (B) longitudinal Performance of the Upper Limb (PUL) 2.0 total scores (maximum 42 points) are plotted versus age. PUL total scores decrease with age, but there is a large variation of scores between Duchenne muscular dystrophy (DMD) patients of similar ages. In (C) FF_{EF} results that were acquired are plotted versus age, as well as FF_{EF} results that were predicted using a logit transformation, linear (mixed) model and logistic transformation. Patients with higher FF_{EF} results at younger ages or faster FF_{EF} increases had steeper predicted FF_{EF} slopes. In (D) predicted FF_{EF} results are plotted versus age and predicted FF_{EF} at age loss of hand-to-mouth movement are shown with a cross. The colors used in (B), (C) and (D) are unique for each participant.

biomarker in different stages of the disease. Predicted FF curves can be used for stratification in clinical trials or as surrogate endpoint, limiting sample size and duration.

The rate of change in FF, for instance one-year change, could be used as biomarker in trials where the therapeutic effect over that period of time can be compared to placebo or another therapy, and the power calculation could be based on the more or less 'linear' middle part of the FF curve as that is where the fastest change is expected to happen. This will require stratification of the cohort with respect to baseline FF, as is now commonly done for functional tests.[10]

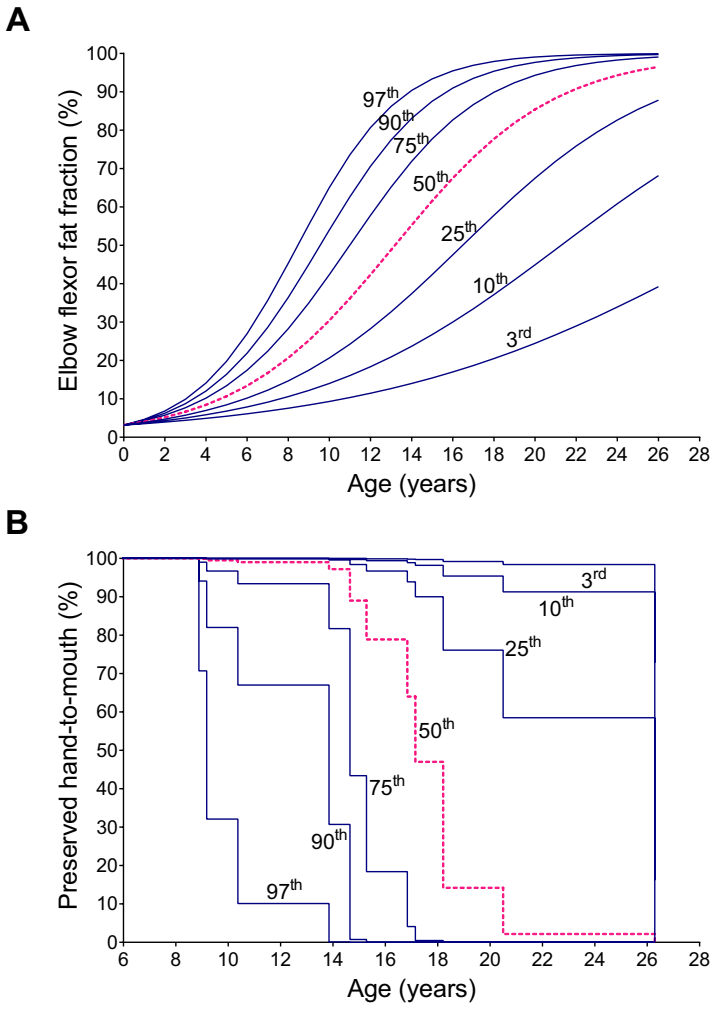


Figure 2. Elbow flexor fat fraction (FF_{EF}) growth chart and survival chart of preserved hand-to-mouth movement

In (A) we generated an FF_{EF} growth chart with a 3rd, 10th, 25th, 50th, 75th, 90th and 97th percentile curve from the predicted FF_{EF} data. (B) Using the resulting hazard ratio, we transformed the predicted FF_{EF} growth curves to a 3rd, 10th, 25th, 50th, 75th, 90th and 97th percentile survival curve for preserved hand-to-mouth movement. A patient on the 3rd percentile in the FF_{EF} growth chart is also on the 3rd percentile in the survival chart.

We assessed the timing of reaching the clinical endpoint via regular phone calls in between clinical assessments. In our experience, patients and caregivers are able to define such important endpoints within a month's precision. This increases the power of our survival analyses compared to standard natural history studies where clinical assessments are performed at six or 12 months intervals only. It reduced the burden for participants and

allowed continuation of the protocol despite COVID-19 related restrictions. The importance of hand-to-mouth movement is stressed by its incorporation in the widely used Brooke upper extremity rating scale, the PUL and DMD Upper Limb Patient Reported Outcome Measure, where patients and families confirmed its clinical relevance.[12, 13]
Limitations of this study are the small sample size, which did not allow modelling the intercept of the FF_{EF} curves. Restrictions due to the COVID-19 pandemic also led to some missing data. It's important to replicate results in other cohorts with different steroid regimes.

In conclusion, FF_{EF} predicted loss of hand-to-mouth movement independent of age in non-ambulant DMD patients. This establishes qMRI FF as biomarker in DMD and potentially facilitates the design of clinical trials, either via stratification or use as surrogate endpoint.

Study Funding
Funding provided by Stichting Spieren for Spieren (grant SvS15).

References

1. McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. Lancet 2018;391:451-461. doi:10.1016/S0140-6736(17)32160-8
2. CHMP. Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. In European Medicines Agency (EMA). Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf. Accessed December 18, 2020.
3. Verhaart IEC, Aartsma-Rus A. Therapeutic developments for Duchenne muscular dystrophy. Nature reviews Neurology 2019;15:373-386. doi:10.1038/s41582-019-0203-3
4. Naarding KJ, Reyngoudt H, van Zwet EW, et al. MRI vastus lateralis fat fraction predicts loss of ambulation in Duchenne muscular dystrophy. Neurology 2020;94:e1386-e1394. doi:10.1212/WNL.0000000000008939
5. Rooney WD, Berlow YA, Triplett WT, et al. Modeling disease trajectory in Duchenne muscular dystrophy. Neurology 2020;94:e1622-e1633. doi:10.1212/WNL.0000000000009244
6. Forbes SC, Arora H, Willcocks RJ, et al. Upper and Lower Extremities in Duchenne Muscular Dystrophy Evaluated with Quantitative MRI and Proton MR Spectroscopy in a Multicenter Cohort. Radiology 2020;295:616-625. doi:10.1148/radiol.2020192210
7. Wood CL, Hollingsworth KG, Hughes E, et al. Pubertal induction in adolescents with DMD is associated with high satisfaction, gonadotropin release and increased muscle contractile surface area. Eur J Endocrinol 2020. doi:10.1530/EJE-20-0709
8. van den Bergen JC, Ginjaar HB, van Essen AJ, et al. Forty-Five Years of Duchenne Muscular Dystrophy in The Netherlands. J Neuromuscul Dis 2014;1:99-109
9. Naarding KJ, Doorenweerd N, Koeks Z, et al. Decision-Making And Selection Bias in Four Observational Studies on Duchenne and Becker Muscular Dystrophy. J Neuromuscul Dis 2020;7:433-442. doi:10.3233/JND-200541
10. Barnard AM, Willcocks RJ, Triplett WT, et al. MR biomarkers predict clinical function in Duchenne muscular dystrophy. Neurology 2020;94:e897-e909. doi:10.1212/WNL.0000000000009012
11. Barnard AM, Willcocks RJ, Finanger EL, et al. Skeletal muscle magnetic resonance biomarkers correlate with function and sentinel events in Duchenne muscular dystrophy. PLoS One 2018;13:e0194283. doi:10.1371/journal.pone.0194283
12. Mayhew A, Mazzone ES, Eagle M, et al. Development of the Performance of the Upper Limb module for Duchenne muscular dystrophy. Dev Med Child Neurol 2013;55:1038-1045. doi:10.1111/dmcn.12213
13. Klingels K, Mayhew AG, Mazzone ES, et al. Development of a patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy: DMD Upper Limb PROM. Dev Med Child Neurol 2017;59:224-231. doi:10.1111/dmcn.13277

Chapter 5

Preserved thenar muscles in non-ambulant Duchenne muscular dystrophy patients

Published in: J Cachexia Sarcopenia Muscle
Jun 2021; 12(3); 694-703; DOI: 10.1002/jcsm.12711

Karin J. Naarding | Kevin R. Keene | Aashley S.D. Sardjoe Mishre | Thom T.J. Veeger
Nienke M. van de Velde | Arina J. Prins | Jedrzej Burakiewicz | Jan J.G.M. Verschuuren
Menno van der Holst | Erik H. Niks | Hermien E. Kan

Abstract

background

Clinical trials in Duchenne muscular dystrophy (DMD) focus primarily on ambulant patients. Results cannot be extrapolated to later disease stages due to a decline in targeted muscle tissue. In non-ambulant DMD patients, hand function is relatively preserved and crucial for daily-life activities. We used quantitative MRI (qMRI) to establish whether the thenar muscles could be valuable to monitor treatment effects in non-ambulant DMD patients.

Methods

Seventeen non-ambulant DMD patients (range 10.2-24.1years) and 13 healthy controls (range 9.5-25.4years) underwent qMRI of the right hand at 3T at baseline. Thenar fat fraction (FF), total volume (TV), and contractile volume (CV) were determined using 4-point Dixon and $T2_{water}$ was determined using multi-echo spin-echo. Clinical assessments at baseline (n=17) and 12 months (n=13) included pinch strength (kg), Performance of the Upper Limb (PUL) 2.0, DMD Upper Limb Patient Reported Outcome Measure (PROM), and playing a video game for 10 minutes using a game controller. Groups differences and correlations were assessed with non-parametric tests.

Results

TV was lower in patients compared to HCs (6.9cm^3, 5.3-9.0cm^3 versus 13.0cm^3, 7.6-15.8cm^3, p=0.010). CV was also lower in patients (6.3cm^3, 4.6-8.3cm^3 versus 11.9cm^3, 6.9-14.6cm^3, p=0.010). FF was slightly elevated (9.7%, 7.3-11.4% versus 7.7%, 6.6-8.4%, p=0.043), while $T2_{water}$ was higher (31.5ms, 30.0-32.6ms versus 28.1ms, 27.8-29.4ms, p<0.001). Pinch strength and PUL decreased over 12 months (2.857kg, 2.137-4.010kg to 2.243kg, 1.930-3.339kg, and 29 points, 20-36 to 23 points, 17-30, both p<0.001), while PROM did not (49 points, 36-57 to 44 points, 30-54, p=0.041). All patients were able to play for 10 minutes at baseline or follow-up, but some did not comply with the study procedures regarding this endpoint. Pinch strength correlated with TV and CV in patients (rho=0.72, rho=0.68) and controls (both rho=0.89). PUL correlated with TV, CV and $T2_{water}$ (rho=0.57, rho=0.51, rho=-0.59).

Conclusions

Low thenar FF, increased $T2_{water}$, correlation of muscle size with strength and function, and the decrease in strength and function over one year indicate that the thenar muscles are a valuable and quantifiable target for therapy in later stages of DMD. Further studies are needed to relate these data to the loss of a clinically meaningful milestone.

Introduction

duchenne muscular dystrophy (DMD) is characterized by progressive replacement of muscle with fat and fibrotic tissue due to the absence of full length dystrophin.[1] This leads to a clinical presentation with a proximal to distal gradient of muscle weakness, in which loss of independent ambulation generally occurs years before loss of upper arm function.[2] Although the first drugs in DMD have now received regulatory approval, there is no cure yet, and ongoing clinical trials focus primarily on ambulant patients.[3] Regulators do not support extrapolation of trial results in ambulant patients to later disease stages, because progressive replacement of muscle by fat is considered irreversible and thus leads to a progressive reduction in target tissue.[4, 5]

Hand function is preserved even longer than upper arm function in DMD, and is crucial for daily-life activities.[6, 7] Many of these activities require grasping motions for which the thenar muscles primarily generate the opposition of the thumb.[8] Since efficacy of new drugs has to be proven separately in non-ambulant patients, preparation for clinical trials in later disease stages is essential. To facilitate the development of these trials, it is necessary to collect natural history data, develop outcome parameters, and study biomarkers that reflect a decline in upper limb function. Quantitative MRI (qMRI) fat fraction (FF) has shown potential as such a biomarker in the lower extremities, and increased FF of the vastus lateralis predicts a decline in ambulatory function.[9-11] Ongoing studies assess the feasibility of qMRI FF in the upper and lower arm, but studies in intrinsic hand muscles are lacking.[12-14] Another qMRI parameter, the T2 relaxation time of the muscle compartment ($T2_{water}$), is indicative for early muscle pathology, since it increases due to inflammation, myocyte swelling, edema and necrosis.[15, 16] Here, we aimed to study the thenar muscles for their value to monitor treatment effects in non-ambulant DMD patients by obtaining natural history of hand function over 1 year, and assessing qMRI parameters compared to healthy controls (HC).

Methods

Participants

Between March 2018 and July 2019, DMD patients were recruited from the Dutch Dystrophinopathy Database (DDD),[17] and via Dutch neurologists, rehabilitation specialists and patient organizations. Inclusion criteria were male, non-ambulant genetically confirmed DMD, aged ≥8 years. Exclusion criteria were MRI contra-indications (e.g. scoliosis surgery or daytime respiratory support), inability to lie still for 45 minutes, exposure to an investigational drug ≤6 months prior to participation, and recent (≤6 months) upper extremity surgery or trauma. Healthy age-matched controls were recruited using posters and advertisements in local media. The study was approved by the local medical ethics committee, and written informed consent was obtained from all patients and legal

representatives prior to inclusion. The investigation was conducted according to the 1964 Declaration of Helsinki and its later amendments. The reported assessments were performed as part of a longitudinal study on outcome measures in non-ambulant DMD patients at Leiden University Medical Center (LUMC): 'Upper extremity outcome measures in non-ambulant DMD patients', registered on ToetsingOnline with ABR number NL63133.058.17 (www.toetsingonline.nl). We report results from baseline (DMD and HC) and 12 months follow-up (DMD only). MRI of the hand muscles was only performed at baseline and clinical assessments were performed at every visit.

MRI acquisition

MRI scans of the right hand were acquired on a 3T scanner (Ingenia, Philips Healthcare, Best, The Netherlands) with two 47 mm surface coils placed on the ventral and dorsal side of the thenar muscles. The standard position for participants was on the right side with the right shoulder in 90° flexion, the elbow in 90° flexion and the forearm in maximum pronation, which was supported by holding a handle (Figure S1). If this position was uncomfortable, patients could choose from lying on the right side with the elbow halfway between pronation and supination or lying in supine position with the forearm either in maximum pronation or halfway between pronation and supination.

The protocol consisted of a 4-point Dixon scan (25 contiguous slices of 4mm thickness, repetition time (TR) 310ms, first echo time (TE) 4.40ms, echo spacing (ΔTE) 0.76ms, flip angle 20°, field of view (FOV) 100x100mm, voxel size 1x1mm, reconstructed voxel size 0.89x0.89mm), a multi-echo spin-echo (MSE) scan (5 slices of 4mm thickness with 8mm slice gap, TR 300ms, TE 8.0ms, ΔTE 8.0ms, FOV 100x100mm, voxel size 1x1mm, reconstructed voxel size 0.89x0.89mm), and a B1 map (5 contiguous slices of 12mm thickness, dual-TR gradient echo with TRs of 30ms and 100ms, TE 2.01ms, FOV 100x100mm, voxel size 1.56x2mm, reconstructed voxel size 0.89x0.89mm). The Dixon scans, MSE scans, and B1 maps were aligned perpendicular to the first metacarpal bone and covered the thenar muscles completely.

MRI analysis

Water and fat images were reconstructed from the Dixon data using in-house developed software (Matlab 2016a, The Mathworks of Natick, Massachusetts, USA) assuming a six-peak lipid spectrum.[18] B0 maps were determined from the phase data of the first and last echoes, and the Goldstein branch cut method was used for phase unwrapping.[19] Scans with major movement artefacts, water/fat swaps, insufficient signal or other artefacts in the thenar were excluded. Regions of interest (ROIs) of the thenar muscles were drawn on all slices of the Dixon scans using online available software (www.mipav.cit.nih.gov) (Figure 1A). The thenar muscles consist of the abductor pollicis brevis, flexor pollicis brevis and opponens pollicis muscle. The ROIs were drawn on the border of this muscle group, except at the dorsal side because the fascia between the different intrinsic hand muscles could not be

observed. This dorsal ROI boundary was defined by a straight line on the palmar side of the tendon of the flexor pollicis longus muscle and the first metacarpal bone. All ROIs of the HCs were drawn by one rater (K.J.N.) and of the DMD patients by two raters (K.J.N. and A.J.P.). Since different positioning of the first metacarpal bone could influence the position and size of a single slice, all qMRI values are reported for the whole thenar muscle and used for subsequent analyses. The total volume (TV; cm^3) of all thenar ROIs combined was assessed from these raw ROIs. For fat fraction measurements (FF; %), an erosion of two voxels was performed for every ROI to avoid contamination of ROIs with subcutaneous fat. A correction for the T1-weighting in the water and fat images was incorporated in the FF calculation, assuming T1 values of 1420ms (correction factor: 1.25) and 371ms (correction factor: 1.05) for water and fat respectively.[20] Thenar FF was then calculated as a weighted mean value by averaging all thenar voxels of all eroded ROIs from the reconstructed fat and water images as:

$$\mathrm{FF} = \frac{1.05 \times \mathrm{Fat}}{1.25 \times \mathrm{Water} + 1.05 \times \mathrm{Fat}}$$

Contractile volume (CV) was calculated by subtracting the fat containing area from TV as follows:

$$\mathrm{CV} = \mathrm{TV} - (\mathrm{TV} \times \mathrm{FF})$$

The MSE scans were fitted using a dictionary fitting algorithm with an Extended Phase Graph model in which the FF, $T2_{water}$ and B1 were fitted, and the flip angle slice profile and through-plane chemical shift displacement were incorporated.[21, 22] This dictionary was created using $T2_{water}$ values from 10-60ms, $T2_{fat}$ values from 120-200ms, FF values from 0-100% and B1 values from 50-140%. The $T2_{fat}$ was calibrated on the subcutaneous fat per slice by creating an automatic fat mask using the last echo of the MSE sequence. The fitted B1 values were visually confirmed to correspond to the B1 map by comparing values from similar locations on the same slices. Uneroded ROIs were copied from the Dixon images, manually checked and adapted in case of movement between scans, and eroded with two voxels before further analysis. Voxels that fitted on the physiological boundaries of the dictionary were excluded. If less than 100 thenar voxels remained for a particular scan, that scan was excluded. Thenar $T2_{water}$ was estimated over all slices as a weighted mean of all thenar voxels (Figure 1B), and for all intrinsic hand muscle voxels (supporting information).

Clinical assessments

All clinical assessments were performed after the MRI examination. Weight was measured using a patient lift with inbuilt scale (Maxi Move™, ArjoHuntleigh, Sweden). Height was calculated from the ulna length for both DMD patients and HCs using the formula proposed by Gauld et al. with 18 as maximum age.[23]

Pinch strength of the right hand was assessed with MyoPinch and grip strength with MyoGrip (Institute of Myology, Ateliers Laumonier, France).[24] Specific strength was defined as the

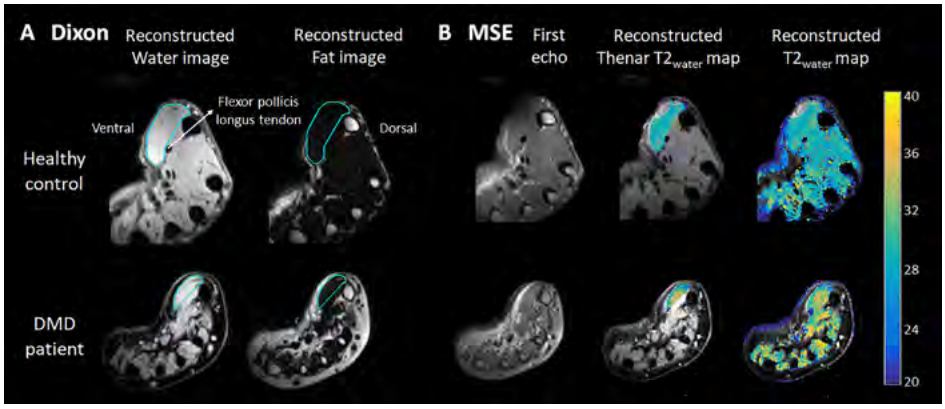


Figure 1. Dixon and multi-echo spin-echo (MSE) MRI acquisitions and analyses
(A) Example of a Dixon reconstructed water and fat image of a healthy control (HC; 16 years old) and a Duchenne muscular dystrophy (DMD) patient (13 years old). ROIs of the thenar muscles were drawn (light blue line) with the dorsal boundary defined as a line on the palmar side of the tendon of the flexor pollicis longus muscle and first metacarpal bone. (B) Example of the first echo from the MSE scan of the same HC and DMD patient and the following reconstructed images: $T2_{water}$ map for voxels within the thenar ROI and $T2_{water}$ map of all hand muscles. MSE voxels that fitted on the physiological boundaries of the dictionary, were excluded.

pinch strength in kg per cm^3 of CV. Performance of the Upper Limb (PUL) 2.0 was performed for the right arm, and only in DMD patients because a maximum score was assumed in HCs. The PUL 2.0 consists of 22 items and yields a maximum total score of 42 points.[25] Items are divided over three dimensions with a maximum score of 12 for the shoulder dimension, 17 for the elbow dimension and 13 for the distal wrist/hand dimension. The adapted DMD Upper Limb Patient Reported Outcome Measure (PROM) questionnaire was used to investigate patient reported upper limb function. The questionnaire contains 32 daily-life activity items which are scored on a three-level scale ('cannot do'; 'with difficulty'; 'easy') with a maximum possible score of 64 points.[26] Furthermore, DMD patients were asked to play a video game (Rocket League®, Psyonix LLC, California, USA) for 10 minutes using a game controller (GC-100XF Wired Gaming Controller, NACON™, Bigben group, France). In this game, patients used their right thumb to steer a vehicle on the screen. Afterwards patients rated tiredness and difficulty on two 10-point numeric rating scales (NRS) ranging from 'not at all tiring/difficult' (score 0) to 'very tiring/difficult' (score 10) with matching facial cartoons based on the Wong-Baker Faces Rating Scale.[27]

Statistical analysis

Interrater variability of qMRI results in DMD patients was assessed using Bland-Altman analyses to determine bias and limits of agreement. Differences in baseline characteristics, qMRI values, specific strength, and pinch strength between DMD patients and HCs were assessed with the Mann-Whitney U test. Differences in clinical assessments over 12 months were assessed using Wilcoxon signed-rank test. Statistical significance was set at $p<0.05$.

Bonferroni-Holm correction was used to correct for multiple comparisons of qMRI results between patients and HCs, and of clinical assessments between baseline and follow-up. Spearman's rank correlation coefficient was used to correlate qMRI values to pinch strength and grip strength in DMD patients and HCs, and to PUL total and distal, PROM total and NRS scores in patients. The correlation was considered very strong (0.9-1.0), strong (0.7-0.9) or moderate (0.5-0.7).[28]

Data availability

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

Results

Characteristics of participants and qMRI data inclusion

Twenty-two DMD patients and 14 HCs participated. Five patients and one HC could not undergo MRI. Characteristics of the remaining 17 patients and 13 HCs are presented in Table 1. Due to unforeseen restrictions during the COVID-19 pandemic, the 12 months follow-up visit could only take place for 13 patients. All patients used corticosteroids at baseline, except for two patients who had temporarily ceased treatment for six weeks prior to the visit due to weight gain and non-compliance.

All HCs and 10 DMD patients could maintain the standard position, while three patients had their forearm positioned halfway between pronation and supination. The MRI scans were acquired in supine position in four DMD patients due to discomfort lying on the side, one with the forearm in maximum pronation and three with their forearm positioned halfway between pronation and supination. Scans in a separate group of HCs showed no or minimal differences in qMRI results between two forearm positions (supporting information and Figure S2). Interrater variability (Figure S3) showed a mean bias in FF of 0.4%, and 95% limits of agreement -1.3 to 2.0%, for TV this was 0.4 cm^3 (-0.8 to 1.7 cm^3), for CV 0.3 cm^3 (-0.8 to 1.5 cm^3), and for $T2_{water}$ 0.1 ms (-0.8 to 0.9 ms).

Four Dixon scans and five MSE scans of DMD patients and four MSE scans of HCs were excluded due to insufficient signal or movement artefacts. Thirteen Dixon scans from patients and HCs, and 12 MSE scans from patients and nine from HCs were included in the analyses (see flowchart in Figure S4).

Thenar FF, muscle size and $T2_{water}$

Thenar FF was only slightly elevated in patients compared to HCs (median 9.7 versus 7.7%, p=0.043; Figure 2A). By contrast, TV and CV were significantly lower in patients (median 6.9 and 6.3 versus 13.0 and 11.9 cm^3, both p=0.010; Figure 2B and C), whereas thenar $T2_{water}$ was higher (median 31.5 versus 28.1 ms, p<0.001; Figure 2D). The higher thenar $T2_{water}$ in DMD patients could be visually confirmed for all intrinsic hand muscles (Figure S5). In

patients, whole thenar TV, CV and specific strength correlated very strongly (rho=0.96-0.97) to these same values determined on only the slice with the largest cross-sectional area (CSA), FF correlated moderately (rho=0.67; Figure S6).

Table 1. Baseline characteristics

	Healthy controls (n=13)	DMD patients (n=17)	p-value
Age, years	15.7 (11.1-20.7)	13.4 (12.5-16.7)	0.536
Calculated height, m	1.73 (1.47-1.75)	1.54 (1.47-1.66)	0.053
BMI, kg/m^2	19.5 (16.4-22.3)	27.5 (23.4-31.6)	**0.001**
Righthanded	12 (92.3%)	13 (76.5%)	
Prednisone intermittent	NA	8 (47.1%)	
Deflazacort intermittent	NA	8 (47.1%)	
Deflazacort daily	NA	1 (5.9%)	
Steroids on-day at baseline	NA	7 (41.1%)	
Steroids off-day at baseline	NA	10 (58.8%)	
Age at start steroid use, years	NA	5.7 (4.6-8.0)	
Age at loss of ambulation, years	NA	11.6 (10.1-12.8)	
		range 8.6-18.9	
Time since loss of ambulation, years	NA	2.6 (1.4-4.0)	

Values are median (first-third quartiles) or number of patients (%). Abbreviations: DMD = Duchenne muscular dystrophy, BMI = body mass index. Statistical significance was set at $p<0.05$.

Clinical assessments

Clinical assessments at baseline and follow-up are presented in Table 2. At baseline, pinch strength in patients was lower compared to HCs (Figure 2E), and it decreased significantly over 12 months from median 2.857 to 2.243 kg (Figure 3A). PUL total also significantly decreased from a median of 29 to 23 points. Grip strength declined from median 8.47 to 6.39 kg and PROM total scores from 49 to 44 points, but both were not significant after correction. PUL distal and NRS tiredness and difficulty scores did not show a decline over 12 months. At baseline, all patients scored NRS tiredness score ≤4 (i.e. 'not really tiring'). Similarly, NRS difficulty score was ≤4 (i.e. 'not really difficult') for all patients except two, who scored 5 and 6 (i.e. 'a little difficult'). Fifteen out of 17 patients were able to play the video game for 10 minutes at baseline, including the patient with the lowest pinch strength (0.723 kg), and 10 out of 13 patients were able at 12 months follow-up. Two patients at baseline and one at follow-up did not want to continue playing after five minutes, but were able to play for 10 minutes at a later study visit. The other two patients at follow-up refused or quit after five minutes, but claimed to easily play for 10 minutes.

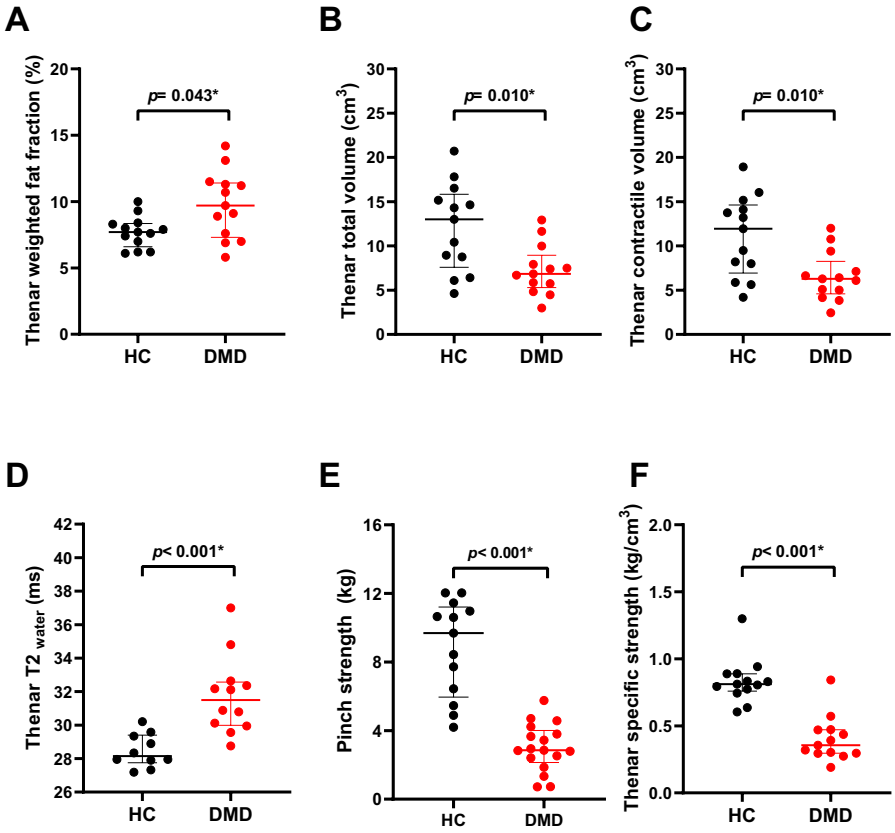


Figure 2. Dixon and pinch strength results of healthy controls (HC) compared to Duchenne muscular dystrophy (DMD) patients

Thenar weighted fat fraction (FF) (A), total volume (TV) (B), contractile volume (CV) (C), $T2_{water}$ (D), pinch strength (E), and specific strength (F), are presented for healthy controls (HCs; black) and DMD patients (red). TV, CV, specific strength and pinch strength were lower in DMD patients. $T2_{water}$ was higher compared to HCs, while FF was only slightly elevated. Statistical significance was set at $p<0.05$. Uncorrected *p*-values are reported and statistical significance after Bonferroni-Holm correction is shown by *

Relations between qMRI and clinical assessments

Table 3 shows correlations between qMRI parameters and clinical assessments. Apart from the smaller TV and CV, specific strength (kg/cm^3), was also significantly lower in patients (median 0.36 versus 0.81 kg/cm^3; Figure 2F). In both groups, FF did not correlate with pinch or grip strength (Figure 3B). In patients, pinch strength correlated strongly with TV (Figure 3C) and moderately with CV. There was a moderate correlation between grip strength and TV and CV, between PUL total and TV (Figure 3D), CV and $T2_{water}$, between PUL distal and TV and CV, between PROM total and TV and CV, and between NRS tiredness and TV. In HCs, correlations of pinch and grip strength with TV and CV were also strong or very strong.

Table 2. Clinical assessments in DMD patients at baseline and 12 months follow-up

DMD patients	Baseline (n=17)	12 months follow-up (n=13)	p-value
Pinch strength, kg	2.857 (2.137-4.010)	2.243 (1.930-3.339)	<0.001*
Grip strength, kg	8.47 (5.25-11.08)	6.39 (4.92-9.93)	0.016
PUL 2.0 total score, points (max 42)	29 (20-36)	23 (17-30)	<0.001*
PUL 2.0 shoulder score, points (max 12)	4 (0-8)	1 (0-6)	0.109
PUL 2.0 elbow score, points (max 17)	13 (10-16)	11 (7-14)	0.001 *
PUL 2.0 distal score, points (max 13)	11 (10-12)	11 (10-12)	0.227
PROM total score, points (max 64)	49 (36-57)	44 (30-54)	0.041
NRS tiring, points (max 10)	0 (0-1) $^{n=15}$	1 (0-4) $^{n=9}$	0.250
NRS difficult, points (max 10)	0 (0-2) $^{n=15}$	0 (0-1) $^{n=9}$	1.000

Values are median (first-third quartiles). Abbreviations: DMD = Duchenne muscular dystrophy, PUL = Performance of the Upper Limb, PROM = Patient Reported Outcome Measure, NRS = Numeric Rating Scale. Statistical significance was set at $p<0.05$. Uncorrected *p*-values are reported and statistical significance after Bonferroni-Holm correction is shown by *. If a certain value was not available for all patient, the number of patients for whom the data was available was presented after the result with *n* = number.

DISCUSSION

The thenar muscles were assessed for their value to monitor treatment effects in DMD using qMRI, strength and functional assessments. TV, CV, specific strength and pinch strength were lower and $T2_{water}$ was higher in DMD patients compared to HCs, while FF was only slightly elevated. Pinch strength and PUL total decreased significantly over 12 months in DMD patients, and there were moderate to strong correlations of qMRI muscle size with pinch strength and PUL total and distal. Despite this decrease in strength, operating a game controller was still possible for all our participants.

In DMD, hand function is preserved years after loss of ambulation and the ability to raise the arms [2]. Previous qMRI studies support this pattern of decline on tissue level by showing a proximal to distal involvement, where thigh, shoulder and upper arm muscles are, on average, more affected than lower leg and forearm muscles.[29, 30] Hand function and corresponding activities are vital for independence, e.g. operating a wheelchair and using a smart phone or computer, and for entertainment. Compared to ambulant patients, non-ambulant patients spend significantly more time (in total 2-4 hours/day) on playing (online) video games, which in recent years has become an important tool for social interaction.[7]

Drugs that preserve muscle and make use of dystrophin products are assumed to have the greatest effect on progression at an early disease stage, because they rely on the presence of sufficient muscle tissue.[4, 5] In addition, measurable disease progression is of vital importance to detect this potential preserving effect of a drug. All aspects of the thenar

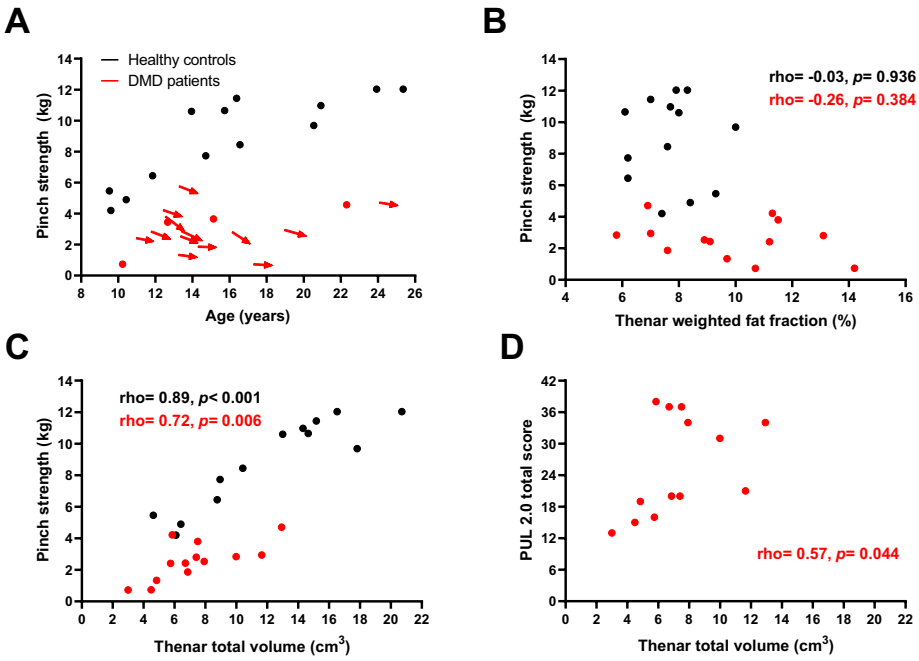


Figure 3. Pinch strength declines over time and correlates with quantitative MRI muscle size in Duchenne muscular dystrophy (DMD)
(A) Pinch strength is plotted versus age for both healthy controls (HC; black) and DMD patients (red). In all patients with 12 months follow-up data, pinch strength decreased over that time-period. (B) No clear relation between the plotted thenar fat fraction and pinch strength can be observed. Thenar total volume (TV) is plotted against pinch strength in (C) and Performance of the Upper Limb (PUL) 2.0 total score in (D). A strong correlation between pinch strength and TV in HCs and DMD patients can be observed, as well as a moderate correlation between PUL 2.0 total score and TV in DMD patients.

muscles that were assessed in this study pointed towards a relative preservation of the thenar muscles, and some also showed measurable disease progression: tissue characteristics as measured by qMRI, strength via pinch strength, function by assessing PUL and PROM, and activities in daily life via playing a video game. On tissue level, we found a slightly increased FF and an elevated $T2_{water}$ in the thenar muscles of DMD patients. $T2_{water}$ reflects active pathological processes, such as inflammation, myocyte swelling, edema and necrosis,[16] and previous qMRI studies of lower extremity muscles showed that young DMD patients have limited fat replacement and elevated $T2_{water}$.[15, 31] Other qMRI techniques might show early signs of muscle pathology and treatment effects in DMD, such as ionic dysregulation by sodium MRI,[32] or pH or phosphodiester alterations by ^{31}P or ^{1}H magenetic resonance spectroscopy.[33, 34] Thenar muscle volume was clearly reduced compared to HCs, which could be inherent to the dystrophic process or related to long-term corticosteroid use.[1, 35]

Table 3. Correlations between qMRI results and clinical assessments

	Fat fraction	Total volume	Contractile volume	$T2_{water}$
Healthy controls				
Pinch strength	rho= -0.03 *p*= 0.936	rho= 0.89 *p*< 0.001	rho= 0.89 *p*< 0.001	rho= -0.39 *p*= 0.266
Grip strength	rho= 0.03 *p*= 0.936	rho= 0.91 *p*< 0.001	rho= 0.91 *p*< 0.001	rho= -0.44 *p*= 0.200
DMD patients				
Pinch strength	rho= -0.26 *p*= 0.384	rho= 0.72 *p*= 0.006	rho= 0.68 *p*= 0.010	rho= 0.16 *p*= 0.618
Grip strength	rho= -0.17 *p*= 0.578	rho= 0.69 *p*= 0.009	rho= 0.67 *p*= 0.013	rho= -0.15 *p*= 0.633
PUL 2.0 total score	rho= -0.22 *p*= 0.480	rho= 0.57 *p*= 0.044	rho= 0.51 *p*= 0.075	rho= -0.59 *p*= 0.043
PUL 2.0 distal score	rho= -0.35 *p*= 0.249	rho= 0.58 *p*= 0.039	rho= 0.53 *p*= 0.065	rho= 0.10 *p*= 0.765
DMD Upper Limb PROM	rho= -0.32 *p*= 0.292	rho= 0.55 *p*= 0.051	rho= 0.52 *p*= 0.070	rho= -0.03 *p*= 0.931
NRS tiredness score	rho= 0.26 *p*= 0.415	rho= -0.54 *p*= 0.068	rho= -0.50 *p*= 0.101	rho= 0.12 *p*= 0.721
NRS difficulty score	rho= -0.18 *p*= 0.569	rho= -0.28 *p*= 0.384	rho= -0.24 *p*= 0.446	rho= 0.02 *p*= 0.951

Abbreviations: DMD = Duchenne muscular dystrophy, PUL = Performance of the Upper Limb, PROM = Patient Reported Outcome Measure, NRS = Numeric Rating Scale.

For the second aspect, muscle strength, we observed a decline in pinch strength over 12 months in our non-ambulant cohort, which is in agreement with the slightly older and weaker cohort described by Seferian et al.[36] As onset of weakness in pinch strength can already be quantified in preschool infants with DMD,[37] this process is apparently slow. The higher baseline pinch strength in our cohort compared to literature could be explained by the younger age at inclusion in the present study, and higher age at and shorter duration since loss of ambulation, and more consistent corticosteroid use.[36] The lower specific strength we observed, was also reported in other muscles of DMD patients, as well as in other muscular dystrophies.[30, 38, 39] The strong correlation between muscle size and strength suggests that therapies that manage to increase muscle volume have the potential to also increase muscle strength.[40]

The PUL 2.0 total has been accepted as a primary endpoint in non-ambulant DMD (e.g. NCT03406780 and NCT04371666). We showed quantifiable disease progression on function level in the upper extremity, reflected by a decline in the PUL 2.0 total within the same range of the non-ambulant cohort of Pane et al.[41] A decline in PUL distal was reported in a much

larger cohort over 24 months, but this was small (0.8 points), not reported as significant and therefore arguably not clinically relevant.[41] We did not observe a decline in the distal domain, supporting the preservation of hand function. The decline in PUL total was not reflected in a significant decline in PROM total score in our data, which could have been caused by our small cohort. Ultimately, it is important that therapies are shown to have an effect on activities in daily life, especially from a regulatory and patient's perspective.[4] Therefore, we studied the ability to play video games for 10 minutes using a game controller as clinical endpoint. None of the patients had reached this endpoint during our study, although some patients were not motivated to complete the 10 minutes or preferred smaller game controllers than used in the current setup. Therefore, the search for a direct connection between results on tissue characteristics, strength and function level with this or another robust clinical endpoint will be continued.

Obtaining qMRI data of the thenar muscles in non-ambulant DMD patients led to some challenges and technical considerations. Contractures of knee flexion, shoulder extension and internal rotation, forearm pronation and wrist flexion caused difficulties in maintaining the desired standardized position or any comfortable position in the MRI scanner. To address the effects of this on our outcome parameters, we studied the effect of two forearm positions on qMRI results separately, and found that thenar FF and $T2_{water}$ were comparable between both positions (supporting information). Even though TV and CV differed significantly between positions, average differences were small (0.55-0.63 cm^3). We also found that TV and CV correlated strongly with these same values determined on only the slice with the largest CSA (supporting information), indicating that (contractile) CSA can be used as a proxy for whole muscle values. Thenar FFs, however, should always be determined via whole muscle analysis, as correlations between single slice and whole muscle were moderate for thenar FF. This could be explained by the known proximal-distal differences in fat replacement in DMD.[42, 43] Furthermore, both the small size of the hand and positioning away from the center of the MRI scanner led to reductions in image quality. The best scan quality was observed when patients were positioned on the right side with the arm placed beside the body with either the elbow in 90° flexion and the forearm halfway between pronation and supination, or the elbow in maximum extension and the forearm in any comfortable position. However, the more severely affected patients in our study were often unable to lie on their right side, which makes the use of qMRI of the thenar muscles as biomarker challenging when using a conventional MR scanner. Drawing ROIs consistently may be challenging because of the small size of the thenar muscles, but the low interrater variability showed that this did not influence our qMRI results. The positioning away from the center of the MRI scanner resulted in both DMD patients and HCs in difficulties with obtaining sufficient B1 for the MSE scans. However, as the $T2_{water}$ values in all intrinsic hand muscles in DMD were also elevated, and values in our HCs are comparable to those of our previous study in upper arm and lower extremity muscles,[22] we are confident that the analysis was sound.

In conclusion, the minimal fat replacement within the thenar muscles and increased $T2_{water}$ indicate that the thenar muscles are in an early stage of muscle pathology in this cohort of non-ambulant patients. The simultaneous decrease in pinch strength and PUL total over one year shows that there is measurable disease progression within the possible duration of a clinical trial. Together with the correlation between muscle size and function, these results indicate that the thenar muscles are a valuable and quantifiable target for systemic or local therapy in later stages of the disease.

Acknowledgement

We acknowledge the Stichting Spieren for Spieren (grant number SvS15) for funding this work.

REFERENCES

1. Cros D, Harnden P, Pellissier JF, Serratrice G. Muscle hypertrophy in Duchenne muscular dystrophy. A pathological and morphometric study. Journal of neurology 1989;236:43-47. doi:10.1007/bf00314217
2. McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. Lancet 2018;391:451-461. doi:10.1016/S0140-6736(17)32160-8
3. Verhaart IEC, Aartsma-Rus A. Therapeutic developments for Duchenne muscular dystrophy. Nature reviews Neurology 2019;15:373-386. doi:10.1038/s41582-019-0203-3
4. CHMP. Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. In European Medicines Agency (EMA). Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf. Accessed December 18, 2020.
5. CDER, CBER. Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment. Guidance for Industry. https://www.fda.gov/media/92233/download. Accessed on May 1, 2019. US Food & Drug Administration (FDA) 2018
6. Connolly AM, Malkus EC, Mendell JR, et al. Outcome reliability in non-ambulatory boys/men with Duchenne muscular dystrophy. Muscle Nerve 2015;51:522-532. doi:10.1002/mus.24346
7. Heutinck L, Kampen NV, Jansen M, Groot IJ. Physical Activity in Boys With Duchenne Muscular Dystrophy Is Lower and Less Demanding Compared to Healthy Boys. J Child Neurol 2017;32:450-457. doi:10.1177/0883073816685506
8. Duncan SF, Saracevic CE, Kakinoki R. Biomechanics of the hand. Hand Clin 2013;29:483-492. doi:10.1016/j.hcl.2013.08.003
9. Barnard AM, Willcocks RJ, Triplett WT, et al. MR biomarkers predict clinical function in Duchenne muscular dystrophy. Neurology 2020;94:e897-e909. doi:10.1212/WNL.0000000000009012
10. Naarding KJ, Reyngoudt H, van Zwet EW, et al. MRI vastus lateralis fat fraction predicts loss of ambulation in Duchenne muscular dystrophy. Neurology 2020;94:e1386-e1394. doi:10.1212/WNL.0000000000008939
11. Rooney WD, Berlow YA, Triplett WT, et al. Modeling disease trajectory in Duchenne muscular dystrophy. Neurology 2020;94:e1622-e1633. doi:10.1212/WNL.0000000000009244
12. Hogrel JY, Wary C, Moraux A, et al. Longitudinal functional and NMR assessment of upper limbs in Duchenne muscular dystrophy. Neurology 2016;86:1022-1030. doi:10.1212/WNL.0000000000002464
13. Ricotti V, Evans MR, Sinclair CD, et al. Upper Limb Evaluation in Duchenne Muscular Dystrophy: Fat-Water Quantification by MRI, Muscle Force and Function Define Endpoints for Clinical Trials. PLoS One 2016;11:e0162542. doi:10.1371/journal.pone.0162542
14. Willcocks RJ, Triplett WT, Forbes SC, et al. Magnetic resonance imaging of the proximal upper extremity musculature in boys with Duchenne muscular dystrophy. Journal of neurology 2017;264:64-71. doi:10.1007/s00415-016-8311-0
15. Arpan I, Willcocks RJ, Forbes SC, et al. Examination of effects of corticosteroids on skeletal muscles of boys with DMD using MRI and MRS. Neurology 2014;83:974-980. doi:10.1212/WNL.0000000000000775
16. Carlier PG. Global T2 versus water T2 in NMR imaging of fatty infiltrated muscles: different methodology, different information and different implications. Neuromuscul Disord 2014;24:390-392. doi:10.1016/j.nmd.2014.02.009
17. van den Bergen JC, Ginjaar HB, van Essen AJ, et al. Forty-Five Years of Duchenne Muscular Dystrophy in The Netherlands. J Neuromuscul Dis 2014;1:99-109
18. Hamilton G, Yokoo T, Bydder M, et al. In vivo characterization of the liver fat (1)H MR spectrum. NMR Biomed 2011;24:784-790. doi:10.1002/nbm.1622
19. 2D phase unwrapping algorithms [computer program]. https://www.mathworks.com/matlabcentral/fileexchange/22504-2d-phase-unwrapping-algorithms. Retrieved November 1, 2020: MATLAB Central File Exchange, 2020.
20. Gold GE, Han E, Stainsby J, et al. Musculoskeletal MRI at 3.0 T: relaxation times and image contrast. AJR Am J Roentgenol 2004;183:343-351. doi:10.2214/ajr.183.2.1830343
21. Weigel M. Extended phase graphs: dephasing, RF pulses, and echoes - pure and simple. J Magn Reson Imaging 2015;41:266-295. doi:10.1002/jmri.24619
22. Keene KR, Beenakker JM, Hooijmans MT, et al. T2 relaxation-time mapping in healthy and diseased skeletal muscle using extended phase graph algorithms. Magn Reson Med 2020;84:2656-2670. doi:10.1002/mrm.28290
23. Gauld LM, Kappers J, Carlin JB, Robertson CF. Height prediction from ulna length. Dev Med Child Neurol 2004;46:475-480. doi:10.1017/s0012162204000787

24. Servais L, Deconinck N, Moraux A, et al. Innovative methods to assess upper limb strength and function in non-ambulant Duchenne patients. Neuromuscul Disord 2013;23:139-148. doi:10.1016/j.nmd.2012.10.022
25. Mayhew AG, Coratti G, Mazzone ES, et al. Performance of Upper Limb module for Duchenne muscular dystrophy. Dev Med Child Neurol 2019. doi:10.1111/dmcn.14361
26. Klingels K, Mayhew AG, Mazzone ES, et al. Development of a patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy: DMD Upper Limb PROM. Dev Med Child Neurol 2017;59:224-231. doi:10.1111/dmcn.13277
27. Wong DL, Baker CM. Pain in children: comparison of assessment scales. Pediatr Nurs 1988;14:9-17
28. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012;24:69-71
29. Forbes SC, Arora H, Willcocks RJ, et al. Upper and Lower Extremities in Duchenne Muscular Dystrophy Evaluated with Quantitative MRI and Proton MR Spectroscopy in a Multicenter Cohort. Radiology 2020;295:616-625. doi:10.1148/radiol.2020192210
30. Wokke BH, van den Bergen JC, Versluis MJ, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. Neuromuscul Disord 2014;24:409-416. doi:10.1016/j.nmd.2014.01.015
31. Hooijmans MT, Niks EH, Burakiewicz J, et al. Elevated phosphodiester and T2 levels can be measured in the absence of fat infiltration in Duchenne muscular dystrophy patients. NMR Biomed 2017;30. doi:10.1002/nbm.3667
32. Gerhalter T, Gast LV, Marty B, et al. (23) Na MRI depicts early changes in ion homeostasis in skeletal muscle tissue of patients with duchenne muscular dystrophy. J Magn Reson Imaging 2019;50:1103-1113. doi:10.1002/jmri.26681
33. Hooijmans MT, Doorenweerd N, Baligand C, et al. Spatially localized phosphorous metabolism of skeletal muscle in Duchenne muscular dystrophy patients: 24-month follow-up. PLoS One 2017;12:e0182086. doi:10.1371/journal.pone.0182086
34. Reyngoudt H, Turk S, Carlier PG. (1) H NMRS of carnosine combined with (31) P NMRS to better characterize skeletal muscle pH dysregulation in Duchenne muscular dystrophy. NMR Biomed 2018;31. doi:10.1002/nbm.3839
35. Schakman O, Gilson H, Kalista S, Thissen JP. Mechanisms of muscle atrophy induced by glucocorticoids. Horm Res 2009;72 Suppl 1:36-41. doi:10.1159/000229762
36. Seferian AM, Moraux A, Annoussamy M, et al. Upper limb strength and function changes during a one-year follow-up in non-ambulant patients with Duchenne Muscular Dystrophy: an observational multicenter trial. PLoS One 2015;10:e0113999. doi:10.1371/journal.pone.0113999
37. Mattar FL, Sobreira C. Hand weakness in Duchenne muscular dystrophy and its relation to physical disability. Neuromuscul Disord 2008;18:193-198. doi:10.1016/j.nmd.2007.11.004
38. Lokken N, Hedermann G, Thomsen C, Vissing J. Contractile properties are disrupted in Becker muscular dystrophy, but not in limb girdle type 2I. Ann Neurol 2016;80:466-471. doi:10.1002/ana.24743
39. Marra MA, Heskamp L, Mul K, et al. Specific muscle strength is reduced in facioscapulohumeral dystrophy: An MRI based musculoskeletal analysis. Neuromuscul Disord 2018;28:238-245. doi:10.1016/j.nmd.2017.11.017
40. Bettica P, Petrini S, D'Oria V, et al. Histological effects of givinostat in boys with Duchenne muscular dystrophy. Neuromuscul Disord 2016;26:643-649. doi:10.1016/j.nmd.2016.07.002
41. Pane M, Coratti G, Brogna C, et al. Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal data. PLoS One 2018;13:e0199223. doi:10.1371/journal.pone.0199223
42. Hooijmans MT, Niks EH, Burakiewicz J, et al. Non-uniform muscle fat replacement along the proximodistal axis in Duchenne muscular dystrophy. Neuromuscul Disord 2017;27:458-464. doi:10.1016/j.nmd.2017.02.009
43. Chrzanowski SM, Baligand C, Willcocks RJ, et al. Multi-slice MRI reveals heterogeneity in disease distribution along the length of muscle in Duchenne muscular dystrophy. Acta Myol 2017;36:151-162

Supplementary material

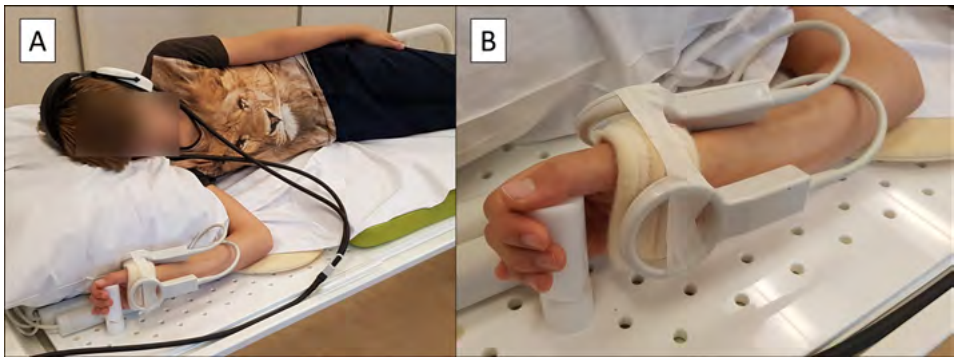


Figure S1. Example of a participant and forearm positioned in maximum pronation
In (A) the standard position is shown, where participants lay on the right side with the right shoulder in 90° flexion, with the right elbow in 90° flexion and the forearm in maximum pronation, which was supported by holding a handle. During the actual scanning procedure a sheet would be placed around the arm to prevent direct contact between the skin and the coils and cables. A detail of the 47 mm surface coils placed on the ventral and dorsal side of the thenar muscles is shown in (B).

Effect of different positioning

Healthy controls (HC) ≥18 years old without MRI contra-indications were included in a separate study to assess the effects of different positioning in the MRI scanner on thenar quantitative MRI (qMRI) results. Medical ethical approval for this protocol was waived, and written informed consents from participants were obtained.

Seven HCs were included. MRI scans of the right hand were acquired, while the participant was positioned on his right side with the right shoulder in 90° flexion, the elbow in 90° flexion, and the forearm in maximum pronation or halfway between pronation and supination. The Dixon scan was acquired in these two positions in all HCs, and the multi-echo spin-echo (MSE) scan was acquired in both positions in three of the HCs. Thenar fat fraction (FF), total volume (TV), contractile volume (CV) and T2 relaxation time of the muscle compartment ($T2_{water}$) were determined as described previously. We assessed differences between qMRI results in the different positions using Wilcoxon signed-rank test.

Thenar FF and $T2_{water}$ results were comparable in both positions (figure S2A and D). Both TV and CV differed significantly between the two positions, although the average difference was small with 0.55 cm^3 for TV and 0.63 cm^3 for CV (Supporting information figure S2B and C).

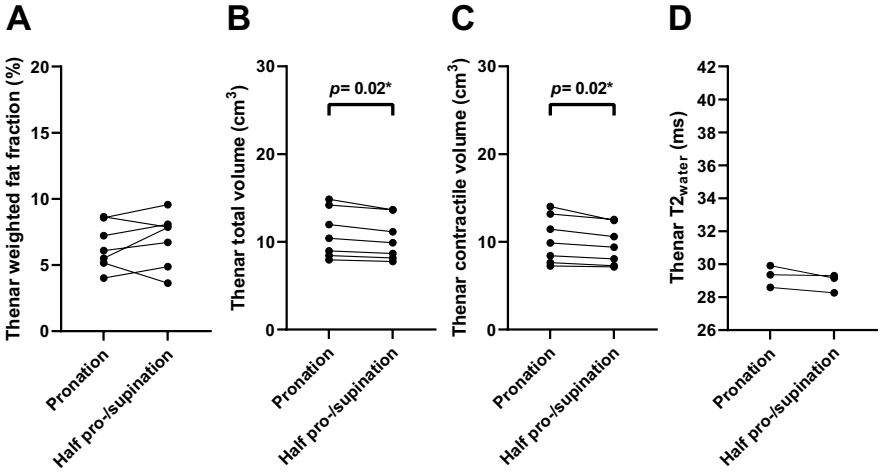


Figure S2. Thenar qMRI results in different forearm positions
Dixon and multi-echo spin-echo whole muscle thenar results are shown for seven HC. Participants are positioned with a 90° elbow angle, and the forearm in maximum pronation or halfway between pronation and supination. Both positions are shown in (A) for FF, (B) for TV, (C) for CV and (D) for $T2_{water}$. Thenar FF and $T2_{water}$ results are comparable in both positions. Total volume and contractile volume differ significantly between both positions, but the resulting difference is small.

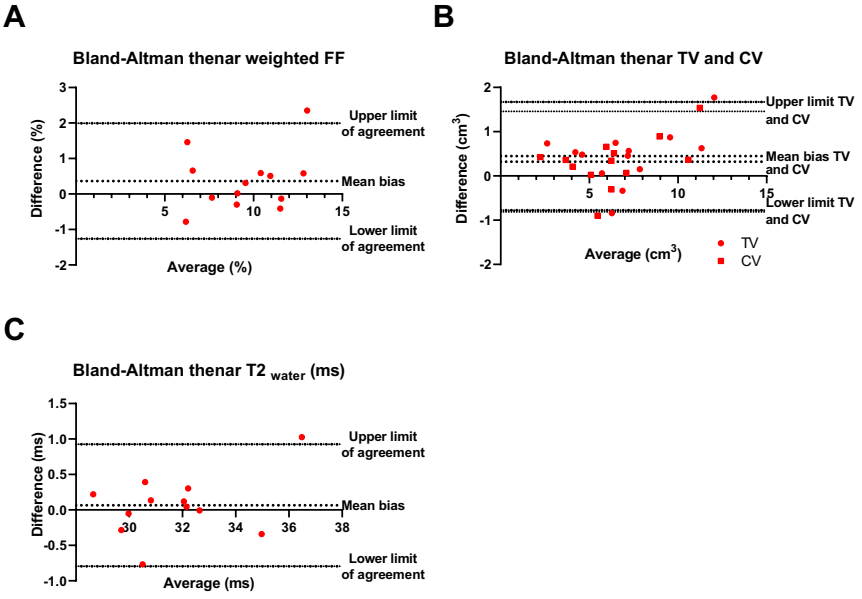


Figure S3 Interrater variability for thenar qMRI results
The thenar ROIs of DMD patients were drawn by two raters: K.J.N. and A.J.P. The Bland-Altman plots for the interrater variability in DMD patients is shown for FF in (A), TV and CV in (B) and $T2_{water}$ in (C). The mean bias and 95% limits of agreement for FF were 0.4% (-1.3 to 2.0%), for TV 0.4 cm^3 (-0.8 to 1.7 cm^3), for CV 0.3 cm^3 (-0.8 to 1.5 cm^3), and for $T2_{water}$ 0.1 ms (-0.8 to 0.9 ms).

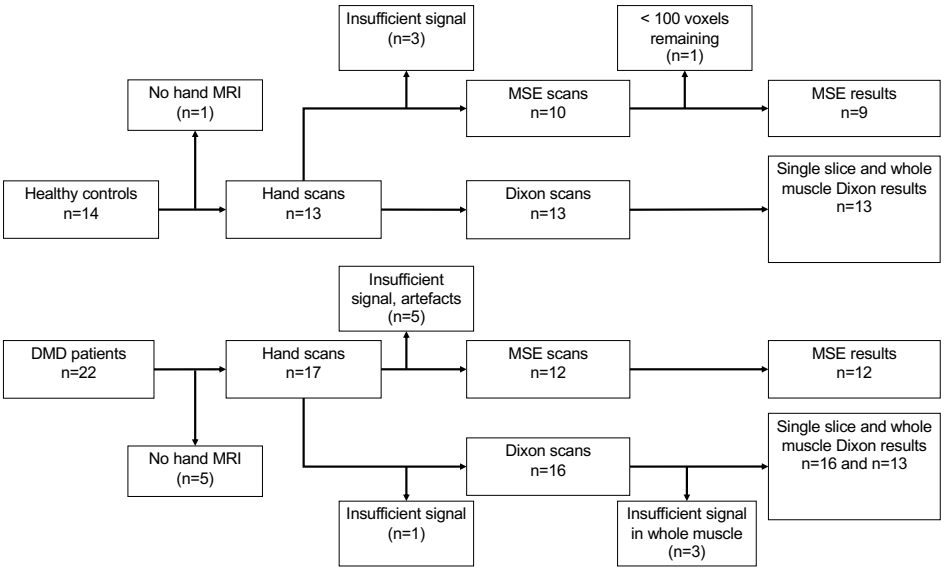


Figure S4 Flowchart of MRI data inclusion
Twenty-two DMD patients and 14 HC participated in the study. Five DMD patients could not undergo the MRI scans of the hand muscles: two patients decided not to undergo the MRI at the baseline visit, two patients were unable to complete the scans due to discomfort, and a technical failure of the MRI led to absence of the scans for the last patient. In one HC no MRI scans were acquired due to the relative contra-indication of dental braces. Four Dixon scans and five MSE scans of DMD patients were excluded due to insufficient signal or movement artefacts, and four MSE scans of HCs were excluded due to insufficient signal or because less than 100 thenar voxels with a sufficient B1 remained. In total 13 Dixon scans from HCs and 13 from DMD patients were included in the analysis of the whole thenar muscles. Similarly, MSE scans from nine HCs and 12 DMD patients were included in the analysis.

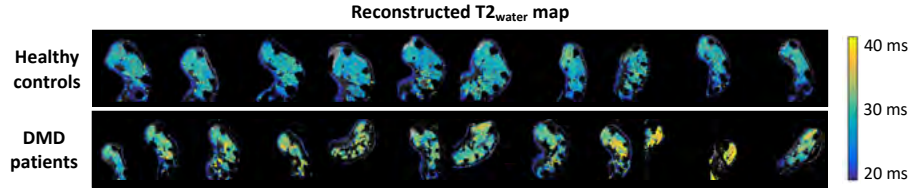


Figure S5 $T2_{water}$ of all hand muscles for HCs and DMD patients
The reconstructed $T2_{water}$ map for one slice per HC and DMD patient is shown in all participants to illustrate the higher $T2_{water}$ values in hand muscles in DMD patients versus HCs.

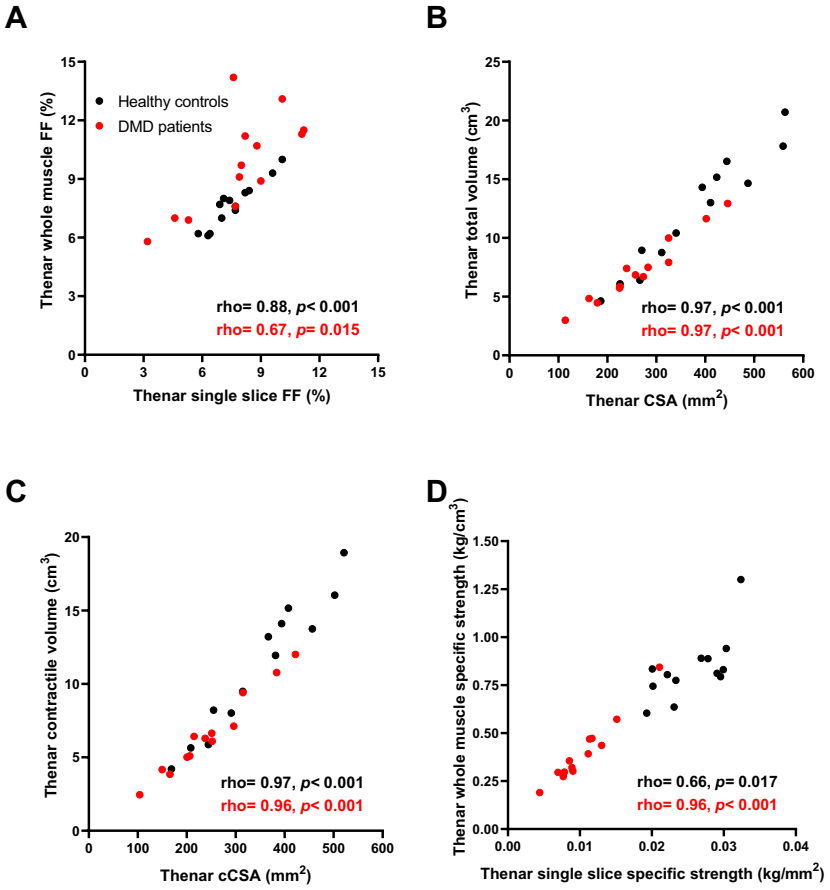


Figure S6 Single slice versus whole muscle Dixon thenar results
Dixon thenar whole muscle results are plotted versus the single slice results, for HCs (black; n=13) and DMD patients (red; n=13). The comparison is presented for FF in (A), CSA and TV in (B), cCSA and CV in (C), and specific strength in (D). All correlations between single slice and whole muscle values were strong, except FF in DMD patients and specific strength in HCs, which were moderate.

Single slice versus whole muscle results

If patient positioning would influence thenar qMRI results, this effect is expected to be larger if results would be determined on one slice instead of the whole muscle. In addition, to reduce scan time in the future, we assessed if qMRI results could be determined only on a single slice. To study this in all DMD patients and HCs, we also assessed all Dixon results for the slice with the largest cross-sectional area (CSA). This yielded a single slice thenar FF, CSA, contractile CSA (cCSA), and specific strength (pinch strength/cCSA). Spearman correlation coefficient was used to assess the correlation of the single slice and whole muscle Dixon results for HCs and DMD patients separately.

All correlations between single slice and whole muscle Dixon results were strong, except for FF in DMD patients and specific strength in HCs, which were moderate (figure S6A until D).

Chapter 6

The black box of technological outcome measures: an example in Duchenne muscular dystrophy

Published in: J Neuromuscul Dis
Jul 2022; 9(4); 555-569; DOI: 10.3233/JND-210767

Karin J. Naarding | Mariska M.H.P. Janssen | Ruben D. Boon | Paulina J.M. Bank
Robert P. Matthew | Gregorij Kurillo | Jay J. Han | Jan J.G.M. Verschuuren
Imelda J.M. de Groot | Menno van der Holst | Hermien E. Kan | Erik H. Niks

Abstract

Background

Outcome measures for non-ambulant Duchenne muscular dystrophy (DMD) patients are limited, with only the Performance of the Upper Limb (PUL) approved as endpoint for clinical trials.

Objective

We assessed four outcome measures based on devices developed for the gaming industry, aiming to overcome disadvantages of observer-dependency and motivation.

Methods

Twenty-two non-ambulant DMD patients (range 8.6-24.1 years) and 14 healthy controls (HC; range 9.5-25.4 years) were studied at baseline and 16 patients at 12 months using Leap Motion to quantify wrist/hand active range of motion (aROM) and a Kinect sensor for reached volume with Ability Captured Through Interactive Video Evaluation (ACTIVE), Functional Workspace (FWS) summed distance to seven upper extremity body points, and trunk compensation (KinectTC). PUL 2.0 was performed in patients only. A stepwise approach assessed quality control, construct validity, reliability, concurrent validity, longitudinal change and patient perception.

Results

Leap Motion aROM distinguished patients and HCs for supination, radial deviation and wrist flexion (range p=0.006 to <0.001). Reliability was low and the manufacturer's hand model did not match the sensor's depth images. ACTIVE differed between patients and HCs (p<0.001), correlated with PUL (rho=0.76), and decreased over time (p=0.030) with a standardized response mean (SRM) of -0.61. It was appraised as fun on a 10-point numeric rating scale (median 9/10). PUL decreased over time (p<0.001) with an SRM of -1.28, and was appraised as fun (median 7/10). FWS summed distance distinguished patients and HCs (p<0.001), but reliability in patients was insufficient. KinectTC differed between patients and HCs (p<0.01), but correlated insufficiently with PUL (rho = -0.69).

Conclusions

Only ACTIVE qualified as potential outcome measure in non-ambulant DMD patients, although the SRM was below the commonly used threshold of 0.8. Lack of insight in technological constraints due to intellectual property and software updates made the technology behind these outcome measures a kind of black box that could jeopardize long-term use in clinical development.

Introduction

duchenne muscular dystrophy (DMD) is typically characterized by progressive muscle weakness in a proximal to distal gradient.[1] Independent ambulation is generally lost years before upper arm function.[2] The progressive impairment of arm function causes difficulties in performing daily-life activities.[3] The first drugs for ambulant DMD patients have received conditional approval, but results on efficacy cannot be extrapolated to later disease stages due to progressive and irreversible reduction in targeted muscle tissue.[4, 5] Therefore, separate trials with dedicated outcome measures for non-ambulant patients need to be performed.

The Performance of the Upper Limb (PUL) 2.0 scale is currently the only outcome measure that is accepted as primary endpoint for non-ambulant DMD patients (e.g. NCT03406780 and NCT04371666).[6] However, the PUL has its limitations: it requires a clinical assessment, is observer-dependent, and has a floor and ceiling effect.[6, 7] Commercial technology with motion-tracking capabilities developed for gaming are being explored as new outcome measures for clinical trials in non-ambulant DMD.[8, 9] They potentially enable measurements at home, are observer-independent, and may overcome disadvantages of ordinal scales, such as the use of non-linear statistics. Furthermore, if the assessment can be performed in the form of a game, this could overcome variability due to lack of motivation in patients and lead to quantification of motions that are closer to the activities of daily living.

To address the lack of outcome measures in non-ambulant DMD, we evaluated several assessment methods based on off-the-shelf motion tracking technologies that could provide easy-to-use and affordable assessments in clinic. We chose Leap Motion which features marker-less hand tracking and Microsoft Kinect v2 which includes full-body tracking capabilities. The Leap Motion was used previously to assess the active range of motion (aROM) of the wrist and hand in healthy controls (HCs). The study revealed high test-retest reliability, but also issues of occlusions with some of the finger joints.[10, 11] Leap Motion aROM has not yet been evaluated as an outcome measure in neuromuscular disorders. For the Kinect v2 sensor, we evaluated three assessment protocols: 'Ability Captured Through Interactive Video Evaluation' (ACTIVE),[12] Functional Workspace (FWS),[13] and Kinect Trunk Compensation (KinectTC).[14] ACTIVE was developed as an outcome measure for neuromuscular diseases through the assessment of the reaching ability of the arm, summarized as a volume of reach during performance of a game activity. ACTIVE was shown previously to be responsive to treatment with nusinersen in spinal muscular atrophy (SMA) patients, and demonstrated excellent test-retest reliability in DMD.[9, 12] However, change over time has not yet been evaluated in DMD. In the FWS protocol, the Kinect is used to assess the ability to touch seven upper extremity body points via a guided video. The methodology for the analysis of the motion has been validated with a standard marker-based motion capture in HCs.[13] The FWS assessment has not been studied previously in DMD patient population. Finally, the KinectTC protocol was applied to assess the participants' trunk compensation during repeated task performance. Patients with neuromuscular

weakness often use their trunk to compensate for the loss of muscle strength in the upper extremity. In this study, DMD participants performed ten hand-to-mouth movements while being tracked by Kinect to quantify trunk compensation. In summary, we assessed the feasibility of Leap Motion aROM of the wrist and hand, ACTIVE, FWS and KinectTC, as outcome measures in non-ambulant DMD patients.

Materials and methods

Participants

Participants were included between March 2018 and July 2019 in a longitudinal study conducted at Leiden University Medical Center (LUMC; ABR number NL63133.058.17, www.toetsingonline.nl). For patients, visits were scheduled at baseline, 12 and 18 months which lasted about four hours each as approved by the medical ethical board. One half day visit was scheduled at baseline for HCs. Due to unforeseen restrictions during the COVID-19 pandemic, the 12 months follow-up visit could only take place for 16 patients and the 18 months follow-up visit for 12. Therefore, we report results from baseline (DMD and HC) and 12 months follow-up (DMD). DMD patients were recruited from the Dutch Dystrophinopathy Database,[15] and via outpatient clinics and patient organizations. Inclusion criteria were male, non-ambulant, genetically confirmed DMD, aged ≥8 years. Exclusion criteria were exposure to an investigational drug ≤6 months prior to participation and recent (≤6 months) upper extremity surgery or trauma. As the study protocol included muscle MRI, patients with MRI contra-indications (e.g. spinal fusion, daytime respiratory support, or the inability to lie still for 45 minutes) were also excluded. Healthy age-matched controls were recruited using posters and advertisements in local media. The study was approved by the local medical ethics committee in accordance with the ethical standards laid down in the 1975 Declaration of Helsinki and its later amendments. Written informed consent had been obtained from all patients and from legal representatives for patients under 16 years of age. Patient inclusion in this study has been reported previously.[16]

Measurements and data analysis

All measurements were performed seated behind a height-adjustable table. DMD patients sat in their own wheelchair and HCs sat on a chair with a backrest and armrests. All unilateral assessments were performed only with the right hand or arm. Height was calculated from the ulna length for both DMD patients and HCs with 18 as maximum age to create scaled scores for some of the outcome measures.[17] Patients performed assessment in the following order at every visit: ACTIVE, FWS, KinectTC, Leap Motion aROM and PUL.

Leap Motion active range of motion

A Leap Motion (Leap Motion Inc., San Francisco, USA) measurement setup was adjusted from Nizamis et al.,[10] with the sensor oriented downward or from the side, instead of upward

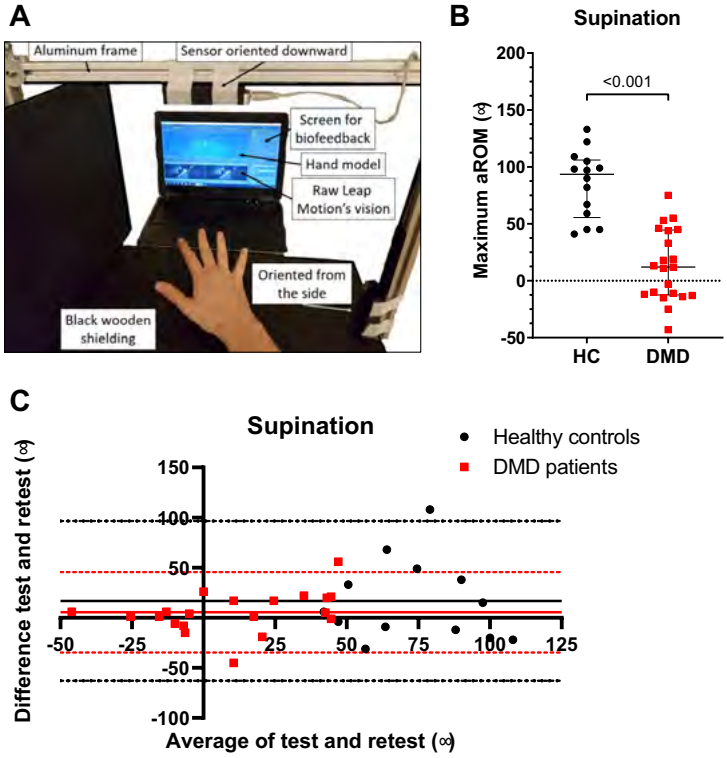


Figure 1. Leap Motion active range of motion (aROM) setup, construct validity and reliability
In (A) the Leap Motion measurement setup is shown with the Leap Motion sensor oriented downward or from the side by using an aluminum frame with mat black wooden shielding. Construct validity and reliability of supination are shown in (B) and (C) respectively. Supination showed the largest difference between Duchenne muscular dystrophy (DMD) patients and healthy controls (HC) and thus the best construct validity. The Bland-Altman plot with mean bias (straight lines) and 95%-confidence intervals (dotted lines) shows that reliability is low for both HCs (round) and patients (square). Average of the two trials is plotted on the x-axis and difference between the trials on the y-axis.

(Figure 1A). This enabled un-occluded tracking of primarily the dorsal side of the hands, and allowed patients to rest their hands on the setup during movement recording. Maximum active range of motion (aROM) was assessed using the Leap Motion in two separate trials of at least three repetitions for the following five movements of the right arm: flexion/extension of the finger joints, thumb abduction/adduction, radial/ulnar deviation, pronation/supination, and wrist flexion/extension. Trials without at least one complete movement were excluded.

The Leap Motion operates based on infra-red stereoscopy, consisting of two infrared cameras that capture motion with 200 frames per second and three infra-red LEDs that provide the illumination. Based on the reconstructed depth image, the Leap Motion software development kit (SDK) provides a skeletal model of the hand with the lower arm. In this

study, Orion Beta v3.2.1 SDK was used to extract the internal hand model. Movements of the elbow, wrist, finger and thumb points were recorded using Brekel Pro Hands software, version 1.35 (Brekel, Amsterdam, The Netherlands). The maximum aROM was determined for both extremes of the five movements by calculating the raw joint angles using custom-made software written in MATLAB (MATLAB R2016a, The MathWorks, Inc., Natick, USA). Screen recordings were captured to compare the raw Leap Motion's depth images with the provided hand model.

ACTIVE Scaled Volume

The reached volume of the arms was determined using the ACTIVE game (software version 2017) and a Microsoft Kinect v2 sensor (Microsoft Corp., Redmond, Washington, USA) mounted at a height of 1.80m and 2.95m in front of the participant. In this game, participants are virtually situated in a cave where they are stimulated to gather as many diamonds as possible. More diamonds can be collected when participants reach further upwards, sideways and forward with their arms and trunk as described previously (Figure 2A).[9] The maximum volume out of three ACTIVE trials of 60 seconds was used. If the last trial yielded the largest volume, a fourth trial was added to the protocol, assuming that the patient had not yet reached his maximal potential. The volume was normalized to create a Scaled Volume score using the participant's calculated height and the reached volume (Figure 2B).[9, 12]

Functional Workspace summed distance

The same Kinect sensor was used to assess the FWS. FWS determines the ability to reach different targets close to the body with the hand (simulating the motions of some common activities of daily living). Custom software was developed at University of California to collect the Kinect skeletal motion data.[13] During each FWS trial, participants were asked to reach with their right hand towards seven upper extremity targets: belt buckle or stomach, back pocket, ipsilateral shoulder, contralateral shoulder, mouth, top of head, and back of head (Figure 3A). Patients were instructed not to use their trunk to assist with their motion. The rigid body model was utilized to define the position for each of the seven landmark targets as described previously.[13] A second order, 1 Hz low-pass Butterworth filter was used to smooth the estimated distance timeseries. Since the tracking of the fingertips by the Kinect is relatively unreliable, the wrist trajectory was used to calculate the Euclidean distance to the expected wrist position at each target location. Based on the target sequence provided by the video instructions and the duration of the hand at each landmark, the minimal distance was extracted for each landmark. The distance was normalized by the individual's hand length to obtain the relative distance measure by using a ratio between hand length and ulna length that has been described for different ages.[18] All the data processing was performed in MATLAB (version R2016a). The summed FWS distance was calculated as the sum of the number of hand lengths distance to the seven targets, where higher scores indicated larger distances from the target.

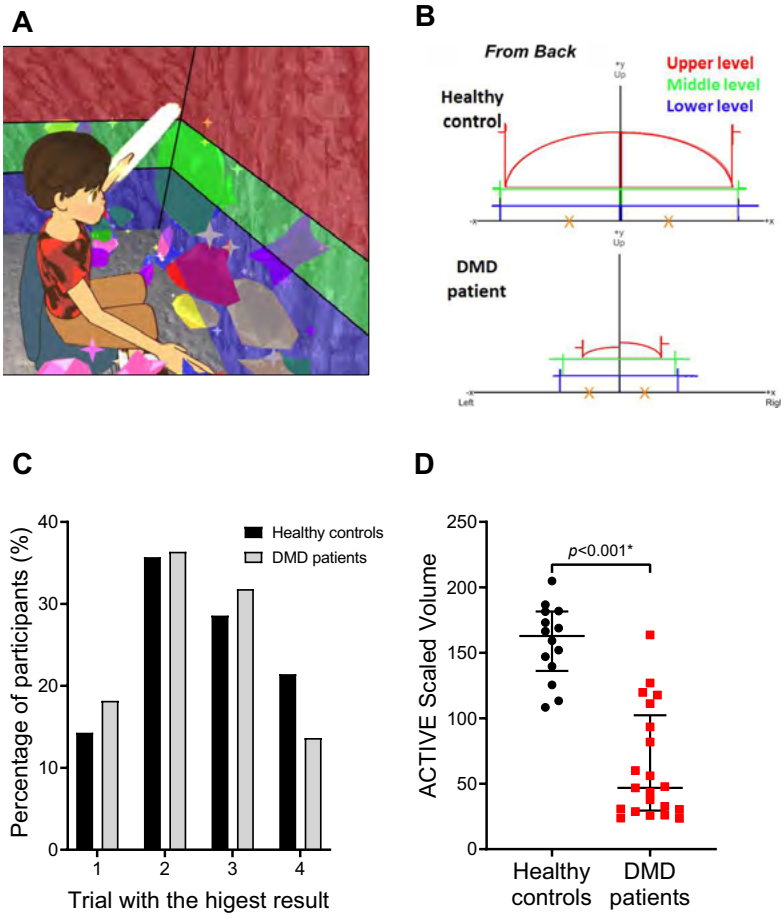


Figure 2. Ability Captured Through Interactive Video Evaluation (ACTIVE) setup and construct validity
In (A) the ACTIVE avatar is shown while the participant is pushing away the walls on the left on the three levels: upper, middle, and lower level. In (B) the reached width and height for the three levels are shown for a healthy control (HC) and Duchenne muscular dystrophy (DMD) patient. In (C) the percentage of participants who reached the largest volume in that trial is presented. The highest result was reached in the third and fourth trials in seven HCs (50%) and ten DMD patients (45%). In (D) the Scaled Volume is shown to be higher for HCs compared to DMD patients (p<0.001).

Kinect Trunk Compensation

KinectTC was assessed by quantifying the participants' trunk compensation during repeated task performance with the upper extremity using the same Kinect sensor. Participants performed ten hand-to-mouth movements using only their right hand whilst holding a 200g cup, similar to the PUL 2.0 hand-to-mouth item. In accordance with this PUL item, participants were instructed to use as little compensation as possible: sitting straight, keeping trunk and head still and primarily using the elbow flexion muscles. As a gold standard, the trunk

compensation (i.e. flexing, lateral flexing or extending) during the movement was visually assessed by a single observer (K.J.N.) and scored as present or absent. Kinect depth data was related to a model of body points by Microsoft Kinect SDK 2.0 software. Movements of wrist, elbow, shoulder, head, and spine body points were recorded using customized Unity3D software (version 5.6.0, Unity Technologies, San Francisco, USA). Trunk distance was defined as the total distance covered by the Kinect 'spine shoulder' point, which was quantified for all hand-to-mouth movement cycles using custom-made software in MATLAB (version R2016a; Figure 4A). Data were resampled to a uniformly distributed time series (30Hz), a 5Hz filter (fourth-order bidirectional Butterworth) was applied, and hand-to-mouth movement cycles were detected based on the distance between the wrist and head point.

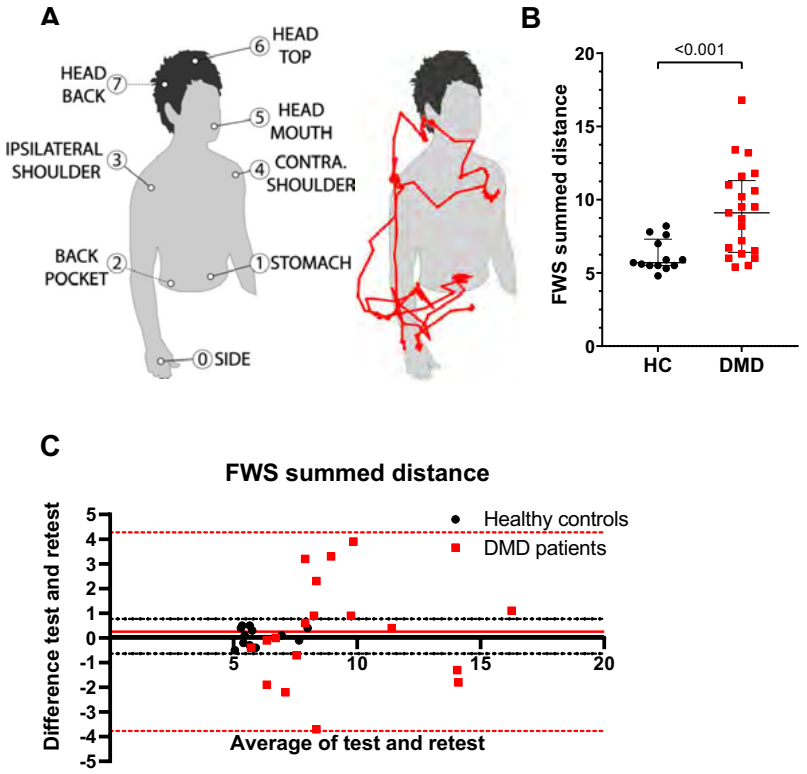


Figure 3. Functional Workspace (FWS) summed distance setup, construct validity and reliability
In (A) the seven upper extremity targets of the FWS are shown and alongside this a typical movement pattern of the right wrist during the FWS is presented in red. Construct validity and reliability of the summed hand length distance to these seven targets are shown in (B) and (C) respectively. This summed distance differed significantly between Duchenne muscular dystrophy (DMD) patients and healthy controls (HC; $p<0.001$). The Bland-Altman plot with mean bias (straight lines) and 95%-confidence intervals (dotted lines) shows that reliability is high for HCs (round), but much lower for patients (square). Average of the two trials is plotted on the x-axis and difference between the trials on the y-axis.

Using a 3D model of the body points, onset and/or offset of detected cycles was corrected manually, if necessary. Cycles without a dip in head-wrist distance and those with artefacts in the movements of the spine shoulder point were excluded. Trunk distance per cycle in mm was divided by the participant's calculated height in m to yield trunk compensation as a ratio to height. The KinectTC outcome was calculated as the average of at least five complete movement cycles without artefacts.

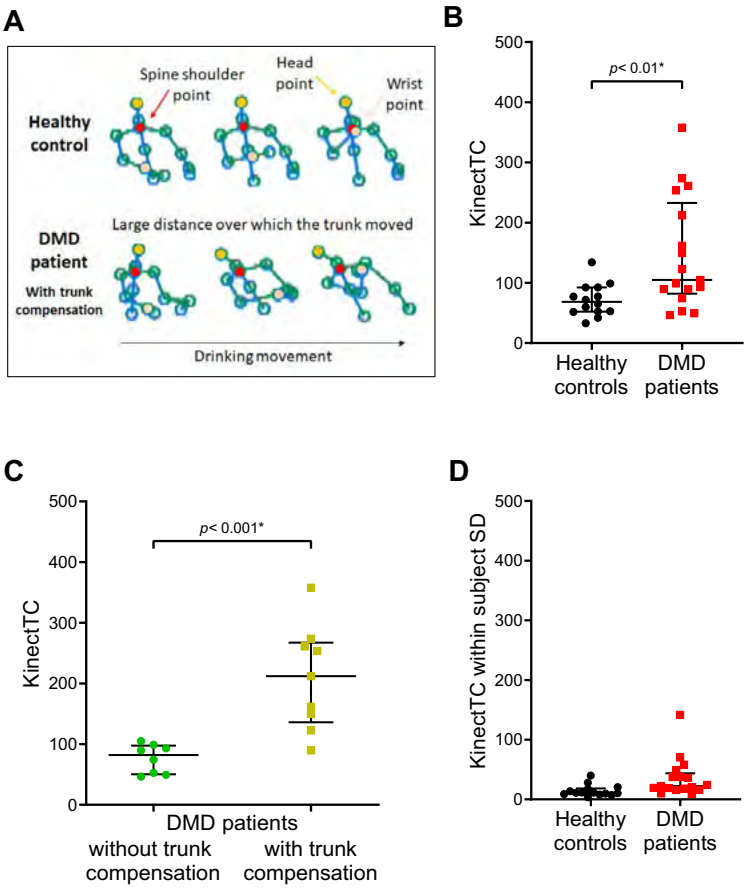


Figure 4. Kinect Trunk Compensation (KinectTC) setup, construct validity and reliability
Recorded spine shoulder, head point, wrist point and other body points for the hand-to-mouth movement of a healthy control (HC) and a Duchenne muscular dystrophy (DMD) patient with trunk compensation are shown in (A). KinectTC (trunk compensation in mm as a ratio to height in m) is shown in (B) to be significantly higher for DMD patients compared to HCs (p<0.01), and in (C) for DMD patients with visually scored trunk compensation compared to patients without visually scored compensation (p<0.001). In (D) the within subject SD is shown for HCs and DMD patients. The average within subject SD is 36 for patients and 14 for HCs.

Performance of the Upper Limb

PUL 2.0 was assessed in DMD patients only. PUL 2.0 was performed for the right arm and consists of 22 items that are divided into a shoulder (12 points), elbow (17 points) and distal wrist/hand dimension (13 points), yielding a maximum total score of 42 points.[19]

Patient perception

After the Leap Motion, ACTIVE, FWS and PUL, patients rated these assessments in the categories fun, annoying and tiring. This was done using 10-point numeric rating scales (NRS) ranging from 'not at all fun/annoying/tiring' (score 0) to 'a lot of fun/very annoying/tiring' (score 10) with matching facial cartoons based on the Wong-Baker Faces Rating Scale.[20]

Statistical analysis and stepwise approach

Leap, ACTIVE, FWS and KinectTC were evaluated in a stepwise approach that first assessed critical requirements for any outcome measure: quality control, construct validity and reliability. If results for this first step were of sufficient quality, the next steps consisted of [2] concurrent validity, [3] longitudinal change, and [4] patient perception. A flowchart of the stepwise approach for all four outcome measures is shown in Figure 5. Results are described as median (interquartile range (IQR) 1st quartile to 3rd quartile), unless otherwise stated.

In the first step, quality control consisted of describing excluded data and evaluating screen recordings for Leap Motion aROM for anatomical inconsistencies by visually comparing the simultaneously recorded raw Leap Motion's depth images and manufacturer's hand model. In case serious quality issues were encountered in part of the assessment, it could be decided to continue the stepwise approach with a specific part of the assessment for which consistent data were available. For ACTIVE, FWS and KinectTC, a comparison with screen recordings was not possible, because the depth images of the Kinect could not be obtained. Next, the construct validity criterion was tested. This was passed if outcomes differed significantly between patients and HCs, and for KinectTC between patients with and without visible trunk compensation using Mann-Whitney U tests. Statistical significance was set at $p<0.05$. Bonferroni-Holm correction was used to correct for multiple comparisons within the aROM assessments. Finally, reliability was assessed using a Bland-Altman analysis to determine mean bias and 95%-confidence interval (CI) of test-retest assessments. Test-retest data were available for Leap Motion aROM and FWS, which was deemed reliable if the 95%-CI in patients did not exceed the difference between patients and HCs. No test-retest data was available for ACTIVE and KinectTC, but reliability of ACTIVE has been determined previously.[9] KinectTC was deemed reliable, if the 95%-CI of the within subject standard deviation (SD) of movement cycles for patients was smaller than the difference between patients with and without trunk compensation.

In the second step, concurrent validity was determined via the correlation of the outcome measures with PUL 2.0 total score using Spearman correlation coefficient. The correlation with PUL should be strong (rho≥0.7).[21]

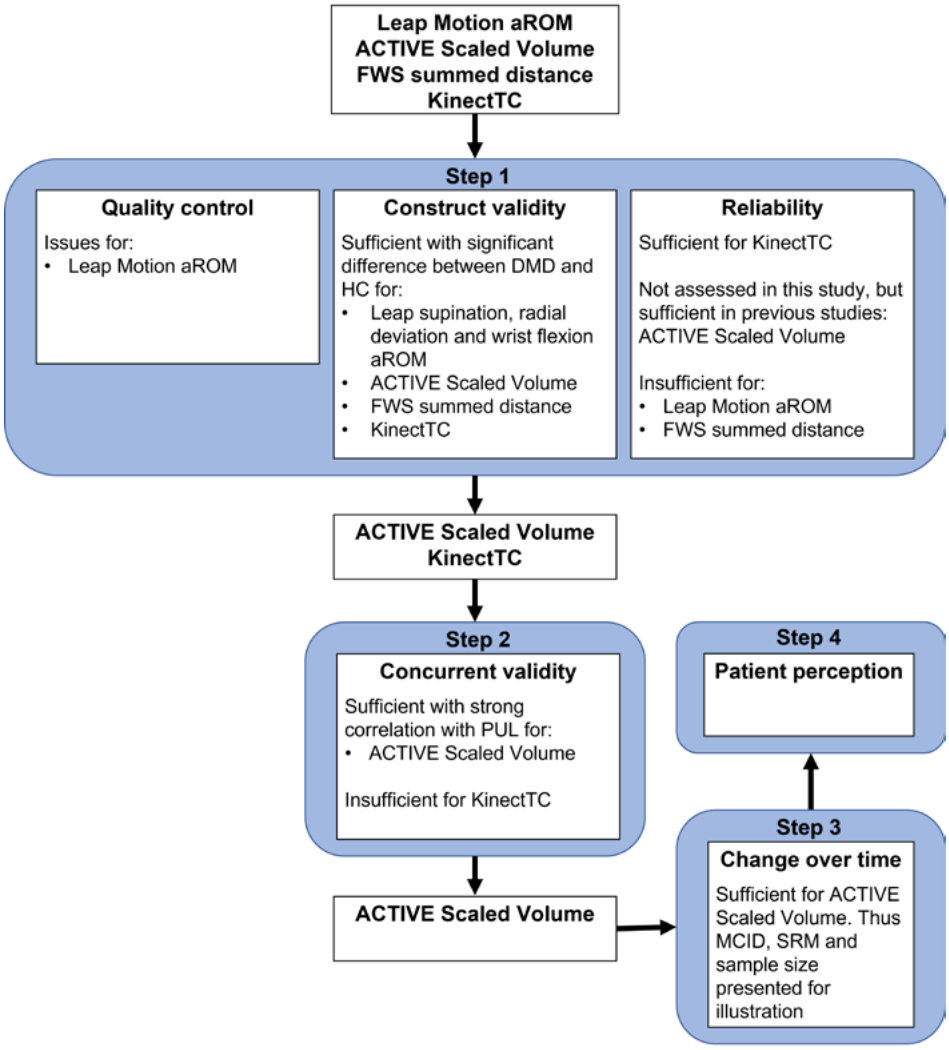


Figure 5. Flowchart of the stepwise approach for all four outcome measures
Leap, ACTIVE, FWS and KinectTC were assessed in a stepwise approach that first tested quality control, construct validity and reliability. If results for this step were of sufficient quality, the next steps consisted of: concurrent validity, longitudinal change and patient perception. Leap did not perform well enough on quality control and reliability and FWS on reliability, so these two measures did not continue after the first step. KinectTC did not have a strong relation with PUL and did not continue further. ACTIVE was analyzed according to the entire stepwise approach.

In the third step, change over time was satisfactory if outcomes showed significant change over 12 months as assessed by the Wilcoxon signed-ranks test. If the change was significant, the size of the change was illustrated using the minimally clinically important difference (MCID), standardized response mean (SRM) and corresponding sample size. MCID was determined via a distribution-based method using one-third of the SD of the baseline values. The SRM was calculated as the mean change over 12 months/SD of that change and should exceed 0.8.[22] Corresponding sample sizes for a potential clinical trial with the measurement as primary outcome measure were calculated using Lehr's formula.[23] In this calculation, we assumed a treatment effect of 50% reduction in disease progression over 12 months with a power of 80% and $\alpha<0.05$ in a 1:1 randomization. For comparison, all change over time values were also determined for PUL.

For the fourth and final step, patient perception was determined. Patient perception was assessed using NRS fun, annoying, and tiring scores of the three outcomes. These were compared to those of the PUL using Wilcoxon signed-ranks test.

Results

Participants

Twenty-two DMD patients and 14 HCs were included in the study. Baseline characteristics and results from all outcome measures are presented in Table 1. One patient with autism spectrum disorder was only able to perform the ACTIVE and PUL assessments. Four patients were unable to perform the described hand-to-mouth movement at baseline, having lost the ability at a median age of 14.5 years (range 8.9-18.2 years). Baseline median PUL total score was 21 points (IQR 19 to 34; Table 1). All patients used glucocorticoids in an intermittent schedule, except one patient who used daily deflazacort. One patient had ceased glucocorticoid treatment six weeks prior to baseline for a total of six months due to weight gain. Ambulation was lost median 2.5 years before baseline visit (range 0.6-5.8 years), at a median age of 11.5 years (range 8.0-18.9 years). Median age at start steroid use was 5.6 years (range 2.5-9.6 years).

Leap Motion active range of motion

Regarding quality control, thumb movements in only the abduction/adduction plane should have led to recorded movements in only one axis, but instead showed unexplainable movements in three axis in all participants. Therefore, these were excluded from further analysis. All Leap Motion test trials, and 98% (123/126) of retest trials in patients and 99% (83/84) in HCs contained a complete movement and were thus included in the analyses. No screen recordings had been captured for the first patient and five HCs. The metacarpophalangeal joints should show maximum flexion angles of about 90 degrees for all fingers,[10] but in our data the maximum flexion angles decreased from about 90 degrees

Table 1. Baseline characteristics and construct validity results from all outcome measures for HCs and DMD patients

	Healthy controls n=14	DMD patients n=22	p-value
Age, years	15.2 (11.5;20.6)	13.4 (12.3;16.2)	0.413
Height, m	1.74 (1.49;1.76)	1.52 (1.45;1.66)	0.016*
Body mass index	18.7 (16.6;22.2)	27.4 (23.6;30.8)	<0.001*
Leap Motion aROM			
Pronation, °	92 (84;112)	89 (82;96) *n=21*	0.400
Supination, °	94 (56;106)	12 (-13;45) *n=21*	**<0.001***
Radial deviation, °	20 (14;28)	10 (6;19) *n=21*	**0.002***
Ulnar deviation, °	42 (36;48)	46 (39;49) *n=21*	0.576
Wrist flexion, °	60 (49;64)	48 (31;53) *n=21*	**0.006***
Wrist extension, °	55 (45;64)	38 (22;53) *n=21*	0.043
ACTIVE Scaled Volume, points	163 (136;182)	47 (30;102) *n=21*	<0.001*
FWS summed distance, hand lengths	5.7 (5.5;7.3) *n=13*	9.1 (6.4;11.3) *n=21*	<0.001*
KinectTC	68 (52;92)	105 (82;233) *n=17*	<0.01*
PUL 2.0 total score, points		21 (19;34)	

Median (1st quartile; 3rd quartile). Differences between patients and HCs were assessed using Mann-Whitney U tests. Statistical significance was set at $p<0.05$ and is shown by *, for Leap Motion aROMs this is after Bonferroni-Holm correction and for clarity uncorrected *p*-values are reported. If a certain value was not available for all patients, the number of patients for whom the data was available was presented after the result with *n* = number. Abbreviations: HC = healthy control, DMD = Duchenne muscular dystrophy, aROM = active range of motion, ACTIVE = Ability Captured Through Interactive Video Evaluation, FWS = Functional Workspace, KinectTC = Kinect Trunk Compensation (trunk compensation in mm as a ratio to height in m), PUL = Performance of the Upper Limb.

for the index finger to about 70 degrees for the little finger in both patients and HCs (Figure S1). Due to these structural problems that could be caused by occlusion of the finger joints,[10] we continued the stepwise approach only for the wrist aROMs. Screen recordings revealed that in some patients the forearm position as recorded by the Leap did not match the position as seen in the screen recordings, which led to incorrect wrist flexion and extension values (Figure S2A). For supination, some patients moved similarly on the screen recordings in both trials, while test and retest values differed as much as 68 degrees due to different estimation of the elbow position (Figure S2B). These examples suggest that aROM values can be inconsistent while the data recordings show three complete movements that visually appear normal. Unfortunately, there was no quantitative method to filter out these incorrect recordings.

Regarding construct validity, significant differences between patients and HCs were found for supination, radial deviation and wrist flexion (Table 1 and Figure 1B), but not for pronation, ulnar deviation and wrist extension.

Table 2. Concurrent validity, change over time and patient perception results

	DMD patients, n=22							
	Correlation with PUL	**12-months change n=16**	**MCID**	**SRM**	**Sample size per study arm**	**NRS fun score n=21**	**NRS tiring score n=21**	**NRS annoying score n=21**
ACTIVE Scaled Volume, points	0.76 (0.47 to 0.90)	-5.6 (-23.4 to 1.3) [n=15]	14.1	-0.61	169	9 (7-10)	6 (4-7)	2 (0-5)
KinectTC	-0.69 (-0.88 to -0.29)	-	-	-	-	-	-	-
PUL 2.0 total score, points	-	-3.0 (-3.8 to -2.0)	2.9	-1.28	39	7 (5-10)	3 (2-4)	1 (0-2)

Correlation values are shown as rho (95%-confidence interval), and 12-months change and NRS scores as median (first-third quartiles). Abbreviations: DMD = Duchenne muscular dystrophy, PUL = Performance of the Upper Limb, MCID = minimally clinically important difference, SRM = standardized response mean, NRS = Numeric Rating Scale, ACTIVE = Ability Captured Through Interactive Video Evaluation, KinectTC = Kinect Trunk Compensation (trunk compensation in mm as a ratio to height in m).

Regarding reliability, radial deviation showed the smallest mean bias (0 degrees) and 95%-CI (-12 to 12 degrees) in patients, followed by wrist flexion (bias -5 degrees; 95%-CI -30 to 20 degrees) and supination (bias 5 degrees; 95%-CI -35 to 46 degrees). The 95%-CI of radial deviation and wrist flexion exceeded the difference between patients and HCs, while supination had only a slightly smaller 95%-CI. In HCs, radial deviation also showed the smallest mean bias (1 degrees) and 95%-CI (-15 to 16). The Bland-Altman plot of supination is presented in Figure 1C, because this aROM showed the largest difference between patients and HCs. The stepwise approach for Leap Motion aROM was not continued after the first step (Figure 5), because quality control showed unresolvable measurement problems and reliability was insufficient with a 95%-CIs that exceeded the differences between patients and HCs.

ACTIVE Scaled Volume

Regarding quality control, we included 96% (70/73) of ACTIVE trials, because all three trials of one patient's baseline visit had to be excluded due to a measurement error. The largest Scaled Volume was achieved in the third or fourth trial in 45% of DMD patients and 50% of HCs (Figure 2C). One patient did not want to perform more than one trial.
Regarding construct validity, ACTIVE Scaled Volume differed significantly between patients and HCs (Table 1 and Figure 2D).
Regarding concurrent validity, the correlation of ACTIVE Scaled Volume with PUL total was strong (rho=0.76; Table 2, Figure 6A).
Regarding change over time, ACTIVE Scaled Volume showed a decrease of median 5.6 points over 12 months (IQR -23.4 to 1.3; p=0.030; n=15), from median 47 (IQR 30 to 102) to median 44 (IQR 29 to 64). MCID for patients was 14.1 points for Scaled Volume (Table 2). The change in Scaled Volume exceeded the MCID in five out of 15 patients and the resulting SRM was -0.61, with a sample size of 169. PUL total changed from median 21 points (IQR 19 to 34) at baseline to median 19 (IQR 17 to 30) at 12 months, and this decrease was significant (-3.0 (IQR -3.8 to -2.0), p<0.001). The MCID for PUL total was 2.9 points, and the change exceeded the MCID in nine out of 16 patients. The SRM was -1.28 and corresponding sample size 39 (Table 2).
Regarding patient perception, ACTIVE was reported to be a lot of fun (median 9) and a little tiring (median 6), but not annoying (median 2). In comparison, patients appraised PUL 2.0 to be fun (median 7), not really tiring (median 3) and not at all annoying (median 1). Only NRS tiring scores differed significantly between ACTIVE and PUL (p=0.002), where ACTIVE was more tiring. All NRS results can be found in Table 2.

Functional Workspace summed distance

Regarding quality control, FWS test and retest trials were captured for 93% (14/15) of HCs. For patients, a test trial was available for 100% (21/21) and retest trial for 86% (18/21). For HCs, 87% (13/15) of retest trials were included, due to the exclusion of a retest trial of one participant who was able to reach all targets in both trials, while the wrist point did not move

in the skeletal data in the retest trial. Primarily in patients, some trials differed substantially between test and retest, but because no reference video was available, it was unclear whether this was caused by the differences in movements or body tracking issues of the Kinect.
Regarding construct validity, FWS summed distance differed significantly between patients and HCs (Table 1 and Figure 3B).
Regarding reliability in patients, Bland-Altman mean bias was 0.3 with 95% limits of agreement of -3.8 to 4.3 hand lengths (n=18; Figure 3C). FWS summed distance was more reliable in HCs with a mean bias of 0.1 and 95% limits of agreement of -0.68 to 0.8 hand lengths (n=12). Stepwise approach for FWS summed distance was not continued after the first step (Figure 5), because of insufficient reliability where the 95%-CI exceeded the difference between patients and HCs.

Kinect Trunk Compensation

Regarding quality control, movements of all participants contained at least five cycles which enabled calculation of the KinectTC value. We included 82% (139/170) of movement cycles for patients and 94% (132/140) for HCs. Some of the cycles that were excluded due to artefacts showed jumps in the spine shoulder point that occurred when the hand and arm occluded this point during the hand-to-mouth movement.
Regarding construct validity, KinectTC differed significantly between patients and HCs and between patients with and without visually scored trunk compensation (Table 1 and Figure 4B and 4C).
Regarding reliability, the average within subject SD was 36 for DMD patients and 14 for HCs (Figure 4D). The 95% CI (19-52) for patients was smaller than the difference between patients with and without trunk compensation.

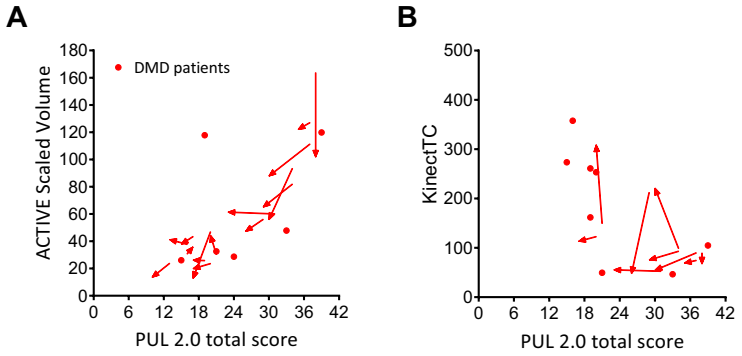


Figure 6. ACTIVE and KinectTC change over time and relation with function tests.
ACTIVE Scaled Volume plotted for baseline and 12 months follow-up of Duchenne muscular dystrophy (DMD) patients (red) against Performance of the Upper Limb (PUL) 2.0 total score in (A). Correlation was strong with PUL 2.0 (rho=0.76). ACTIVE Scaled Volume did decrease significantly over 12 months. KinectTC scaled trunk distance (trunk compensation in mm as a ratio to height in m) is plotted for baseline and 12 months follow-up of DMD patients against PUL 2.0 total score in (B). Correlation was moderate with PUL 2.0 (rho=-0.69).

Regarding concurrent validity, KinectTC showed a moderate correlation with PUL total (rho=-0.69, Figure 6B). Therefore the stepwise approach was not continued (Figure 5).

Discussion

We studied the feasibility of Leap Motion aROM, ACTIVE, FWS and a new measure, KinectTC, as outcome measures in non-ambulant DMD patients. At this time, current versions of Leap Motion aROM and FWS summed distance were shown to be unreliable as outcome measure for clinical trials in DMD. KinectTC correlated insufficiently with functional measures. Only ACTIVE showed promise as outcome measure in non-ambulant DMD due to its correlation with a functional outcome scale, decline over 12 months, and patient appraisal. However, the SRM was lower than for PUL.

Low reliability of Leap Motion aROM in both DMD patients and HCs can at least partly be explained by the incorrect estimation of the forearm and elbow motion. It was observed that the elbow was not in view during the wrist flexion/extension motion. This was a consequence of our choice to position the sensor on the side to allow patients to rest their hands. In the Leap Motion developer archive it is stated that the elbow position is estimated in case the elbow is not in view.[24] While this might not cause problems when playing games, this caused the Leap Motion to not provide sufficiently reliable results on wrist aROM to be used as outcome measure.

FWS summed distance was reliable in HCs, but not in DMD patients. This may be due to the fact that Kinect had more difficulty to reliably recognize participants and their motion in a wheelchair, and to recognize the more subtle movements of severely affected patients. Finally, we also observed that patients responded differently to the instruction not to use trunk compensation. A potential solution could be to instruct patients to move the hand to the seven targets as they would in daily life, and thus use as much compensation as they choose. On the other hand this could also introduce more difficulties for the Kinect to follow extreme movement of patients due to occlusion. The custom software with the rigid model and analysis in MATLAB are still under development and further adjustments could improve the reliability in patients.[13] The concept to track different movements of the upper extremities close to the body that simulate functional movements used frequently in daily-life seems worthwhile to explore further whilst trying to improve the software and analysis.

KinectTC was applied as the first outcome to quantify compensation strategies in DMD. In our study, it fell short as outcome measure, because it did not show sufficient correlation with functional measures.
Occlusion is one of the challenges in using camera-based marker-less tracking devices, such as Leap Motion and Kinect. In KinectTC, small and sometimes larger jumps in the spine

shoulder point were observed when the hand and arm occluded this point during the hand-to-mouth movement. The Kinect was also better positioned to register lateral flexion than flexion or extension of the trunk, although Kinect was shown to be comparable to 3D motion analysis in assessing both lateral and anteroposterior movements.[25] KinectTC did seem to give insight into use of trunk compensation in patients with a PUL total score of about 18-36 points. Since a measure similar to KinectTC was able to show response to the use of an arm support in three DMD patients,[14] KinectTC could provide useful additional data in clinical care to support decisions of therapies and supportive devices. However, our data do not support its use as outcome measure in clinical trials.

ACTIVE showed the most promise as an outcome measure in non-ambulant DMD. Our study supports that patients should perform at least three or four ACTIVE trials, since 45% of patients did not achieve the highest Scaled Volume in the first or second attempt. In a previous study in SMA patients, only two trials were performed per participant.[12] In our study we could not determine test-retest reliability, but this was previously demonstrated to be excellent in a small population of eight DMD patients.[9] ACTIVE was responsive to disease progression, but the SRM was lower than the commonly used threshold of 0.8, and also lower when compared to the PUL. As a consequence, the sample size when using ACTIVE (169) was also much larger than that for PUL (39).[22] In the study in SMA patients using ACTIVE, the MCID was 4.5-10.9 and the predicted sample size was 28 patients.[12] This MCID was smaller than our value of 14.1, which is potentially caused by our diverse patient population leading to large baseline SD. Their predicted required sample size of 28 patients was also much smaller than ours of 169 patients, most likely because our calculation was based on a 50% reduction in disease progression, while the SMA calculation was based on an improvement of median 15.9 points caused by nusinersen. In terms of enjoyability, patients showed no clear preference for ACTIVE or PUL, but PUL was reported less tiring than ACTIVE. The order of assessments is unlikely to have influenced these results, since PUL was performed later than ACTIVE. Also ACTIVE and PUL were both suited for our patient with autism spectrum disorder. ACTIVE showed promise as outcome measure in non-ambulant DMD, but sensitivity to change was lower than the commonly used threshold.

Hardware from the gaming industry has a limited production time, and the production of the Kinect v2 sensor by Microsoft that was applied in this study was discontinued in 2015.[26] There are other sensors available that could also be used to determine the same outcomes. Some sensors require the use of markers, which is perhaps less time-efficient than using marker-less sensors. While switching to a different sensor is possible, a validation process or separate study should be conducted to determine characteristics of the outcome measure when using the new sensor, such as construct validity, reliability, concurrent validity and change over time.

Use of software from the gaming industry has limitations. The provided SDK software is not adjusted to correctly track persons with particular limitations, which poses challenges when

using these devices in patient populations. For instance, DMD patients were sometimes not recognized by the SDK for ACTIVE, FWS and KinectTC software, because they sat in a wheelchair with armrests and a headrest from which they were hard to distinguish by the software. The presence of intellectual property poses an additional limitation. Constraints have been found in the Leap Motion software for the hand model that we used, Orion Beta v3.2.1.[10] We were not allowed to get details about these constraints, and the terms and conditions for use of the software did not allow modification of the SDK to analyze the raw data differently, unless a Development License was procured. The current supplied version, v4.1.0, is different from the one we used, but a clear list of changes is not provided.
The ACTIVE software was developed by researchers from Research Institute at Nationwide Children's Hospital, Ohio, USA. The used software (version 2017) is still operational, but was updated in 2019, after the start of this study. In this 2019 software version a new avatar was added, which provides real-time feedback to patients. This potentially leads to similar problems of comparability between results obtained with different software versions. When the analysis algorithm is not transparent, the same software version should be used continuously in clinical trials, or a new study should be conducted each time a new software version is used, to obtain data on the properties of this adjusted outcome measure. The ACTIVE yields a total volume and volume for the different levels, but data cannot be checked afterwards by replaying the movement of the wrist or hand points that are used to acquire these volumes. Transparency of analysis algorithms would enable updates, for instance to deal with bugs, if a standard validation process is in place to ensure that the changes made do not affect the outcomes.

The data presented illustrate the black box of commercial software or software otherwise protected by intellectual property, and the obstacles when using this software for sustainable scientific applications, such as use for outcome measures in future clinical trials. The term black box is used increasingly in this era of big data and medical algorithms.[27] For researchers, to develop a new device and get approval to use it as primary outcome measure in clinical trials is a lengthy process, as shown by the stride velocity 95th centile.[28] Commercial parties are able to develop and improve devices and software fast and the research field and patient organizations are looking to profit from this speed by collaborating. Examples are the recent World Duchenne Organization meeting about the use of wearables and a study using Apple watches to collect activity data, potentially to use as outcome measure.[29] The presented obstacles for use of hardware from commercial entities or software protected by intellectual property should incite the debate on future directions for the outcome measure research field and possible solutions for this black box, such as transparency of analysis algorithms.

Limitations of this study are the small study cohort for which not all follow-up visits could take place due to the COVID-19 pandemic. The change over time results were therefore based on small numbers. Although some trials and movement cycles for the different outcomes had to be excluded, this led only to exclusion of <5% of all data gathered.

In summary, of the evaluated technological outcome measures in their current iteration of development or version, ACTIVE showed the most promise due to a strong correlation with functional measures, change over 12 months and appraisal of being fun. However, the SRM was lower than commonly used thresholds. PUL met all criteria satisfactorily and had a SRM above this threshold. Outcome measures based on hardware and software from the gaming industry can indeed overcome problems such as observer dependence and lack of motivation. However, lack of insight in detailed operations of the software and hardware compounded by intellectual property constraints, and possible software updates and hardware discontinuation, make these outcome measures a black box and could jeopardize their long-term applicability in clinical trials.

Acknowledgement

This study was supported by Stichting Spieren for Spieren (grant number SvS15).

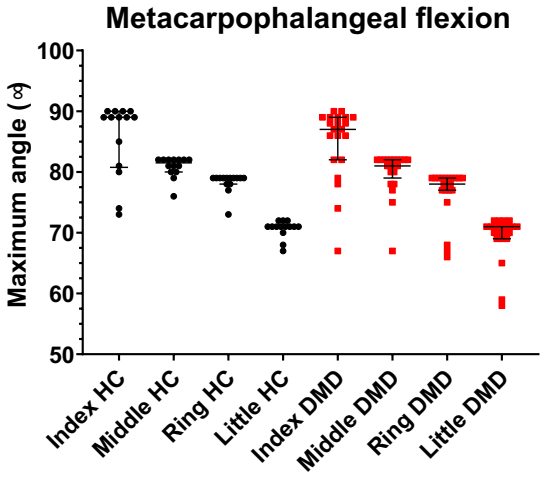
References

1. Birnkrant DJ, Bushby K, Bann CM, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and neuromuscular, rehabilitation, endocrine, and gastrointestinal and nutritional management. Lancet Neurol 2018;17:251-267. doi:10.1016/S1474-4422(18)30024-3
2. McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. Lancet 2018;391:451-461. doi:10.1016/S0140-6736(17)32160-8
3. Brooke MH, Fenichel GM, Griggs RC, et al. Duchenne muscular dystrophy: patterns of clinical progression and effects of supportive therapy. Neurology 1989;39:475-481. doi:10.1212/wnl.39.4.475
4. CHMP. Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. In European Medicines Agency (EMA). Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf. Accessed December 18, 2020.
5. Verhaart IEC, Aartsma-Rus A. Therapeutic developments for Duchenne muscular dystrophy. Nature reviews Neurology 2019;15:373-386. doi:10.1038/s41582-019-0203-3
6. Pane M, Mazzone ES, Fanelli L, et al. Reliability of the Performance of Upper Limb assessment in Duchenne muscular dystrophy. Neuromuscul Disord 2014;24:201-206. doi:10.1016/j.nmd.2013.11.014
7. Pane M, Coratti G, Brogna C, et al. Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal data. PLoS One 2018;13:e0199223. doi:10.1371/journal.pone.0199223
8. Han JJ, de Bie E, Nicorici A, et al. Reachable workspace and performance of upper limb (PUL) in duchenne muscular dystrophy. Muscle Nerve 2016;53:545-554. doi:10.1002/mus.24894
9. Lowes LP, Alfano LN, Crawfis R, et al. Reliability and validity of active-seated: An outcome in dystrophinopathy. Muscle Nerve 2015;52:356-362. doi:10.1002/mus.24557
10. Nizamis K, Rijken NHM, Mendes A, et al. A Novel Setup and Protocol to Measure the Range of Motion of the Wrist and the Hand. Sensors (Basel) 2018;18. doi:10.3390/s18103230
11. Gamboa E, Serrato A, Castro J, et al. Advantages and Limitations of Leap Motion from a Developers', Physical Therapists', and Patients' Perspective. Methods Inf Med 2020;59:110-116. doi:10.1055/s-0040-1715127
12. Alfano LN, Miller NF, Iammarino MA, et al. ACTIVE (Ability Captured Through Interactive Video Evaluation) workspace volume video game to quantify meaningful change in spinal muscular atrophy. Dev Med Child Neurol 2020;62:303-309. doi:10.1111/dmcn.14230
13. Matthew RP, Seko S, Kurillo G, et al. Reachable Workspace and Proximal Function Measures for Quantifying Upper Limb Motion. IEEE J Biomed Health Inform 2020;24:3285-3294. doi:10.1109/JBHI.2020.2989722
14. Kooren PN, Dunning AG, Janssen MM, et al. Design and pilot validation of A-gear: a novel wearable dynamic arm support. J Neuroeng Rehabil 2015;12:83. doi:10.1186/s12984-015-0072-y
15. van den Bergen JC, Ginjaar HB, van Essen AJ, et al. Forty-Five Years of Duchenne Muscular Dystrophy in The Netherlands. J Neuromuscul Dis 2014;1:99-109
16. Naarding KJ, Doorenweerd N, Koeks Z, et al. Decision-Making And Selection Bias in Four Observational Studies on Duchenne and Becker Muscular Dystrophy. J Neuromuscul Dis 2020;7:433-442. doi:10.3233/JND-200541
17. Gauld LM, Kappers J, Carlin JB, Robertson CF. Height prediction from ulna length. Dev Med Child Neurol 2004;46:475-480. doi:10.1017/s0012162204000787
18. Aldegheri R, Agostini S. A chart of anthropometric values. J Bone Joint Surg Br 1993;75:86-88. doi:10.1302/0301-620X.75B1.8421044
19. Mayhew AG, Coratti G, Mazzone ES, et al. Performance of Upper Limb module for Duchenne muscular dystrophy. Dev Med Child Neurol 2019. doi:10.1111/dmcn.14361
20. Wong DL, Baker CM. Pain in children: comparison of assessment scales. Pediatr Nurs 1988;14:9-17
21. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012;24:69-71
22. Cohen J. Statistical power analysis for the behavioral sciences.: Lawrence Erlbaum Associates, 1988.
23. Morrow JM, Sinclair CD, Fischmann A, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. Lancet Neurol 2016;15:65-77. doi:10.1016/S1474-4422(15)00242-2
24. Arms. https://developer-archive.leapmotion.com/documentation/csharp/devguide/Leap_Overview.html. Accessed on May 28, 2020. [online].
25. Clark RA, Pua YH, Fortin K, et al. Validity of the Microsoft Kinect for assessment of postural control. Gait Posture 2012;36:372-377. doi:10.1016/j.gaitpost.2012.03.033
26. Microsoft. Microsoft to consolidate the Kinect for Windows experience around a single sensor. Accessed

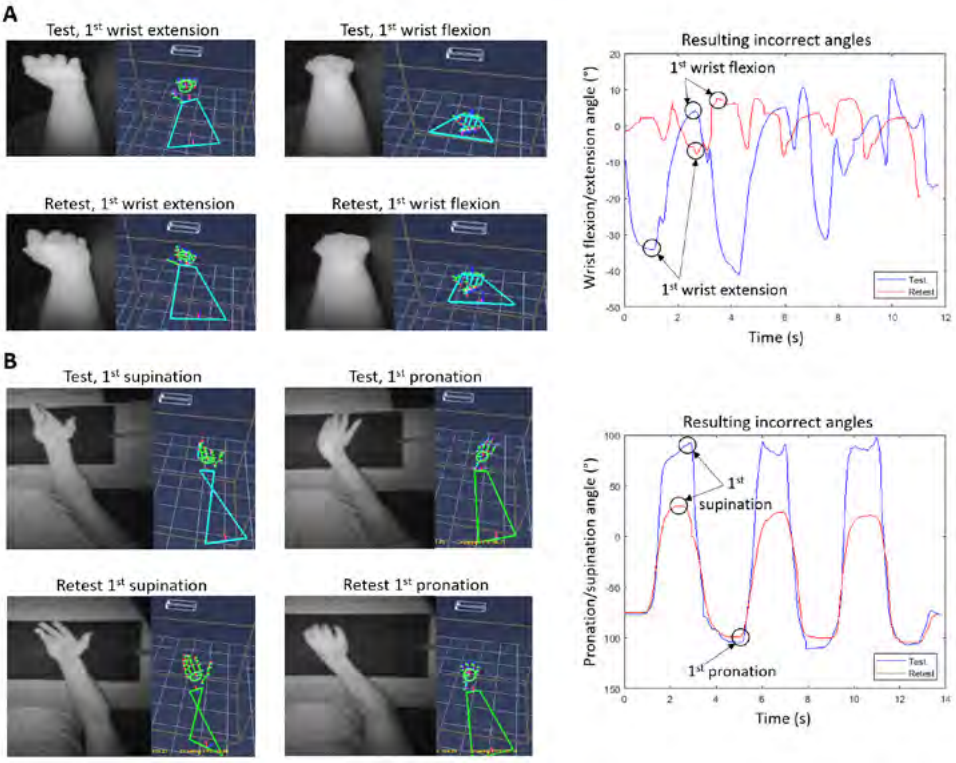
on May 5, 2020 at: https://docs.microsoft.com/en-us/archive/blogs/kinectforwindows/microsoft-to-consolidate-the-kinect-for-windows-experience-around-a-single-sensor [online].
27. Price WN. Big data and black-box medical algorithms. Sci Transl Med 2018;10. doi:10.1126/scitranslmed.aao5333
28. CHMP. Qualification opinion on stride velocity 95th centile as a secondary endpoint in Duchenne Muscular Dystrophy measured by a valid and suitable wearable device*. In European Medicines Agency (EMA). Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-stride-velocity-95th-centile-secondary-endpoint-duchenne-muscular-dystrophy_en.pdf. Accessed June 16, 2021.
29. UK D. KINEDMD is a study developing an activity monitoring biomarker. Available at: https://www.duchenneuk.org/outcome-measures/. Accessed July 30, 2021. [online].

Supplementary material

Below two supplementary figures can be found.



Supplementary Figure 1. Structural problems in Leap Motion assessments
The metacarpophalangeal joints seem to show a maximum flexion angle which decreases per finger with the largest maximum flexion angle for the index finger and the smallest angle for the little finger. This decrease was seen in both healthy controls (HC; black) and Duchenne muscular dystrophy (DMD) patients (red).



Supplementary Figure 2. Measurement problems in Leap Motion assessments
In black-and-white the depth images from the screen recordings are shown and next to this is the registered hand model at the same time. The graphs present resulting angles over the time-period of the measurement. (A) Wrist flexion/extension is performed at test and retest by a DMD patient, showing that the forearm moves with the hand in the hand model, while it is kept still in reality and the wrist itself only flexes or extends a small amount. This is more extreme in the retest assessment and leads to very low wrist flexion and extension angles. This retest assessment was excluded because no representative complete movements with a peak and dip could be recognized. (B) Supination angles of a HC are 68 degrees lower at pronation/supination retest compared to the 1st test, due to incorrect measurement of the elbow position during retest.

Chapter 7

Summary and general discussion

Summary

The overall aim of this thesis was to identify outcome measures in Duchenne muscular dystrophy (DMD), specifically for non-ambulant patients, that are able to detect a clinically relevant difference in a relatively short period of time compatible with the duration of a clinical trial. The use of such outcome measures could lead to smaller sample sizes in such trials with a lower burden for patients.

In **chapter 2**, we reviewed the considerations provided by patients and/or caregivers for not taking part in three observational studies on patients with DMD and one study on patients with Becker muscular dystrophy (BMD). We first assessed if age, travel-time, *DMD* gene mutation and age at loss of ambulation, derived from the national patient registry, the Dutch Dystrophinopathy Database, were comparable between participants and non-participants. This showed that participating patients were overall representative of the eligible sub-population for their study. Exceptions were the lack of patients with distal mutations upstream of exon 63 in all studies, and a younger age of participants in the study that investigated upper extremity outcome measures in non-ambulant DMD patients (chapter 4, 5 and 6). This suggests that studying more advanced disease stages in DMD could be more challenging. The most frequently reported considerations were 'Burden of protocol' (38%), ' MRI' (30%), and 'Travel-time' (19%).
Our results highlighted that nationwide patient recruitment registries can be used to compare participants and non-participants to ensure that observational research is representative of the whole patient cohort. Furthermore, the results suggest that to facilitate and increase patient participation several factors could be addressed: 1) optimizing involvement of patients in the design of new studies, 2) improving the MRI experience, and 3) integrating observational research and clinical care.

In **chapter 3**, we assessed the additive predictive value of vastus lateralis (VL) fat fraction (FF) measured using quantitative MRI (qMRI) to age on loss of ambulation (LoA) in two cohorts: one from the Leiden University Medical Center (LUMC; n=19) and the other from Cincinnati Children's Hospital Medical Center (CCHMC; n=15). We found an excellent interobserver reliability for VL FF determined by qMRI, which supports the feasibility of including muscle qMRI data from multiple centers in studies in DMD. We found VL FF to have added predictive value to age on LoA in the cohort from the LUMC (hazard ratio 1.15, 95% confidence interval 1.05-1.26, p=0.003). This is important because it suggests a direct relation between an important disease milestone and the outcome measure qMRI FF, which is required to use this outcome measure as primary endpoint in clinical trials.
Results could not be replicated in the cohort from the CCHMC (hazard ratio 0.96, 95% confidence interval 0.84-1.10, p=0.569). This may be explained by a limited number of LoA events (three) that occurred in this cohort of less severely affected participants. VL FF results were presented in growth charts, which could be used to stratify patients in clinical trials

with a small number of participants. Although results should be confirmed in a larger cohort with prospective determination of the disease milestone, our results support the use of FF assessed with qMRI as a primary endpoint or stratification tool in clinical trials in DMD.

In **chapter 4** the same approach was used as in chapter 3, but this time in a prospective study and for the relation between qMRI FF of an upper extremity muscle and loss of a milestone in non-ambulant patients with DMD (n=20). We assessed the additive predictive value of elbow flexor FF (48 MRIs) to age on loss of hand-to-mouth movement. Four-point Dixon MRI scans of the right upper arm were performed at baseline and at the 12-, 18-, or 24-month follow-up. Loss of hand-to-mouth movement was determined at study visits and by phone calls every 4 months. Elbow flexor FF predicted loss of hand-to-mouth movement on top of age in non-ambulant DMD patients (hazard ratio 1.12, 95% confidence interval 1.04-1.21, p=0.002). This result further established the relation between qMRI muscle FF and important disease milestones in DMD, thereby backing the clinical relevance of a potential effect of a therapy on qMRI muscle FF. It thus further supports use of qMRI muscle FF as primary endpoint in DMD and potentially facilitates the design of clinical trials via stratification based on disease severity and progression in qMRI FF.

In **chapter 5**, we presented qMRI results of the hand muscles from the longitudinal upper extremity outcome measure study in non-ambulant DMD patients. Fat replacement was minimal (9.7% versus 7.7%, p=0.043) and the T2 relaxation time of the muscle compartment ($T2_{water}$) was increased (31.5 ms versus 28.1 ms, p<0.001) compared to healthy controls. These results indicated that the thenar muscles were in an early stage of muscle pathology in our cohort of non-ambulant patients. Furthermore, the decrease in pinch strength (2.857 kg to 2.243 kg, p<0.001) and Performance of the Upper Limb (PUL) 2.0 total score (29 points to 23 points, p<0.001) over one year showed that there was measurable disease progression within the possible duration of a clinical trial. At follow-up, all participants still had useful function of the hands. Together with the moderate to strong correlation between muscle size and function, these results indicate that the thenar muscles are a valuable and quantifiable target for systemic or local therapy even in later stages of the disease. As a next step in outcome measure research, a direct relation between muscle qMRI or pinch strength and an important disease milestone still needs to be established in DMD.

In **chapter 6** we described results from the longitudinal upper extremity outcome measure study in non-ambulant DMD patients, concerning four innovative new outcome measures of upper extremity motor function using devices from the gaming industry. These outcome measures were developed in the form of a game and they provide a continuous outcome parameter without a maximum score, in order to overcome disadvantages of current outcome measures, such as a floor and ceiling effect, observer dependency, and motivational issues. Active range of motion (aROM) of the wrist and hand was determined using the Leap Motion sensor, and the Microsoft Kinect v2 sensor was used to determine three other

outcome measures. A stepwise approach was used to assess all technological outcome measures on quality control, construct validity, reliability, concurrent validity, longitudinal change and patient perception. The Ability Captured Through Interactive Video Evaluation (ACTIVE) game was used to determine the reached volume of the arms via the Kinect sensor, and showed the most promise in the stepwise approach. ACTIVE differed between patients and healthy controls ($p<0.001$), declined significantly over 12 months (5.6 points, $p=0.030$), and was appraised as being fun by patients. All four outcome measures were also correlated with and compared to results from PUL 2.0. There was a strong correlation of ACTIVE with PUL (rho=0.76). However, the standardized response mean (SRM) of ACTIVE was below 0.8, which is a commonly used threshold to determine responsiveness over time. PUL 2.0 performed similar to ACTIVE on the pervious items of the stepwise approach, but had an SRM above 0.8. Outcome measures based on hardware and software from the gaming industry can overcome problems such as observer dependence and lack of patient motivation. However, lack of insight in constraints of the software and hardware due to intellectual property, and possible software updates and hardware discontinuation, make these outcome measures a black box that could jeopardize their use in clinical trials.

General discussion

For the rare and fatal muscle wasting disease DMD, there is currently no fully approved therapy available beside glucocorticoids .[1] Improved care and glucocorticoids have led to a longer life-expectancy for patients, and therefore they go longer through life in the non-ambulant phase.[2] In this non-ambulant phase patients are progressively limited in their upper extremity functioning and therefore in their independence.[3] This leads to an urgent medical need for non-ambulant DMD patients. Many clinical trials with new drugs currently focus on the ambulant phase of the disease, although there are exceptions (NCT04371666, NCT01027884, NCT04004065). Attaining an effective and approved drug is a long and expensive process. Due to a progressive reduction in muscle tissue to be targeted by drugs, separate clinical trials need to be performed in non-ambulant patients, therefore specific outcome measures are required for this disease stage.[4, 5] These outcome measures should demonstrate sufficient reliability, construct validity, concurrent validity, longitudinal change, accessibility and clinical relevance. For this, natural history data of these outcome measures is required.

The 'best' upper extremity outcome measure

There is no ideal outcome measure. Which upper extremity outcome measure is optimal for a clinical trial differs depending on the targeted population, mechanism of action of the therapy, and the goal of the study.[6] Based on the results in this thesis, the following different types of outcome measures and the types of studies they would be appropriate for will be discussed: 1. clinical outcome measures that require specific patient related tasks, such as Performance of the Upper Limb, 2. patient reported outcome measures, and 3. biomarkers that for example reflect tissue characteristics, such as FF measured using muscle MR.

Clinical outcome measures

Clinical outcome measures often are more aligned with symptoms experienced by the patient compared to biochemical changes at tissue level, such as an increase in dystrophin. Therefore, providing clinical relevance is more straightforward.

Strength measurements are clinical outcome measures that have a relatively direct relation with a treatment effect, but are not inherently clinically meaningful. Pinch strength MyoPinch is an example of such an outcome measure, that is promising in an older and more progressed non-ambulant population in which the PUL 2.0 has reached its ceiling effect. The observed decline in pinch strength over 12 months in our non-ambulant cohort (chapter 5), was in agreement with the slightly older and weaker cohort described by Seferian et al.[7] The patients' view on strength tests was similar to PUL 2.0. After establishing clinical relevance of pinch strength via its ability to predict loss of a milestone, it could be used as outcome measure in the (late) non-ambulant phase.

The PUL 2.0 is an established outcome measure in non-ambulant patients. In chapters 4, 5 and 6, we showed an annual decline of 3 points in our cohort of patients who lost their

ambulation up to six years before study entry. This was similar to the decline in a recent study including 90 non-ambulant DMD patients, in which patients had an intermediate to fast decline (4-8 points) in the patients with entry level 2 to 6, and a small decline in patients with entry level 0-1 (1 point).[8] In our cohort, patients had an entry level of 2-5 points. The predicted sample size of 39 participants per study arm for PUL is smaller than for other previously studied outcome measures in non-ambulant DMD, which demonstrates the PUL's sensitivity to change.[9] The patient's view on the PUL was also demonstrated to be favorable (chapter 6). Despite disadvantages of observer-dependency and a floor and ceiling effect, we propose the PUL 2.0 as the preferred choice of outcome measure in studies that include per study arm 40-50 non-ambulant DMD patients with an PUL 2.0 entry level of 2-6 points.

In our data, the ACTIVE had a much larger predicted sample size per study arm (169 patients, chapter 6), compared to a study in spinal muscular atrophy (SMA, 28 patients).[10] This seems to demonstrate a higher sensitiveness to change of ACTIVE in these patients treated with nusinersen, although it can probably be attributed to the effectiveness of nusinersen.[10] The patient's view of ACTIVE was similar to PUL, except that ACTIVE was more tiring. An important limitation for using outcome measures, such as ACTIVE, are software updates which are out of control of the researchers. At the time of writing this thesis, a software update for ACTIVE had already taken place and the Microsoft Kinect v2 sensor had been taken out of production.[11] To overcome the constraints imposed by software updates, it is essential to share analysis algorithms with the community or even make them openly available, and/or to have a standardized validation process to ensure that software updates do not affect the outcome measure results. Although PUL was more sensitive to change in the non-ambulatory stage compared to ACTIVE, ACTIVE might be more sensitive to change in the (late) ambulatory stage, where patients have more retained shoulder function.

The PUL and ACTIVE are examples of outcome measures that are determined by observers and cameras/computers respectively. Observer dependent outcome measures, such as PUL and North Star Ambulatory Assessment, require training, are prone to inter- and intra-observer variability, and often have a ceiling effect.[8, 12] Their outcome also often has an ordinal scale, which hampers the use of linear statistics.[8, 12] Outcome measures determined by cameras/computers, such as ACTIVE and MyoPinch, require limited training and have continuous scales, but do require a sometimes expensive setup. They are also prone to software updates and sometimes have difficulty to register patient movements properly. Currently, patients also need to visit study sites to collect data for these outcome measures, while patient representatives stimulate the use of measurements that can also be performed from home, thereby removing the burden of traveling (chapter 2).[6] Future studies should aim to apply clinical outcome measures within the home environment.

Patient reported outcome measures

Outcome measures can also be patient reported questionnaires that can be performed from home. The DMD Upper Limb Patient Reported Outcome measure (PROM) did not show

significant change over 12 months in our small cohort of non-ambulant patients (*n*=13; chapter 5). A major advantage of the PROM is that data can be acquired off-site and that it can be filled in by parents for patients with neuropsychiatric comorbidity, such as autism spectrum disorder. Off-site completed forms are sensitive to missing data, but this can be overcome by prompt checking of received forms. The use of PROMs is advocated by patient representatives and the European Medicines Agency (EMA) and U.S. Food and Drug Administration (FDA), although currently it should still be supported by objective or functional outcomes.[4-6, 13] If PROM would demonstrate change over time in a larger cohort of non-ambulant patients, it could be the preferred outcome measure in large studies that strive for limited site visits and wish to be able to include non-ambulant DMD patients with neuropsychiatric comorbidity.

Biomarkers

The biomarker qMRI muscle FF reflects the increasing amount of fatty tissue within skeletal muscle with disease progression in DMD, and therefore is seen as a marker of loss of muscle tissue. It has been studied extensively in ambulant DMD.[14-30] qMRI FF data of the upper extremities in DMD patients is still sparse, because of the smaller muscle mass and position away from the center of the MRI, which decreases the image quality and frequently causes artefacts.[29] In non-ambulant DMD patients there is the extra challenge of contractures, burden of transport and travelling, difficulty to find a comfortable position in the scanner, and with increasing disease severity contra-indications of spinal fusion and daytime assistive ventilation become more prevalent (chapter 5).[6] Chapter 4 and 5 added to the available literature and suggested positioning on the side to overcome the issue of positioning the upper extremity muscles away from the center of the MRI.

Currently, qMRI FF is not yet approved by regulatory agencies as primary endpoint for clinical trials. Regulators state that they will consider all outcome measures, but that clinical relevance has to be demonstrated. An extensive and time-consuming reviewing process is required for formal approval, such as for the stride velocity 95th centile, which is now approved as secondary endpoint.[6, 31] To establish clinical relevance of qMRI muscle FF we studied and demonstrated its additive predictive value to age on loss of the disease milestones ambulation and hand-to-mouth movement (chapter 3 and 4). This predictive value to age was required in addition to previous correlations between FF and clinical outcome measures, because DMD is a progressive disease and any parameter that changes consistently over time, such as shoe-size, would correlate to a declining functional measure. By assessing the additive predictive value of qMRI muscle FF on top of the predictive value of age, we demonstrated the independent predictive value of qMRI muscle FF. The additive predictive value of qMRI FF of the lower extremities to LoA was also demonstrated by a recent study which showed a similar relation.[32]

The vastus lateralis is the preferred muscle to study in the ambulant phase, because of its sensitiveness to change in that phase and relation with ambulation.[18, 33] For the upper extremities, there is no consensus on which muscle (group) to use, but based on the

demonstrated clinical relevance of qMRI FF of the elbow flexor muscle group, we propose to use this muscle in the early non-ambulant phase. After approval by the regulatory agencies to use qMRI muscle FF as primary endpoint, its sensitiveness to disease progression could lead to shorter trials with potentially a lower burden for patients, although MRI remains less preferred by some patients. It would be most suitable for studies that seek to include a small number of patients from the age of 5 years old up to the early non-ambulant phase.

Determining clinical relevance of outcome measures used in clinical trials

As discussed previously, clinical relevance is essential for each outcome measure, before it can be used in clinical trials. An ideal way to demonstrate clinical relevance in a progressive disease is via the relation of an outcome measure to predict the moment of loss of an important disease milestone. Such a milestone needs to exist for every disease stage. Loss of the milestone has to be unequivocal and easy to determine, and preferably it can both be reported by patients and determined by an observer, which is the case for the ability to ambulate and bring a filled glass to the mouth. This requires extensive natural history studies with long-term follow-up, which leads to a high burden for participants and thus problems in participation (chapter 2). These large studies are especially necessary to determine the exact clinical relevance of an outcome measure, such as which percentage of slower increase in FF due to a therapy would lead to 2 years later LoA. A solution for these long-term studies with a high burden for patients is to integrate natural history studies in visits that take place as part of the outpatient clinical care (chapter 2). When important outcome measures are integrated in such visits, these can also be used for clinical trial phase IV, the post-marketing surveillance phase. In the Netherlands, for this reason the Duchenne Center Netherlands has set-up a training program for physiotherapists and occupational therapists from Academic Medical Centers and rehabilitation centers, to be able to gather high quality data as part of the regular outpatient clinical care. Data can be collected prospectively in the national patient registry, the Dutch Dystrophinopathy Database, and in the national biobank for DMD and BMD.

Even for the seemingly straightforward milestone LoA, different definitions have been used in literature, such as the inability to perform the 10 meter walk/run test at hospital visits or a score ≤25% on the D1 subscale about standing position and transfers of the Motor Function Measure.[20, 28, 34, 35] Similar to another study[34], we defined LoA as the patient being unable to walk 5 meter without assistance or orthoses, and determined this by conducting a detailed interview at each hospital visit (chapter 3). Because it has a large impact on daily life when patients become wheelchair bound, in our view it was feasible to determine LoA to a month's precision. Nowadays, captured photo's/video's on smartphones can also aid in defining this disease milestone. To increase precision without increasing the burden for patients and continue milestone determination when restrictions were in place due to the COVID-19 pandemic, a phone call each quarter could be performed. We propose for all

stakeholders to use similar methodology to determine LoA, because it allows for off-site determination of this milestone.

For the upper extremity, different definitions of an hand-to-mouth function have been included in the Brooke upper extremity scale, PUL and PROM, and have also been used as milestone.[3, 36-38] Our definition was moving a filled glass (total weight 200gram) independently to the mouth using the right hand and allowing support of the elbow on a table (chapter 4). We chose the right hand for a direct relation with the elbow flexor muscles of the right upper arm on qMRI, but for other purposes it can be advisable to use this milestone for either the dominant or both hands. [3, 36, 37] More than 40% of hand-to-mouth movement can be attributed to the biceps brachii and brachialis muscles.[39] Moving a small weight to the mouth is part of vital daily life activities, such as drinking, eating and performing personal hygiene such as brushing teeth unaided. Patients and families assisted in the development of the PUL and PROM, and incorporation of the hand-to-mouth movement in these outcome measures confirmed its clinical relevance for patients and parents.[37, 38] The proposed method of determination of the milestone is again via a detailed interview at each hospital visit and a phone call each quarter.

Previous studies used 'no useful function of the hands' as upper extremity disease milestone of hand function in the late non-ambulatory stage.[3, 40] In our longitudinal upper extremity outcome measure study in DMD, no patients have reached this level of inability and a previous study showed that the time interval between loss of hand-to-mouth function (median 15 years) and loss of useful function of the hands (median 23 years) is large.[3] Compared to ambulant patients, non-ambulant patients spend significantly more time on playing (online) video games, which in recent years has become an important tool for social interaction.[41] We proposed a disease milestone of hand function that should fall within the described time interval and is more applied to daily life, i.e. the ability to play a video game for 10 minutes using a game controller (chapter 5). Developers of game controllers focus increasingly on accessibility to all, which has led to new devices such as the Xbox Adaptive Controller.[42, 43] Amongst our study participants, many patients switched over the course of the study from a traditional game controller to smaller game controllers, such as the Nintendo Switch.[42] A longer follow-up duration in our study should clarify whether playing video games for 10 minutes using a game controller is a useful addition to the disease milestone toolbox in DMD.

Future perspectives

The ideal outcome measure does not exist, but there is an optimal outcome measure for every clinical trial in non-ambulant DMD patients based on the mechanism of action of the therapy, disease phase and the study design.[6] Currently, the PUL 2.0 seems to be the preferred primary outcome measure for skeletal muscle function in a clinical trial in larger non-ambulant population with a PUL entry level of 2-6.

Several gaps in research are still present and need to be addressed. Demonstrating clinical relevance is important for all outcome measures and is supported by regulators. The PUL 2.0 has already been used as primary endpoint in clinical trials, but it has not yet been established what a decline of one or two points on this scale means functionally for the patient or which difference would be clinically relevant. This could be studied by assessing the ability of PUL 2.0 to predict loss of an important disease milestone. This is also the case for the MyoPinch, PROM, and ACTIVE. We proposed hand to mouth function and gaming as disease milestones for the upper extremity. Although the PROM and PUL and the items they contain have been developed with feedback from patient representatives, no extensive studies on patient preferences in disease milestones have been performed.[37, 38] This knowledge on disease milestones would aid in developing studies to assess clinical relevance of a detected change and of outcome measures as a whole, such as the PUL 2.0, MyoPinch, PROM, and ACTIVE. For PROM, also sensitivity to change over time in a larger cohort of non-ambulant DMD patients is required before it can be considered as primary endpoint in clinical trials. Furthermore, the sensitivity to change of ACTIVE is currently too low for it to be considered as primary endpoint in clinical trials. A study in a less severely affected DMD patient population with more retained shoulder function would provide insight whether ACTIVE is more sensitive to change in that population. For qMRI FF, clinical relevance in relation to important disease milestones in general has been established. It is, however, important to have a quantitative estimation on the percentage change in FF that is needed for a single year delay in reaching a disease milestone. qMRI muscle FF could also be used as primary endpoint to study the effect of a local therapy and compare differences between a treated and untreated arm in small trials. As a first step, the natural history of qMRI muscle FF in the left and right arm has to be compared.

Some of the described gaps in research can be addressed by integrating important outcome measures, which are assessed partly for research, into visits as part of the outpatient clinical care. In the Netherlands, a biobank has been set-up by the university medical centers to capture clinical care data without requiring extra effort from patients, so that additional studies that are integrated in care do not have to gather these data separately.

In this thesis, we found that the PUL 2.0 seems to be the preferred outcome measure for skeletal muscle function in clinical trials for non-ambulant DMD patients at this time. The assistance of patients and parents in the development of PUL 2.0 supports its clinical relevance, but its ability to predict loss of a milestone would further aid in establishing the clinical relevance of PUL 2.0. qMRI muscle FF was shown to detect clinically relevant change, but poses practical disadvantages in non-ambulant patients and is less preferred by patients and therefore not ideal for large patient groups. MyoPinch, PROM and ACTIVE were also identified as promising outcome measures, but gaps in knowledge need to be addressed first. Use of these outcome measures could lead to smaller sample sizes and/or a shorter duration of trials with a lower burden for patients. Finally, even in advanced stages of the

disease, clinically relevant muscles are relatively preserved and therefore warrant our effort to search for effective treatments for these patients as well.

References

1. Birnkrant DJ, Bushby K, Bann CM, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and neuromuscular, rehabilitation, endocrine, and gastrointestinal and nutritional management. Lancet Neurol 2018;17:251-267. doi:10.1016/S1474-4422(18)30024-3
2. Gloss D, Moxley RT, Ashwal S, Oskoui M. Practice guideline update summary: Corticosteroid treatment of Duchenne muscular dystrophy Report of the Guideline Development Subcommittee of the American Academy of Neurology. Neurology 2016;86:465-472. doi:10.1212/Wnl.0000000000002337
3. McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. Lancet 2018;391:451-461. doi:10.1016/S0140-6736(17)32160-8
4. CDER, CBER. Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment. Guidance for Industry. https://www.fda.gov/media/92233/download. Accessed on May 1, 2019. US Food & Drug Administration (FDA) 2018
5. CHMP. Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. In European Medicines Agency (EMA). Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf. Accessed December 18, 2020.
6. Straub V, Mercuri E, Grp DOMS. Report on the workshop: Meaningful outcome measures for Duchenne muscular dystrophy, London, UK, 30-31 January 2017. Neuromuscular Disorders 2018;28:690-701. doi:10.1016/j.nmd.2018.05.013
7. Seferian AM, Moraux A, Annoussamy M, et al. Upper limb strength and function changes during a one-year follow-up in non-ambulant patients with Duchenne Muscular Dystrophy: an observational multicenter trial. PLoS One 2015;10:e0113999. doi:10.1371/journal.pone.0113999
8. Pane M, Coratti G, Brogna C, et al. Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal data. PLoS One 2018;13:e0199223. doi:10.1371/journal.pone.0199223
9. Connolly AM, Florence JM, Zaidman CM, et al. Clinical trial readiness in non-ambulatory boys and men with duchenne muscular dystrophy: MDA-DMD network follow-up. Muscle Nerve 2016;54:681-689. doi:10.1002/mus.25089
10. Alfano LN, Miller NF, Iammarino MA, et al. ACTIVE (Ability Captured Through Interactive Video Evaluation) workspace volume video game to quantify meaningful change in spinal muscular atrophy. Dev Med Child Neurol 2020;62:303-309. doi:10.1111/dmcn.14230
11. Microsoft. Microsoft to consolidate the Kinect for Windows experience around a single sensor. Accessed on May 5, 2020 at: https://docs.microsoft.com/en-us/archive/blogs/kinectforwindows/microsoft-to-consolidate-the-kinect-for-windows-experience-around-a-single-sensor [online].
12. Pane M, Mazzone ES, Fanelli L, et al. Reliability of the Performance of Upper Limb assessment in Duchenne muscular dystrophy. Neuromuscul Disord 2014;24:201-206. doi:10.1016/j.nmd.2013.11.014
13. CHMP. Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies. In European Medicines Agency (EMA). Available at: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf. Accessed December 31, 2021.
14. Godi C, Ambrosi A, Nicastro F, et al. Longitudinal MRI quantification of muscle degeneration in Duchenne muscular dystrophy. Ann Clin Transl Neurol 2016;3:607-622. doi:10.1002/acn3.319
15. Bonati U, Hafner P, Schadelin S, et al. Quantitative muscle MRI: A powerful surrogate outcome measure in Duchenne muscular dystrophy. Neuromuscul Disord 2015;25:679-685. doi:10.1016/j.nmd.2015.05.006
16. Akima H, Lott D, Senesac C, et al. Relationships of thigh muscle contractile and non-contractile tissue with function, strength, and age in boys with Duchenne muscular dystrophy. Neuromuscul Disord 2012;22:16-25. doi:10.1016/j.nmd.2011.06.750
17. Wokke BH, van den Bergen JC, Versluis MJ, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. Neuromuscul Disord 2014;24:409-416. doi:10.1016/j.nmd.2014.01.015
18. Willcocks RJ, Rooney WD, Triplett WT, et al. Multicenter prospective longitudinal study of magnetic resonance biomarkers in a large duchenne muscular dystrophy cohort. Ann Neurol 2016;79:535-547. doi:10.1002/ana.24599
19. Burakiewicz J, Sinclair CDJ, Fischer D, et al. Quantifying fat replacement of muscle by quantitative MRI in muscular dystrophy. Journal of neurology 2017. doi:10.1007/s00415-017-8547-3
20. Fischmann A, Hafner P, Gloor M, et al. Quantitative MRI and loss of free ambulation in Duchenne muscular dystrophy. Journal of neurology 2013;260:969-974. doi:10.1007/s00415-012-6733-x

21. Forbes SC, Walter GA, Rooney WD, et al. Skeletal muscles of ambulant children with Duchenne muscular dystrophy: validation of multicenter study of evaluation with MR imaging and MR spectroscopy. Radiology 2013;269:198-207. doi:10.1148/radiol.13121948
22. Mercuri E, Pichiecchio A, Counsell S, et al. A short protocol for muscle MRI in children with muscular dystrophies. Eur J Paediatr Neurol 2002;6:305-307
23. Hogrel JY, Wary C, Moraux A, et al. Longitudinal functional and NMR assessment of upper limbs in Duchenne muscular dystrophy. Neurology 2016;86:1022-1030. doi:10.1212/WNL.0000000000002464
24. Wary C, Azzabou N, Giraudeau C, et al. Quantitative NMRI and NMRS identify augmented disease progression after loss of ambulation in forearms of boys with Duchenne muscular dystrophy. NMR Biomed 2015;28:1150-1162. doi:10.1002/nbm.3352
25. Willcocks RJ, Triplett WT, Forbes SC, et al. Magnetic resonance imaging of the proximal upper extremity musculature in boys with Duchenne muscular dystrophy. Journal of neurology 2017;264:64-71. doi:10.1007/s00415-016-8311-0
26. Hooijmans MT, Niks EH, Burakiewicz J, et al. Non-uniform muscle fat replacement along the proximodistal axis in Duchenne muscular dystrophy. Neuromuscul Disord 2017;27:458-464. doi:10.1016/j.nmd.2017.02.009
27. Barnard AM, Willcocks RJ, Finanger EL, et al. Skeletal muscle magnetic resonance biomarkers correlate with function and sentinel events in Duchenne muscular dystrophy. PLoS One 2018;13:e0194283. doi:10.1371/journal.pone.0194283
28. Barnard AM, Willcocks RJ, Triplett WT, et al. MR biomarkers predict clinical function in Duchenne muscular dystrophy. Neurology 2020;94:e897-e909. doi:10.1212/WNL.0000000000009012
29. Forbes SC, Arora H, Willcocks RJ, et al. Upper and Lower Extremities in Duchenne Muscular Dystrophy Evaluated with Quantitative MRI and Proton MR Spectroscopy in a Multicenter Cohort. Radiology 2020;295:616-625. doi:10.1148/radiol.2020192210
30. Wood CL, Hollingsworth KG, Hughes E, et al. Pubertal induction in adolescents with DMD is associated with high satisfaction, gonadotropin release and increased muscle contractile surface area. Eur J Endocrinol 2020. doi:10.1530/EJE-20-0709
31. CHMP. Qualification opinion on stride velocity 95th centile as a secondary endpoint in Duchenne Muscular Dystrophy measured by a valid and suitable wearable device*. In European Medicines Agency (EMA). Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-stride-velocity-95th-centile-secondary-endpoint-duchenne-muscular-dystrophy_en.pdf. Accessed June 16, 2021.
32. Rooney WD, Berlow YA, Triplett WT, et al. Modeling disease trajectory in Duchenne muscular dystrophy. Neurology 2020. doi:10.1212/WNL.0000000000009244
33. Rooney WD, Berlow YA, Triplett WT, et al. Modeling disease trajectory in Duchenne muscular dystrophy. Neurology 2020;94:e1622-e1633. doi:10.1212/WNL.0000000000009244
34. Mazzone ES, Pane M, Sormani MP, et al. 24 month longitudinal data in ambulant boys with Duchenne muscular dystrophy. PLoS One 2013;8:e52512. doi:10.1371/journal.pone.0052512
35. Vuillerot C, Girardot F, Payan C, et al. Monitoring changes and predicting loss of ambulation in Duchenne muscular dystrophy with the Motor Function Measure. Dev Med Child Neurol 2010;52:60-65. doi:10.1111/j.1469-8749.2009.03316.x
36. Brooke MH, Griggs RC, Mendell JR, et al. Clinical trial in Duchenne dystrophy. I. The design of the protocol. Muscle Nerve 1981;4:186-197. doi:10.1002/mus.880040304
37. Klingels K, Mayhew AG, Mazzone ES, et al. Development of a patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy: DMD Upper Limb PROM. Dev Med Child Neurol 2017;59:224-231. doi:10.1111/dmcn.13277
38. Mayhew A, Mazzone ES, Eagle M, et al. Development of the Performance of the Upper Limb module for Duchenne muscular dystrophy. Dev Med Child Neurol 2013;55:1038-1045. doi:10.1111/dmcn.12213
39. Kawakami Y, Nakazawa K, Fujimoto T, et al. Specific tension of elbow flexor and extensor muscles based on magnetic resonance imaging. Eur J Appl Physiol Occup Physiol 1994;68:139-147
40. Brooke MH, Fenichel GM, Griggs RC, et al. Duchenne muscular dystrophy: patterns of clinical progression and effects of supportive therapy. Neurology 1989;39:475-481. doi:10.1212/wnl.39.4.475
41. Heutinck L, Kampen NV, Jansen M, Groot IJ. Physical Activity in Boys With Duchenne Muscular Dystrophy Is Lower and Less Demanding Compared to Healthy Boys. J Child Neurol 2017;32:450-457. doi:10.1177/0883073816685506
42. Stoner G. For physically disabled gamers, the Switch is incredibly accessible. Here's why. Available at: https://www.washingtonpost.com/video-games/2020/04/21/accessibility-gaming-nintendo-switch/. Accessed June 18, 2021. The Washington Post April 21 2020.
43. Gaming that is accessible for all. Available at: https://www.xbox.com/en-US/community/for-everyone/accessibility. Accessed June 18, 2021. [online].

Appendices

NEDERLANDSE SAMENVATTING

Het overkoepelende doel van dit proefschrift was om uitkomstmaten te identificeren voor Duchenne spierdystrofie (DMD), die met name in rolstoelgebonden patiënten een klinische relevant verschil kunnen vaststellen in de korte periode die binnen medicijnonderzoeken beschikbaar is. Het gebruik van dergelijke uitkomstmaten zou kunnen leiden tot kleinere patiëntengroepen in medicijnonderzoeken en een lagere belasting voor patiënten.

In **hoofdstuk 2** hebben we onderzocht welke redenen patiënten en/of hun verzorgers opgaven om niet mee te doen aan drie observationele studies voor patiënten met DMD en één studie voor patiënten met Becker spierdystrofie (BMD). Eerst hebben we onderzocht of leeftijd, reistijd, *DMD* genmutatie en leeftijd van verlies van loopfunctie overeenkomstig waren tussen mensen die wel of niet meededen aan de studies. We verkregen deze informatie uit de landelijke patiëntendatabase, de Dutch Dystrophinopathy Database. Het bleek dat deelnemers over het algemeen representatief waren voor de patiëntgroepen die aan elk onderzoek mee konden doen. Wel deden in geen van de onderzoeken patiënten mee met mutaties die gerelateerd zijn aan meer leer- en gedragsproblemen (exonen 63 t/m 79), en waren deelnemers jonger dan patiënten die niet meededen in de studie naar uitkomstmaten van de armen in rolstoelgebonden DMD patiënten (hoofdstuk 4, 5 en 6). Dit suggereert dat het onderzoek met patiënten in latere ziektestadia van DMD uitdagender kan zijn wat betreft de inclusie. De vaakst gerapporteerde redenen om niet mee te doen met onderzoek waren de 'Belasting van het studieprotocol' (38%), 'MRI' (30%), en 'reistijd' (19%).
Onze resultaten benadrukken dat landelijke patiëntendatabases kunnen worden gebruikt om deelnemers aan onderzoek te vergelijken met patiënten die niet meedoen, waardoor vastgesteld kan worden dat observationeel onderzoek representatief is voor de gehele patiëntenpopulatie. Daarnaast suggereren de resultaten dat de volgende factoren deelname van patiënten aan onderzoek zou kunnen faciliteren en doen toenemen: 1) betrokkenheid van patiënten in het ontwerp van nieuwe studies optimaliseren, 2) de MRI ervaring verbeteren, 3) observationeel onderzoek integreren in de patiëntenzorg.

In **hoofdstuk 3** hebben we onderzocht of het vetpercentage (FF) in de vastus lateralis (VL) spier gemeten met kwantitatieve MRI (qMRI) een aanvullend voorspellende waarde heeft bovenop de leeftijd op verlies van de loopfunctie. Dit werd onderzocht in twee cohorten, één van het Leids Universitair Medisch Centrum (LUMC; n=19) en de andere van Cincinnati Children's Hospital Medical Center (CCHMC; n=15). Er was een uitstekende interbeoordelaars-betrouwbaarheid voor het meten van de VL FF met qMRI. Dit ondersteunt de haalbaarheid van het gebruik van spier qMRI data in multicenter onderzoeken. VL FF bleek aanvullend voorspellende waarde te hebben op de leeftijd van verlies van loopfunctie in het LUMC cohort (hazard ratio 1.15, 95% betrouwbaarheidsinterval 1.05-1.26, p=0.003). Dit is van belang, omdat het een directe relatie tussen een belangrijke ziektemijlpaal en de

uitkomstmaat qMRI FF suggereert, wat noodzakelijk is om deze uitkomstmaat als primair eindpunt in medicijnonderzoeken te kunnen gebruiken.
Hetzelfde resultaat werd niet gevonden in het CCHMC cohort (hazard ratio 0.96, 95% betrouwbaarheidsinterval 0.84-1.10, *p*=0.569). Dat zou verklaard kunnen worden door het kleine aantal patiënten dat de loopfunctie had verloren (drie) in dit cohort van minder ernstig aangedane patiënten. We toonden VL FF resultaten in groeicurves, welke gebruikt zouden kunnen worden om patiënten te stratificeren in medicijnonderzoeken met een klein aantal deelnemers. Hoewel de resultaten in een groter cohort met prospectieve vaststelling van de ziektemijlpaal zouden moeten worden bevestigd, ondersteunen onze resultaten het gebruik van met qMRI gemeten spier FF als primair eindpunt of methode voor stratificatie binnen medicijnonderzoek voor DMD.

In **hoofdstuk 4** werd dezelfde methode als in hoofdstuk 3 gebruikt, maar dan in een prospectieve studie en voor de relatie tussen qMRI FF van een armspier en het bereiken van een ziektemijlpaal bij rolstoelgebonden patiënten met DMD (n=20). We onderzochten de aanvullend voorspellende waarde van elleboogflexie FF (48 MRI's) op de leeftijd van verlies van de hand-naar-de-mond beweging. Vier-punten Dixon MRI scans van de rechter bovenarm werden verricht op het eerste meetpunt en bij controle na 12, 18 of 24 maanden. Verlies van de hand-naar-de-mond beweging werd vastgesteld bij studie bezoeken en door middel van telefonische controle elke 4 maanden. Elleboogflexie FF bleek aanvullend voorspellende waarde te hebben op de leeftijd van verlies van de hand-naar-de-mond beweging in rolstoelgebonden DMD patiënten (hazard ratio 1.12, 95% betrouwbaarheidsinterval 1.04-1.21, *p*=0.002). Dit resultaat bevestigde de relatie tussen qMRI spier FF en belangrijke ziekte mijlpalen in DMD, en ondersteunt daarmee de klinische relevantie van een eventueel effect van een behandeling op qMRI spier FF. Daarmee wordt het gebruik van qMRI spier FF als primair eindpunt in DMD ondersteund. Ook kan qMRI spier FF de ziekte-ernst en progressie helpen bepalen en kan dit gebruikt worden voor stratificatie en zo mogelijk het ontwerp van nieuwe studies vergemakkelijken.

In **hoofdstuk 5** hebben we qMRI resultaten beschreven van de handspieren van rolstoelgebonden DMD patiënten die meededen aan de longitudinale studie naar uitkomstmaten van de armen. Er was minimale vervetting (9.7% versus 7.7%, *p*=0.043) en de T2 relaxatietijd in spierweefsel ($T2_{water}$) was toegenomen (31.5 ms versus 28.1 ms, *p*<0.001) vergeleken met de gezonde controle. Deze resultaten wezen erop dat de ernst van spierpathologie in de duimmuis spieren zich in een vroeg stadium bevond in ons cohort van rolstoelgebonden patiënten. Daarnaast toonden de afname in duimknijpkracht (2.857 kg naar 2.243 kg, *p*<0.001) en in Performance of the Upper Limb (PUL) 2.0 totale score (29 punten naar 23 punten, *p*<0.001) over de periode van een jaar dat er meetbare ziekteprogressie is binnen de mogelijke tijdsduur van een medicijnonderzoek. Bij de controleafspraken konden alle patiënten de handen nog functioneel inzetten. Samen met de matige tot sterke correlatie tussen spieromvang en functie, wijzen deze resultaten erop

dat de duimmuis spieren een waardevol en meetbaar behandeldoel zijn voor systemische of lokale therapie, zelfs bij een gevorderd ziektestadium. Als vervolgstap in het uitkomstmaten onderzoek in DMD zou vastgesteld moeten worden of er een directe relatie bestaat tussen qMRI van de duimmuisspier of duimknijpkracht en een belangrijke ziektemijlpaal.

In **hoofdstuk 6** beschreven we resultaten van vier innovatieve uitkomstmaten van armfunctie welke werden gemeten met apparaten uit de gaming industrie. Deze data werd verzameld binnen de longitudinale studie naar uitkomstmaten van de armen in rolstoelgebonden DMD patiënten. De vier uitkomstmaten werden ontwikkeld in de vorm van een spel en leverden een continue uitkomstparameter op zonder maximumscore, wat tot doel had nadelen van huidige uitkomstmaten te vermijden, zoals een bodem- en plafond-effect, afhankelijkheid van een observator, en problemen met motivatie. Voor het vaststellen van de actieve range of motion (aROM) van de pols en hand werd een Leap Motion sensor gebruikt, en voor de andere drie uitkomstmaten werd een Microsoft Kinect v2 sensor gebruikt. Een stapsgewijze analyse werd gebruikt om de vier technologische uitkomstmaten te beoordelen op gebied van kwaliteitscontrole, construct validiteit, betrouwbaarheid, concurrente validiteit, verandering over de tijd en patiënten ervaring. Het Ability Captured Through Interactive Video Evaluation (ACTIVE) spel werd gebruikt om het bereikte volume van de armen te meten met de Kinect sensor, en dit vertoonde de meeste potentie in de stapsgewijze analyse. ACTIVE verschilde tussen patiënten en gezonde controles (p<0.001), nam significant af over 12 maanden (5.6 punten, p=0.030), en werd door patiënten als leuk ervaren. Alle vier uitkomstmaten werden ook gecorreleerd met en vergeleken met resultaten van de PUL 2.0. Er was een sterke correlatie tussen ACTIVE en PUL (rho=0.76). Echter was de gestandaardiseerde gemiddelde reactie (SRM) van ACTIVE onder de 0.8. De SRM is een veel gebruikte drempel, waarbij een waarde boven 0.8 wijst op een goede reactie over de tijd. PUL 2.0 gaf vergelijkbare resultaten als ACTIVE op de meeste onderdelen van de stapsgewijze analyse, maar had een SRM boven de 0.8. Uitkomstmaten die gebaseerd zijn op hardware of software van de gaming industrie kunnen nadelen van huidige uitkomstmaten vermijden, zoals afhankelijkheid van een observator en problemen met motivatie. Echter zorgt gebrek aan inzicht in de beperkingen van software en hardware door intellectueel eigendom, en mogelijke software updates en productiestops voor hardware, ervoor dat deze uitkomstmaten een black box zijn wat hun gebruik in medicijnonderzoek in gevaar kan brengen.

List of publications

Journal publications

Geraedts VJ, van Hilten JJ, Marinus J, et al. Stimulation challenge test after STN DBS improves satisfaction in Parkinson's disease patients. Parkinsonism & related disorders 2019;69:30-33. doi:10.1016/j.parkreldis.2019.10.014

Naarding KJ, Reyngoudt H, van Zwet EW, et al. MRI vastus lateralis fat fraction predicts loss of ambulation in Duchenne muscular dystrophy. Neurology 2020;94:e1386-e1394. doi:10.1212/WNL.0000000000008939

Keene KR, Beenakker JM, Hooijmans MT, et al. T2 relaxation-time mapping in healthy and diseased skeletal muscle using extended phase graph algorithms. Magn Reson Med 2020;84:2656-2670. doi:10.1002/mrm.28290

Naarding KJ, Doorenweerd N, Koeks Z, et al. Decision-Making And Selection Bias in Four Observational Studies on Duchenne and Becker Muscular Dystrophy. J Neuromuscul Dis 2020;7:433-442. doi:10.3233/JND-200541

Naarding KJ, Keene KR, Mishre ASDS, et al. Preserved thenar muscles in non-ambulant Duchenne muscular dystrophy patients. J Cachexia Sarcopeni 2021;12:694-703. doi:10.1002/jcsm.12711

Naarding KJ, van der Holst M, van Zwet EW, et al. Association of Elbow Flexor MRI Fat Fraction With Loss of Hand-to-Mouth Movement in Patients With Duchenne Muscular Dystrophy. Neurology 2021;97:E1737-E1742. doi:10.1212/Wnl.0000000000012724

Veeger TTJ, Zwet EW, Mohamad D, et al. Muscle architecture is associated with muscle fat replacement in Duchenne and Becker muscular dystrophies. Muscle & Nerve 2021;64:576-584. doi:10.1002/mus.27399

Naarding KJ, Janssen MMHP, Boon RD, et al. The Black Box of Technological Outcome Measures: An Example in Duchenne Muscular Dystrophy. Journal of Neuromuscular Diseases 2022;9:555-569. doi:10.3233/Jnd-210767

Published conference abstracts

Naarding K, Reyngoudt H, van Zwet E, et al. Higher MRI muscle fat fraction at similar age is associated with earlier loss of ambulation in Duchenne muscular dystrophy. Neuromuscular Disorders 2018;28:S41-S41. doi:10.1016/j.nmd.2018.06.063

Naarding K, Van der Holst M, Van de Velde N, et al. Patient perception of outcome measures for non-ambulant Duchenne muscular dystrophy patients. Neuromuscular Disorders 2019;29:S180-S180. doi:10.1016/j.nrnd.2019.06.501

Naarding K, Veeger T, Mishre AS, et al. MRI brachialis contractile cross-sectional area is correlated strongest to elbow flexion in non-ambulant Duchenne muscular dystrophy patients. Neuromuscular Disorders 2019;29:S156-S156. doi:10.1016/j.nmd.2019.06.419

Veeger T, van Zwet E, al Mohamad D, et al. Association between the progression of muscle fat replacement and muscle architecture in Duchenne muscular dystrophy. Neuromuscular Disorders 2020;30:S96-S96. doi:10.1016/j.nmd.2020.08.172

Curriculum Vitae

Karin Naarding werd geboren in 's-Gravenzande op 7 oktober 1990, als dochter van Pim en Erna Naarding. Vanwege een kennis met Duchenne spierdystrofie schreef ze op het VWO een verslag over de eventuele mogelijkheden van gentherapie bij deze ziekte. Ze behaalde haar VWO diploma aan het Vechtdal College te Hardenberg in 2008, en startte datzelfde jaar met haar studie geneeskunde aan de Rijksuniversiteit Groningen.

Nadat ze de bachelor cum laude had afgesloten, verbreedde ze haar kennis middels een zelfgeorganiseerde minor op gebied van management, economie en rechten, en een master honours programma op het gebied van leiderschap. Haar wetenschappelijke stage van de master geneeskunde vond plaats bij de afdeling neurologie van het Amsterdam UMC, locatie Academisch Medisch Centrum (AMC). Deze stage ging over de waarde van een scoring in vier condities bij patiënten met de ziekte van Parkinson, die behandeld werden met diepe hersenstimulatie.

Karin studeerde in 2015 af als arts en besloot eerst klinische ervaring op te doen. Vanwege een interesse in de neurologie ging ze als arts-assistent niet in opleiding tot neuroloog werken in het Onze Lieve Vrouwe Gasthuis locatie oost, te Amsterdam.

Vanwege een blijvende interesse in Duchenne spierdystrofie en het effect van deze ziekte op het leven van patiënten, wilde ze hier graag onderzoek naar doen. In 2017 startte ze als promovenda in het Leids Universitair Medisch Centrum. Het werk op het gebied van uitkomstmaten bij Duchenne spierdystrofie heeft geresulteerd in dit proefschrift.

In 2021 is ze getrouwd met haar partner, Jurjen IJska. In de hoop een positieve impact te kunnen hebben op het dagelijks leven van patiënten wilde zij zich verder specialiseren in de revalidatiegeneeskunde. Ze is sinds maart 2021 in opleiding tot revalidatiearts in de onderwijs- en opleidingsregio van het Amsterdam UMC, locatie AMC.

Dankwoord

Tot een afgerond proefschrift kom je niet alleen, en hierbij wil ik graag iedereen heel erg bedanken die hier op welke manier dan ook bij heeft geholpen!

Allereerst wil ik alle deelnemers en hun familie bedanken. Meedoen aan onderzoek kost moeite en tijd, en jullie zijn allemaal bereid geweest om dit te investeren. Ik neem mooie en leerzame herinneringen mee van de studiebezoeken.

Veel dank aan mijn promotor Jan en copromotoren Erik en Hermien, die me de vrijheid hebben gegeven om er mijn eigen onderzoek van te maken. Jan, bedankt voor de verfrissende inzichten en enthousiaste nieuwe ideeën die je steeds weer naar voren bracht. Erik, dankjewel voor het gestelde vertrouwen, de begeleiding, en je kritische en verhelderende kijk op alles. Hermien, wat fijn dat je altijd zo benaderbaar bent geweest en me eindeloos van inhoudelijke input en schrijftips wilde voorzien.

Dank aan alle collega's voor de steunende, leuke en inspirerende gesprekken tijdens en na werktijd, in het bijzonder aan mijn kamergenootjes van het spierballencentrum (J3-166) en Gert's kast (J3-164). Mink en Maryam, jullie zijn vanaf dag één mijn maatjes in dezelfde fase van onze promotie geweest, wat een lol en struggles hebben we samen gehad. Nienke, altijd tijd voor gezelligheid en een luisterend oor over en weer, we hebben heel wat afgeskyped. Zaïda en Nadine, wat hebben jullie toch een blije energie, sinds Nadine's aanwezigheid zijn de katten niet meer weg te denken.

Nathalie, wat was het fijn samenwerken met jou, zowel met artikelen als onderwijs geven. Kevin, altijd tijd voor een melig moment, Subway of programmeerprobleem, maar dan moet het wel echt ingewikkeld zijn. Thom en Aashley, dank voor alle programmeerhulp en gezelligheid. Jordi, jij was mijn redder in nood met de MRI en zorgde altijd voor een blij begin van mijn scandag. Menno, dank voor alles wat je me geleerd hebt, je fantastische humeur en omgang met de patiënten. Mariska en Linda, dank voor de fijne samenwerking, door jullie kwamen de bewegingsanalyses tot leven.

De hele neuromusculaire groep, dank voor de fantastische congressen. Chiara, Martijn en Umesh, wat fijn dat jullie zo benaderbaar waren. Yvonne, Marjolein, Anne-Marie en Martha, bedankt voor het fijne samenwerken en alle praktische oplossingen.

Nu terug naar hoe het zaadje voor dit proefschrift is geplant. Op mijn tiende verhuisden we naar Dedemsvaart, waar we kennismaakten met onze nieuwe buren: Jan, Hennie en hun zoon Erik. Erik Vrielink was een verwoed vogelspotter en heeft met kinbesturing een eigen website gebouwd en bijgehouden. Een afdruk van zijn hand en voet van de kinderleeftijd sieren de omslag van dit proefschrift en verwijzen naar het belang van het kunnen bewegen

van je armen en benen. Erik had de ziekte van Duchenne. Ik wil Jan en postuum Hennie en Erik bedanken voor de inspiratie en de hulp met de omslag van dit proefschrift, waardoor het verhaal nu rond is.

Als laatste dank aan mijn lieve familie en schoonfamilie. Mijn betrokken ouders wiens trots en steun ik altijd gevoeld heb. En Jurjen, mijn onverzettelijke rots in de branding.