



RESEARCH ARTICLE

Responsiveness of outcome measures in myotonic dystrophy type 1

Kirsten L. Knak , Aisha M. Sheikh, Nanna Witting & John Vissing 

Rigshospitalet, Department of Neurology, Juliane Maries Vej 28, Copenhagen, DK-2100, Denmark

Correspondence

Kirsten L. Knak, Rigshospitalet, Department of Neurology, Copenhagen, Denmark. Tel: +45 35456217; Fax: 35 45 61 38; E-mail: kirsten.lykke.knak@regionh.dk

Received: 21 April 2020; Revised: 7 June 2020; Accepted: 18 June 2020

Annals of Clinical and Translational Neurology 2020; 7(8): 1382–1391

doi: 10.1002/acn3.51129

Abstract

Objective: As myotonic dystrophy type 1 (DM1) evolves slowly and interventional trials often have a short duration, responsive outcomes in DM1 are needed. The objective of this study was to determine the responsiveness of muscle strength, balance, and functional mobility measurements after a 1-year follow-up period in individuals with DM1. **Methods:** Sixty-three adults with noncongenital DM1 completed the following assessments at baseline and at 1-year follow-up: Handheld dynamometry (lower limbs), stationary dynamometry (lower limbs), step test, timed-up-and-go test (TUG), modified clinical test of sensory integration and balance (mCTSIB), feet-together stance, tandem stance, one-leg stance, 10-meter walk test, and sit-to-stand test. **Results:** Change was captured by stationary dynamometry (proximal flexor and extensor muscles), handheld dynamometry (proximal flexor and distal extensor muscles), TUG, and mCTSIB ($P \leq 0.04$). Ceiling or floor effects were shown for most static balance tests. **Interpretation:** Overall, adequate responsiveness was shown for both muscle strength dynamometers, TUG and mCTSIB. These outcomes are therefore likely candidate endpoints for clinical trials lasting 1 year. Most static balance tests are not responsive and not recommended in a heterogeneous DM1 population.

Introduction

Myotonic dystrophy type 1 (DM1) is the most common muscular dystrophy in adults,¹ characterized by distal limb, facial and bulbar muscle weakness, myotonia and multisystemic affection involving cognitive impairment, cardiac disease, metabolic abnormalities, and cataracts.² There is an unmet need for evidence-based outcomes in individuals with DM1^{3,4} as disease-modifying clinical trials are emerging.⁵ Validity and reliability of muscle strength, balance, and functional mobility outcomes have recently been established in DM1,⁶ but knowledge about responsiveness is lacking, which hampers the possibility to design and pick appropriate endpoints for interventional trials. Responsiveness is a tool's ability to detect change in a condition over time.⁷ Responsiveness of timed-up-and-go test (TUG) and handheld dynamometry (HHD) in the lower limbs has previously been investigated in DM1,⁸ but the follow-up period was 9 years. Because clinical trials often have a maximum duration of 1 year, the challenge is to identify responsive endpoints within 1 year despite the slowly progressive nature of DM1.⁴ Only the 30-second sit-to-stand

test (STS) and 10-meter walk test (10mWT, walk/run max pace) have been investigated after 1 year.⁹ Responsiveness of commonly used endpoints such as stationary dynamometry, step test, feet-together stance, tandem stance, one-leg stance, modified clinical test of sensory integration and balance (mCTSIB), 10mWT (walk, fast pace), and 10-times STS in DM1 is still unknown.

The objective of this study was to investigate responsiveness of muscle strength, balance, and functional mobility measurements after 1 year in individuals with noncongenital DM1.

Methods

Patients

From November 2017 to September 2019, 63 individuals with DM1 were recruited from a DM1 cohort⁶ at the Rigshospitalet ($n = 60$) and Aarhus University Hospital ($n = 3$) in Denmark (see Fig. 1). The inclusion criteria were genetically confirmed DM1 (CTG repeats > 80), 18–60 years, able to stand up from a chair with no arm

support, able to walk at least 10 meters (with or without gait aids) and reside close to Copenhagen or Aarhus. The exclusion criteria were congenital DM1 (defined as disease onset before 1 year), cognitive impairment preventing test adherence, non-DM1-related disorders or medicine consumption which confound muscle strength, balance, or functional test results, abuse of drugs or alcohol within 3 months, serious medical illness (e.g., symptomatic coronary artery disease and cancer), pregnancy, and clinically significant medical illness within 30 days.

Clinical assessments

Assessment of muscle strength, balance, and functional mobility were done twice separated by 1 year. All assessments have previously been described in detail.⁶ The same order of tests and procedures were repeated at follow-up by the same assessor for each patient at the same time of the day. Neither the patient nor the assessor were blinded to the test results, but recall bias is limited after 1 year, and therefore likely did not influence the results. The

patients were asked to wear closed, flat comfortable shoes and asked to refrain from exhausting or unusual physical activity the day before each visit to eliminate bias from muscle soreness or fatigue.

Muscle strength measurements

Maximal isometric muscle torque was tested with HHD (microFET2, Hoggan Scientific, LLC, Salt Lake City, UT) and stationary dynamometry (Biodex System 3 or 4 PRO, Biodex Medical Systems, NY). Newton from HHD was converted to Newton-meter: $Newton\text{-meter} (Nm) = Newton (N) * meter (m)$. Muscle strength was tested over the ankle, knee, and hip joints in the dominant leg. Two practice trials followed by three recorded trials were performed with standardized encouragement.

Balance measurements

The static balance measurements comprised of 60-second feet-together stance, 40-second tandem stance, 40-second one-leg-stance with eyes open and closed, and mCTSIB

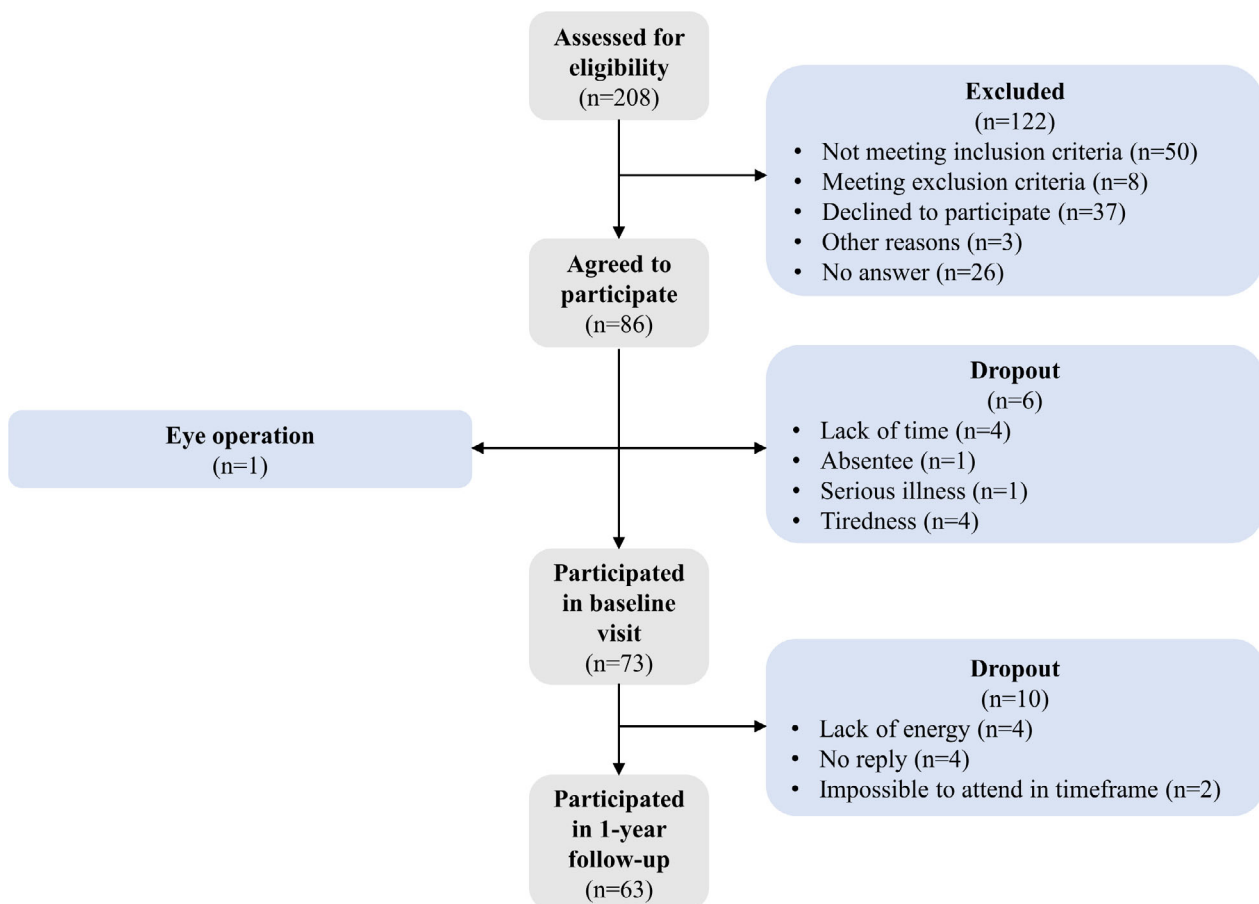


Figure 1. Flowchart of recruitment of patients in this study.

(eyes open and closed on a firm and foam surface) on a balance platform (BioSway Portable Balance System 950-460, Biodex Medical Systems, NY). The dynamic balance measurements consisted of 15-second step test and TUG. Two recorded trials were conducted with no encouragement. All balance measurements were performed with comfortable shoes and insoles and ankle-foot orthosis (AFO) were allowed. Other habitual walking aids such as a cane were only allowed for the TUG.

Functional mobility measurements

The functional mobility measurements included 10-times STS and 10mWT (both at the fastest possible pace). Two recorded trials were conducted with no encouragement.

Questionnaires

The International Physical Activity Questionnaire (short version) was applied to investigate the physical activity level.¹⁰ At follow-up, the patients completed a patient-reported global rating scale (GRS) to investigate whether a change had occurred or not from a subjective perspective, which the objective outcomes were compared to. The GRS questions for the objective tests were as follows: (I) Ankle muscle strength tests: Has your muscle strength in the lower limb/crus changed since the last visit (more difficulties standing on toes, tendency to stumble, slapping foot, walking longer distances)? (II) Knee and hip muscle strength tests: Has your muscle strength in the thigh and buttock changed since the last visit (more difficulties with rising from a chair (use of arms), climbing stairs (use of arms), walking longer distances)? (III) Dynamic balance tests: Has your balance during movement changed since the last visit (tendency to fall, need to lean on objects)? (IV) Static balance tests: Has your balance when you are standing still changed since the last visit (tendency to fall, need to lean on objects)? (V) STS test: Has your ability to rise from a chair changed since the last visit (more difficulties with rising from a chair (use of arms))? (VI) 10mWT: Has your ability to walk shorter distances changed since the last visit (e.g., slower walking at home)? The possible answers were as follows: (I) Much deterioration, (II) Some deterioration, (III) Stability, (IV) Some improvement, or (V) Much improvement. Each objective test was compared to the GRS-question that addressed the investigated construct of the objective test.

Statistical analysis

Linear mixed model was conducted to investigate statistically significant change between baseline and follow-up

(mean \pm SE) with family as a random effect, visit as a covariate and with unstructured covariance to account for repeated measurements over time in the same patients. If the model assumptions were not fulfilled, data were log₂-arithmetic transformed for analyses and antilog₂-arithmetic back-transformed for interpretation. In case of genuine outliers, sensitivity analysis of data with and without outliers was conducted. Based on a previous study,¹¹ floor and ceiling effects were defined as >15% of the patients scoring the lowest or highest score, respectively.

Secondary analyses

Receiver operating characteristic curve and area under the curve were conducted using GRS as anchor for whether a change had occurred or not. The model assumptions were checked, and the five GRS categories were dichotomized into Worse versus Stable/Better. Area under the curve estimates how good the measurements are to correctly classify change or no change compared to the anchor,¹² and the following guideline was applied¹³: 0.50 = no discrimination; 0.50-0.70 = poor discrimination; 0.70-0.80 = acceptable discrimination; 0.80-0.90 = excellent discrimination; 0.90-1.00 = outstanding discrimination.

Subanalyses of patients able to perform ankle dorsal flexion in the stationary dynamometry were done, because a floor effect was found in 21% of patients. This is relevant for clinical trials with an inclusion criterion of preserved dorsal ankle flexion strength to perform dynamometry assessment. Subanalyses based on age, age at onset of disease and CTG repeat size might also be relevant, but was deselected to avoid the risk of mass-significance with multiple testing.

A difference between patients who completed the study versus patients who declined to participate or dropped out was tested by unpaired t-test (continuous data), Mann-Whitney test (ordinal data), and Fisher's exact test (dichotomous data) if the model assumptions were fulfilled.

Mann-Whitney test was used to analyze if there was a difference in cognition and apathy among patients with agreement and disagreement between the subjective GRS ratings and the objective measurements, respectively. Because these secondary analyses were only exploratory, Bonferroni correction was not applied.

Ethics

The Regional Committee of Health Research Ethics in Denmark approved the study (H-17017556) and informed written consent was obtained.

Table 1. Demographic data.

| | |
|--|------------|
| Sex, no. | |
| Female | 30 |
| Male | 33 |
| Age (years), mean (SD) | 41 (10) |
| BMI ¹ , median (IQR ²) | 24 (21-27) |
| MIRS ³ , no. | |
| Grade 1 | 0 |
| Grade 2 | 13 |
| Grade 3 | 2 |
| Grade 4 | 42 |
| Grade 5 | 6 |
| Walking aid, no. | |
| Insoles | 1 |
| AFO ⁴ | 8 |
| Three-wheeled scooter | 1 |
| Cane | 1 |
| Walker | 1 |
| AES-S ⁵ , median (IQR ²) | 12 (8-16) |
| Apathy, no. | 0 |
| STROOP ⁶ , median (IQR ²) | |
| Word score | 32 (27-37) |
| Cognitive impairment, no. | 24 |
| Color score | 34 (31-40) |
| Cognitive impairment, no. | 11 |
| Color-Word score | 37 (34-45) |
| Cognitive impairment, no. | 3 |
| Interference score | 50 (50-51) |
| Cognitive impairment, no. | 0 |

Missing values for STROOP word (n = 11) due to test was not implemented (n = 7), invalid score (n = 1), patient was color blind (n = 2), and patient was unable to read (n = 1).

Missing values for STROOP color, color/word and inferences (n = 10) due to test was not implemented (n = 7), patient was color blind (n = 2), and patient was unable to read (n = 1).

¹BMI = Body mass index (kg/m^2), $\frac{\text{weight}(\text{kg})}{\text{height}(\text{m}^2)}$.

²IQR = Interquartile range.

³MIRS = Muscular impairment rating scale. Grade 1 = no muscular impairment, grade 2 = minimal weakness, grade 3 = distal weakness, grade 4 = mild to moderate proximal weakness, grade 5 = severe proximal weakness.²²

⁴AFO = Ankle-foot orthosis.

⁵AES-S = Apathy evaluation scale (Self-rated). A score > 34 = apathy. Missing values (n = 8) due to test was not implemented (n = 7) and incomplete test (n = 1).

⁶Verbal STROOP color and word test (Adult version). A higher score means better cognitive performance. The 95% CI for normal cognition score measured by STROOP is 30.4 to 69.91.

Results

For demographic data, see Table 1. The follow-up visit was performed after a median of 12 months (IQR 11.75-12.5 months), and the time of day between baseline and follow-up varied with a median of 0.5 hours

(IQR 0.25-2.00 hours). For feet-together stance, tandem stance, and one-leg-stance eyes open, inference statistics were not calculated because of null-inflation (i.e., no true variation among patients) due to ceiling effects.

Muscle strength measurements

A change in muscle strength from baseline to follow-up was captured by stationary dynamometry in the flexor and extensor muscles over the knee and hip joints, and by HHD in the ankle plantar flexors, knee flexors (without 2 outliers), and hip flexors ($P \leq 0.03$) (Table 2).

Twenty-one percent of the patients were unable to overcome the threshold in the stationary dynamometry for ankle dorsal flexors. Subanalysis without these patients did not change the results regarding change.

Balance measurements

All static balance tests, except mCTSIB, showed either a ceiling effect (maximum score) or a floor effect (minimum score) at both baseline and follow-up in many patients (Fig. 2). The balance tests that were able to detect changes at follow-up were the dynamic balance test TUG and the static balance test mCTSIB ($P \leq 0.035$) (Table 3).

Functional mobility measurements

None of the functional mobility tests (10mWT and STS) captured a change at follow-up ($P \geq 0.88$) (Table 3).

Secondary analyses

Outcome measures against GRS

The objective outcome measure results were generally not reflected in the subjective perceptions of change or no change as measured by GRS, because only the 10mWT and ankle plantar flexors with stationary dynamometry reached acceptable agreement with the GRS (area under the curve > 0.70, Fig. 3).

For the 10mWT, the cognition was better in the patients with disagreement between 10mWT change and GRS change (median, IQR; 43.00, 39.25-48.00) versus the patients with agreement between the 10mWT and the GRS (38.25, 35.50-41.13) ($P = 0.02$). For the ankle dorsal flexors with HHD, apathy was less pronounced in patients with disagreement between HHD change and GRS change (median, IQR; 9, 6-12) compared to the patients with agreement between the HHD and the GRS

Table 2. Baseline and follow-up muscle strength measures and their change.

| | Baseline Mean (SD) | FU Mean (SD) | Change Mean (95% CI) for absolute and percentage differences | P-value |
|---|-------------------------------------|-------------------------------------|--|-----------|
| Stationary dynamometry¹ | | | | |
| (Nm) | | | | |
| Ankle plantar flexors | 27.25 (14.80; 39.20) ² | 26.95 (14.60; 47.10) ² | 1.03 (0.94; 1.12); 3% (-6%; 12%) ³ | 0.52 |
| Ankle dorsal flexors | 16.80 (7.70; 27.30) ² | 17.15 (8.50; 27.80) ² | (0.97; 1.06); 1% (-3%; 6%) ³ | 0.50 |
| Knee extensors | 139.78 (64.49) | 133.50 (66.73) | -6.02 (-11.06; -0.98); -4.31% (-7.91%; -0.70%) | 0.02*↓ |
| Knee flexors | 61.14 (27.22) | 58.20 (26.37) | -2.87 (-4.95; -0.79); -4.69% (-8.09%; -1.30%) | 0.009*↓ |
| Hip extensors | 128.40 (98.20; 176.70) ² | 106.75 (72.00; 143.60) ² | -28.65 (-37.88; -19.42); -20.16% (-26.66%; -13.67%) | <0.0001*↓ |
| Hip flexors | 69.42 (28.42) | 85.59 (30.45) | 16.35 (12.41; 20.29); 23.55% (17.88%; 29.23%) | <0.0001*↑ |
| HHD⁴ (Nm) | | | | |
| Ankle plantar flexors | 18.08 (10.37; 25.72) ² | 15.96 (10.36; 24.65) ² | -1.33 (-2.35; -0.31); -7.17% (-12.66%; -1.68%) | 0.01*↓ |
| Ankle dorsal flexors | 18.21 (5.70; 25.74) ² | 18.14 (6.66; 27.54) ² | (0.93; 1.10); 1% (-7%; 10%) ³ | 0.79 |
| Knee extensors | 105.30 (38.24) | 108.58 (41.77) | 3.61 (-1.31; 8.53); 3.43% (-1.24%; 8.10%) | 0.16 |
| Knee flexors | 68.14 (26.41) | 65.73 (25.60) | -2.27 (-5.03; 0.49); -3.33% (-7.39%; 0.72%) | 0.11 |
| Hip extensors | 64.26 (22.63) | 59.93 (19.98) | -3.58 (-7.32; 0.16); -5.57% (-11.40%; 0.25%) | 0.065 |
| Hip flexors | 65.32 (20.43) | 61.30 (22.36) | -3.83 (-7.16; -0.50); -5.86% (-10.96%; -0.76%) | 0.03*↓ |

Note. *P-value ≤ 0.05. ↑improvement, ↓deterioration. The 95% CI for difference is based on SE. FU = 1-year follow-up.

¹Stationary dynamometry: Ankle plantar flexors: missing values (n = 1) because of technical issues. Ankle dorsal flexors: missing values (n = 1) because of technical issues. Knee extensors: missing values (n = 2) because of not conducted due to knee pain (n = 1) and technical issues (n = 1). Knee flexors: missing values (n = 1) because of technical issues. Hip extensors: missing values (n = 1) because of technical issues. Hip flexors: missing values (n = 1) because of technical issues.

²Median (IQR) because mean - (1.96*SD) resulted in a negative value which is meaningless.

³Antilog₂: ratio geometric mean (95% CI); percentage mean (95% CI).

⁴HHD: Ankle plantar flexors: missing values (n = 0). Ankle dorsal flexors: missing values (n = 0). Knee extensors: missing values (n = 1) because of knee pain. Knee flexors: missing values (n = 0). Hip extensors: missing values (n = 2) because of patient compensation (n = 1) and tester unable to hold position (n = 1). Hip flexors: missing values (n = 1) because tester was unable to hold position (n = 1).

(13, 10-17) (P = 0.048). For the other outcomes, there was no difference regarding cognition or apathy between the two groups (P ≥ 0.078).

Comparison with dropouts

There was no difference between patients who completed the study versus dropouts regarding sex, age, BMI, and level of muscle affection (Muscular Impairment Rating Scale (MIRS)) (P ≥ 0.15). The patients who completed the study were younger (41 ± 10 years) than the patients who declined to participate (47 ± 9 years) (P = 0.008), but there was no difference in sex (P = 0.49).

Physical activity level

There was no significant difference in physical activity level from baseline (median, IQR, 375.0, 200.0 to 620.0 min) to follow-up (median, IQR, 372.5, 180.0 to 810.0 min) (P = 0.32).

Discussion

The main findings of this study are that muscle strength and the balance measurements TUG and mCTSIB are responsive to change over a 1-year period in a cohort of noncongenital DM1. In contrast, all other static balance

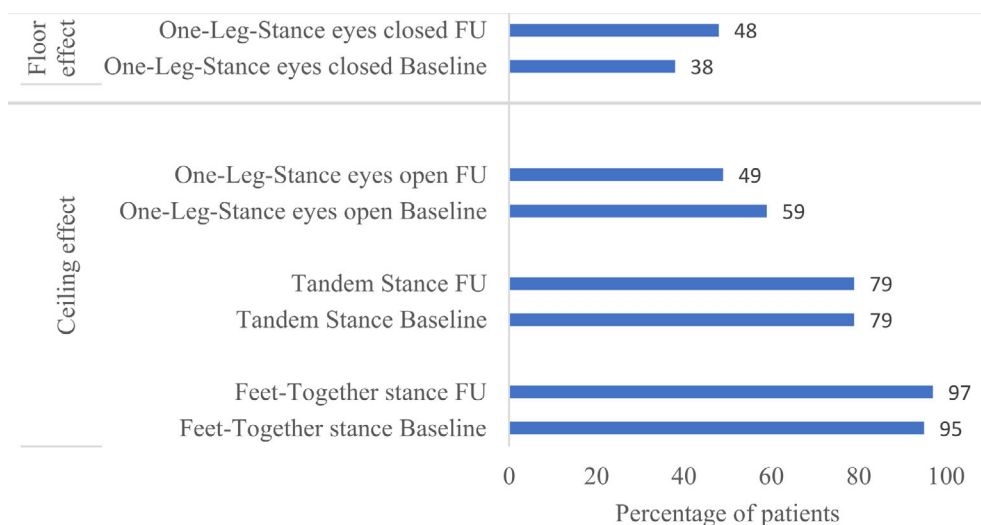


Figure 2. Ceiling and floor effects. The percentage of patients with either ceiling or floor effects is shown for both baseline and 1-year follow-up (FU).

measurements than the mCTSIB are unresponsive because of either ceiling or floor effects.

Muscle strength measurements

Both stationary and handheld dynamometers captured a change in the proximal leg muscles. The study shows that although distal muscles are well-known to be more affected in DM1, proximal muscles capture change better than distal, likely because distal muscles are very weak in DM1 and have reached an end-stage. Thus, 3–21% of the patients were too weak to exceed the threshold of the dynamometers, and therefore unresponsive to change. Moreover, 76% of the patients demonstrated proximal weakness (MIRS ≥ 4). Although HDD and stationary dynamometry agreed for most measures, there was some discrepancy for a few tests. Thus, change in strength was only recorded by HDD ankle plantar flexors, which could relate to easier patient-tester cooperation and smaller strength variation within the patients assessed by HDD, whereas the change in knee extensors, which was only captured by stationary dynamometry, could be due to tester-independency of stationary dynamometry. A significant increase in hip flexor strength assessed by stationary dynamometry was found in this study as well as in a previous DM1 study.⁸ However, this is considered a spurious type II error finding because of no change in physical activity level in this study and the loss of hip flexor strength recorded by HDD in the present- and a previous DM1 study.¹⁴ This study demonstrated loss of strength in knee and hip flexors and ankle plantar flexors assessed by HDD. Our study did not show change in knee extensors

and ankle dorsal flexors with HDD, but studies of longer duration in DM1 have shown this.^{8,14} It has previously been shown that 1 year is too short to register progression in these muscles in DM1.¹⁵ Compared to previous findings of significant decline in the knee extensors, but not in the knee flexors using stationary dynamometry in DM,¹⁶ this study found a reduction of strength in both knee extensors and flexors. The discrepancy may be because Lindemann et al.¹⁶ did not specify DM-type, investigated isokinetic torque, and had a smaller sample size, which reduces power.

The HDD has limitations when a subject is stronger than the investigator, but in this study, this problem was minimal as only one patient could overcome the assessor. Reversely, HDD is superior to stationary dynamometry as illustrated by 21% of the patients who could not exceed the threshold in the stationary dynamometry, whereas this number was only 13% for HDD.

Balance measurements

The TUG and mCTSIB captured change, but the step test and one-leg-stance eyes closed test, measuring different aspects of the same construct, failed to do so. The failure of capturing a change in the one-leg-stance eyes closed test was probably caused by a floor effect, missing values, and heterogeneity of the patients' performances. After 5–9 years of observation, deterioration in DM1 has been shown not only in the TUG^{8,14} but also in the step test.¹⁴

The ceiling effects in the feet-together stance, tandem stance, and one-leg-stance eyes open and the floor effect

Table 3. Baseline and follow-up balance and functional mobility measures and their change.

| | Baseline Mean (SD) | FU Mean (SD) | Change Mean (95% CI) for absolute and percentage differences | P-value |
|---|-----------------------------------|-----------------------------------|--|-----------------|
| Dynamic balance | | | | |
| TUG ¹ (s) | 8.11 (1.68) | 8.45 (1.41) | 0.35 (0.17; 0.53); 4.32% (2.14%; 6.49%) | 0.0003*↓ |
| Step Test ² (no.) | 18.53 (5.77) | 18.77 (6.12) | 0.24 (−0.27; 0.75); 1.30% (−1.45%; 4.05%) | 0.37 |
| Static balance | | | | |
| mCTSIB ³ (deg.) | 0.99 (0.25) | 1.05 (0.25) | 0.06 (0.001; 0.119); 6.06% (0.12%; 12.00%) | 0.035*↓ |
| Feet-together stance ⁴ (s) | 60 (60; 60) ⁵ | 60 (60; 60) ⁵ | NA ⁶ | NA ⁶ |
| Tandem stance ⁷ (s) | 40 (40; 40) ⁵ | 40 (40; 40) ⁵ | NA ⁶ | NA ⁶ |
| One-leg-stance eyes open ⁸ (s) | 40.00 (11.87; 40.00) ⁵ | 38.59 (12.19; 40.00) ⁵ | NA ⁶ | NA ⁶ |
| One-leg-stance eyes closed ⁹ (s) | 5.09 (3.70; 12.47) ¹⁰ | 7.83 (3.60; 17.56) ¹⁰ | 1.14 (0.84; 1.54); 14% (−16%; 54%) ¹¹ | 0.42 |
| Functional mobility | | | | |
| 10mWT ¹² (s) | 5.588 (1.33) | 5.582 (1.54) | −0.009 (−0.15; 0.13); −0.16% (−2.62%; 2.29%) | 0.90 |
| STS ¹³ (s) | 16.15 (4.32) | 16.09 (5.02) | −0.06 (−0.75; 0.63); −0.37% (−4.62%; 3.88%) | 0.88 |

Note. *P-value ≤ 0.05. ↓deterioration. The 95% CI for difference is based on SE. FU = 1-year follow-up.

¹TUG: missing values (n = 0).

²Step test: missing values (n = 1) because patient did not finish the test due to knee pain.

³mCTSIB: missing values (n = 14) because of technical issues (n = 2), inability to complete the test (n = 10), and inability to initiate the test (n = 2).

⁴Feet-together stance: missing values (n = 0).

⁵Median (IQR) because of null inflation (i.e., no true variation across patients).

⁶Impossible to estimate due to null inflation (i.e., no true variation across patients).

⁷Tandem stance: missing values (n = 0).

⁸One-leg-stance eyes open: missing values (n = 0).

⁹One-leg-stance eyes closed: missing values (n = 31) because of inability to initiate the one-leg-stance eyes open test (n = 1) and inability to stand ≥ 30 s in one-leg-stance eyes open test (n = 30) which qualified for the one-leg-stance eyes closed test.

¹⁰Median (IQR) because mean − (1.96*SD) resulted in a negative value which is meaningless.

¹¹Antilog₂: ratio geometric mean (95% CI); percentage mean (95% CI).

¹²10mWT: missing values (n = 0).

¹³STS: missing values (n = 1) because of incomplete test.

in the one-leg-stance eyes closed test suggests that these tests are unresponsive outcomes in a heterogeneous DM1 cohort. A ceiling effect has also previously been shown in the feet-together stance in DM1.¹⁷ The mCTSIB was close to reaching the threshold for floor effects with 13-14% of the patients being unable to complete the test. This indicates that the subparts of mCTSIB with eyes closed and standing on a foam surface may be too challenging for patients with more severe balance impairments.

Functional mobility measurements

None of the functional mobility tests in this study showed change in 1 year, which is at variance with other studies in DM1 using these tests.^{9,18} The discrepancies may be due to differences in test methods,^{9,18} unspecified DM-

type,¹⁸ and a larger sample size.⁹ Changes in the 10mWT after a longer period of time has previously been shown in DM1.¹⁴

General discussion

Outcome measures against GRS

The objective measurements were generally poorly reflected by the subjective scoring of changes in muscle strength, balance, and functional mobility using the GRS after 1 year. This relationship was better matched when the observation period was 9 years,⁸ where larger changes occur that can be more easily perceived. However, overall the true agreement between the objective measurements and subjective assessments in this study is somewhere

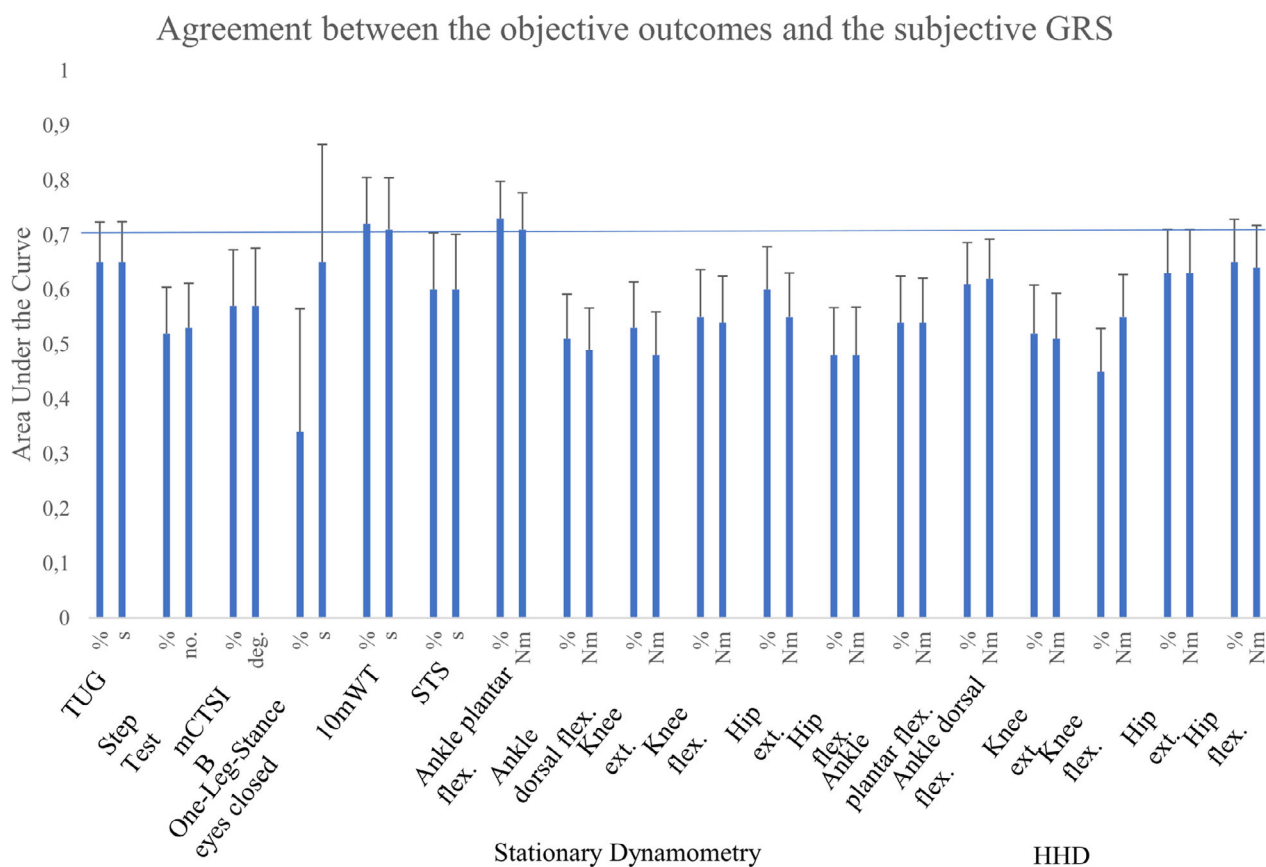


Figure 3. Agreement between the objective muscle strength, balance and functional mobility measurements and the subjective GRS. Area under the curve (Y-axis) is reported for both absolute and relative change values for each outcome measure with 95% CI (X-axis). Flex.=flexors, ext.=extensors.

between no agreement to excellent agreement with 95% confidence. This uncertainty suggests that GRS as an anchor of change or no change is unsuitable in DM1. Thus, defining responsive outcome measures using an anchor for the slowly progressive DM1 disease within a clinical trial duration of 1 year has proven problematic. This may be caused by the several limitations of GRS such as recall bias (inaccuracy of retrieving previous experiences¹⁹),⁸ response-shift bias (a shift in internal perceptions⁸), difficulty to perceive slow, gradual decline,⁸ well-being at the day of rating, cognition,⁸ and apathy. Moreover, the classification accuracy of change or no change by the objective measurements can only be as good as the anchor. Cognition was impaired in 5-38% of the patients in this study, but none of the patients reached the threshold for apathy (apathy score > 34²⁰). However, the patients with disagreement between GRS and the objective measurements did not show lower cognition, which suggests that the impact of cognition may be less than anticipated.

Implications for clinical trials

For clinical trials it is important to select the most responsive outcomes, so that small therapeutic effects are not concealed, and larger study cohorts can be avoided. Thus, based on this study, the best outcomes for clinical trials within 1 year are the stationary dynamometry or HHD with measurements of the proximal muscle groups and the balance assessments TUG or mCTSIB, because these tests capture subtle, but highly significant changes after 1 year of no intervention. For functional mobility, modified or other mobility tests, or novel outcomes such as gait analyses could be investigated for responsiveness to define additional responsive mobility outcomes for clinical trials of 1 year. However, the outcome measures that did not capture change after 1 year with no intervention may still capture a change after a 1-year interventional trial if the treatment is very effective, and therefore can still be considered as outcome. Responsiveness may be improved in a more homogeneous sample,²¹ but the

present sample was not large enough for a subgroup analysis.

The strength of the heterogeneous DM1 cohort investigated in this study is that it is generalizable to the majority of the noncongenital DM1 population.

Study limitations

The study was limited by dropouts, but since these constituted only 14% of the sample and did not differ clinically from the completers, this is not considered to influence conclusions significantly. The patient-rated GRS, in contrast to clinician-rated GRS, may be a limitation in DM1 patients due to possible symptoms such as lack of insight, apathy, and impaired cognition, but on the other hand patient-rated GRS represents the patients' own perceptions, which should be acknowledged. Efforts should therefore be directed at developing more suitable patient-reported-outcomes for DM1.

Conclusion

In conclusion, both muscle strength dynamometers and the balance measurements TUG and mCTSIB showed reasonable responsiveness by detecting subtle changes after 1 year in the slowly, progressive disease, DM1. All static balance measurements are not recommended as responsive outcomes due to either ceiling or floor effects, except the mCTSIB.

Acknowledgment

We thank grants from Axel Muusfeldt's Foundation, Familien Hede Nielsen's Foundation, and Rigshospitalet's Research Foundation. We also thank the patients for participating in the study.

Authors' contributions

Aisha Munawar Sheikh contributed with acquisition of data and drafting the manuscript. Nanna Witting contributed with design of the study, analysis of data, and drafting the manuscript. John Vissing contributed with design of the study, analysis of data, and drafting the manuscript.

Conflicts of interest

KLK reports grants from Axel Muusfeldt's Foundation, grants from Familien Hede Nielsen's Foundation, grants from Rigshospitalet's Research Foundation during the conduct of the study. AMS, NW, and JV have nothing to disclose.

References

1. Hammarén E, Kjellby-Wendt G, Kowalski J, Lindberg C. Factors of importance for dynamic balance impairment and frequency of falls in individuals with myotonic dystrophy type 1 – A cross-sectional study – Including reference values of Timed Up & Go, 10m walk and step test. *Neuromuscul Disord* 2014;24(3):207–215.
2. Andersen G, Ørngreen MC, Preisler N, et al. Muscle phenotype in patients with myotonic dystrophy type 1: Muscle Phenotype in DM1. *Muscle Nerve* 2013;47(3):409–415.
3. Foff EP, Mahadevan MS. Therapeutics development in myotonic dystrophy type 1. *Muscle Nerve* 2011;44(2):160–169.
4. Gagnon C, Meola G, Hébert LJ, et al. Report of the second Outcome Measures in Myotonic Dystrophy type 1 (OMMYD-2) international workshop San Sebastian, Spain, October 16, 2013. *Neuromuscul Disord* 2015;25(7):603–616.
5. Nguyen C-TE, Campbell C. Myotonic dystrophy type 1. *Can Med Assoc J* 2016;188(14):1033.
6. Knak K, Sheikh A, Andersen H, et al. Intra-rater reliability and validity of outcome measures in myotonic dystrophy type 1. *Neurology* In press.
7. Portney L, Watkins M. Validity of Measurements. In: *Foundations of Clinical Research - Applications to Practice*. 3rd Edition. New Jersey: Pearson Education; 2009.
8. Kierkegaard M, Petitclerc É, Hébert LJ, et al. Responsiveness of performance-based outcome measures for mobility, balance, muscle strength and manual dexterity in adults with myotonic dystrophy type 1. *J Rehabil Med* 2018;50(3):269–277.
9. Jimenez-Moreno AC, Nikolenko N, Kierkegaard M, et al. Analysis of the functional capacity outcome measures for myotonic dystrophy. *Ann Clin Transl Neurol* 2019;6(8):1487–1497.
10. International Physical Activity Questionnaire. International Physical Activity Questionnaire [Internet]. 2020. Available from: https://sites.google.com/site/theipaq/questionnaire_links.
11. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60(1):34–42.
12. Portney L, Watkins M. Statistical Measures of Validity. In: *Foundations of Clinical Research - Applications to Practice*. 3rd Edition. New Jersey: Pearson Education; 2009.
13. Laerd Statistics. ROC Curve [Internet]. Available from: <https://statistics.laerd.com/premium/spss/blr/binomial-logistic-regression-in-spss-12.php#roc>.
14. Hammarén E, Kjellby-Wendt G, Lindberg C. Muscle force, balance and falls in muscular impaired individuals with myotonic dystrophy type 1: A five-year prospective cohort study. *Neuromuscul Disord* 2015;25(2):141–148.

15. Sedehizadeh S, Brook JD, Maddison P. Body composition and clinical outcome measures in patients with myotonic dystrophy type 1. *Neuromuscul Disord* 2017;27(3):286–289.
16. Lindeman E, Leffers P, Spaans F, et al. Deterioration of motor function in myotonic dystrophy and hereditary motor and sensory neuropathy. *Scand J Rehab Med* 1995;27:59–64.
17. Hammarén E, Ohlsson JA, Lindberg C, Kjellby-Wendt G. Reliability of static and dynamic balance tests in subjects with myotonic dystrophy type 1. *Adv Physiother* 2012;14(2):48–54.
18. Nitz JC, Burns YR, Jackson RV. A longitudinal physical profile assessment of skeletal muscle manifestations in myotonic dystrophy. *Clin Rehabil* 1999;13(1):64–73.
19. Wikipedia. Recall bias. 2020. Available from: https://en.wikipedia.org/wiki/Recall_bias
20. Shirley Ryan Abilitylab. Apathy Evaluation Scale (Self, Informant, and Clinician Versions). Available from: <https://www.sralab.org/rehabilitation-measures/apathy-evaluation-scale-includes-three-forms-self-informant-and-clinician>
21. Lencioni T, Piscoquito G, Rabuffetti M, et al. Responsiveness of gait analysis parameters in a cohort of 71 CMT subjects. *Neuromuscul Disord* 2017;27(11):1029–1037.
22. Mathieu J. Assessment of a disease-specific muscular impairment rating scale in myotonic dystrophy. *Neurology* 2001;56:336–340.