Walking test outcomes in adults with genetic neuromuscular diseases: A systematic literature review of their measurement properties

Nawale Hadouiri, Isabelle Fournel, Christel Thauvin-Robinet, Agnès Jacquin-Piques, Paul Ornetti, Mathieu Gueugnon

European Journal of Physical and Rehabilitation Medicine 2024 Feb 01
DOI: 10.23736/S1973-9087.24.08095-X

SYSTEMATIC REVIEW

Walking test outcomes in adults with genetic neuromuscular diseases: a systematic literature review of their measurement properties

Nawale HADOUIRI [1, 2, 3] *, Isabelle FOURNEL [4, 5], Christel THAUVIN-ROBINET [2, 6, 7], Agnès JACQUIN-PIQUES [8], Paul ORNETTI [9, 10], Mathieu GUEUGNON [3, 10]

[1]Department of Physical Medicine and Rehabilitation, Dijon-Bourgogne University Hospital, Dijon, France; [2]UMR-Inserm 1231, Génétique des Anomalies du Développement (GAD), Bourgogne Franche-Comté University, Dijon, France; [3]INSERM, CIC 1432, Clinical Investigation Center, Plurithematic Module, Technological Investigation Platform, Dijon-Bourgogne University Hospital, Dijon, France; [4]Clinical Investigation Center, CHU Dijon, Dijon, France; [5]INSERM, CIC 1432, Module Epidémiologie Clinique, Dijon, France; [6]Fédération Hospitalo-Universitaire Médecine Translationnelle et Anomalies du Développement (TRANSLAD), CHU Dijon Bourgogne, Dijon, France; [7]Centre de Référence Maladies Rares "Maladies neurogénétiques", CHU Dijon Bourgogne, Dijon, France; [8]Centre de Compétences Maladies Rares "Maladies neuromusculaires", Department of Neurology, Dijon University Hospital, Dijon, France; [9]Department of Rheumatology, Dijon-Bourgogne University Hospital, Dijon, France; [10]INSERM, UMR1093-CAPS, Bourgogne Franche-Comté University, Dijon, France

*Corresponding author: Nawale Hadouiri, Department of Physical Medicine and Rehabilitation, Dijon-Bourgogne University Hospital, Boulevard du Maréchal de Lattre de Tassigny 2, 21000 Dijon, France. E-mail: nawale.hadouiri@chu-dijon.fr

ABSTRACT

INTRODUCTION: Neuromuscular diseases (NMDs) include a large group of heterogeneous diseases. NMDs frequently involve gait disorders, which affect quality of life. Several walking tests and tools have been described in the literature, but there is no consensus regarding the use of walking tests and tools in NMDs or of their measurement properties for walking outcomes. The aim of this review is to present an overview of walking tests, including their measurement properties when used in adults with inherited or genetic NMDs. The aim is to help clinicians and researchers choose the most appropriate test for their objective.

EVIDENCE ACQUISITION: A systematic review was conducted after consulting MEDLINE (*via* PubMed), EMBASE, Science direct, Google Scholar and Cochrane Central Register of Controlled Trials databases for published studies in which walking outcome measurement properties were assessed. The validity, reliability, measurement error and responsiveness properties were evaluated in terms of statistical methods and methodological design qualities using the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) guidelines.

EVIDENCE SYNTHESIS: We included 46 studies in NMDs. These studies included 15 different walking tests and a wide variety of walking outcomes, assessed with six types of walking tools. Overall, the 6MWT was the most studied test in terms of measurement properties. The methodological design and statistical methods of most studies evaluating construct validity, reliability and measurement error were "very good." The majority of outcome measurements were valid and reliable. However, studies on responsiveness as minimal important difference or minimal important change were lacking or were found to have inadequate methodological and statistical methods according to the COSMIN guidelines.

CONCLUSIONS: Most walking outcomes were found to be valid and reliable in NMDs. However, in view of the growing number of clinical trials, further studies are needed to clarify additional measurement properties.

(*Cite this article as:* Hadouiri N, Fournel I, Thauvin-Robinet C, Jacquin-Piques A, Ornetti P, Gueugnon M. Walking test outcomes in adults with genetic neuromuscular diseases: a systematic literature review of their measurement properties. Eur J Phys Rehabil Med 202 Feb 01. DOI: 10.23736/S1973-9087.24.08095-X)

KEY WORDS: Neuromuscular diseases; Walking; Charcot-Marie-Tooth disease.

Introduction

Neuromuscular diseases (NMDs) constitute a large group of rare diseases[1] that are mainly of genetic origin and whose prevalence is often below 1 case per 2000 people.[1] To date, more than 600 genes are known to be implicated.[2] The most frequently encountered NMDs in adulthood are the Charcot-Marie-Tooth (CMT) disease (prevalence ranging from 9.7 to 82.3 per 100,000 persons in the Caucasian population),[3] myotonic dystrophy type 1 (DM1) and type 2 (DM2) (prevalence ranging from 5 to 20 per 100,000 persons)[4] and facioscapulohumeral muscular dystrophy (FSHD) (prevalence ranging from 4 to 10 per 100,000 persons).[5] All of these NMDs have a common deficiency according to the International Classification of Functioning (ICF),[6] which is muscle weakness in the upper and/or lower limbs consecutive to the defect of the motor unit.[7] The affected muscle topography depends on the type of NMD (for example, CMT patients experience distal upper limb and proximal and distal lower limb muscle weakness and atrophy).[8] Gait disorders, which are one of the main disabilities that result from lower limb muscle weakness, have a considerable impact on quality of life.[9] This makes gait a relevant endpoint for the functional evaluation of NMD patients. Few interventional trials have assessed walking improvement in inherited NMDs compared to other neurological diseases. However, for the last few years, there has been a lot of hope around gene therapy following advances in the diagnosis of NMDs, which increasingly opens the field of prospects to more ambitious clinical trials.[10]

It is important for clinicians and researchers to use valid, reliable, feasible, and responsive walking tests to assess walking disability, changes over time, and the effectiveness of interventions. In the clinical trials that have been published so far, various walking tests and tools have been used to assess walking disabilities in individuals with NMDs. Many walking tests are classified as short tests. For example, the 10-m test is a simple test used to measure locomotor capacity in clinical and research settings and in which the time taken to complete the test or the mean velocity is assessed. On the other hand, there are also prolonged tests. For instance, in the 2-Minute Walk Test (2WMT) and the 6-Minute Walk Test (6MWT), individuals are instructed to walk as far as possible in 2 and 6 minutes, respectively. These prolonged tests, which are submaximal exercise tests, are used to assess aerobic capacity and endurance, with more ecological properties (*i.e.*, to have a better picture of everyday life) than short tests. Indeed, because most activities of daily living are performed at submaximal levels of exertion, prolonged tests like the 6MWT may better reflect the functional exercise level and motor performance for daily physical activities.[11] However, these walking tests and tools only evaluate walking capacities in the strict sense (*i.e.*, what people can do in a standardized and controlled environment). There has been a recent interest in developing home-based monitoring devices that could be used in future therapeutic trials in NMDs to assess their walking capability (*i.e.*, what people can do in their daily environment) and motor performance (*i.e.*, what a person actually does do in his/her daily environment).[12]

In interventional trials in adults, the 6MWT seems to be the most frequently used walking test in NMDs such as CMT,[13, 14] DM1[15, 16] and DM2,[17] late-onset Pompe disease (LOPD),[18, 19] FSHD[20-22] and other muscular dystrophies.[23] In these studies, the 6MWT distance was the only recorded walking assessment variable for analyzing changes in walking capacity after a medical or physical intervention. The acquisition of walking speed was also observed in individuals with NMDs with the 10-m walking test performed at a comfortable and/or fast speed[14, 18, 24-28] or with the 2MWT.[29] A simple stopwatch was used to assess walking speed.[14, 15, 18, 27, 28, 30] However, complex tools such as a 3D motion analysis system can more thoroughly evaluate the progression of walking capacities, i.e., in CMT[31-33] or in hereditary spastic paraplegia (HSP),[26, 34, 35] through kinetic, kinematic and electromyographic (EMG) data acquisition.[31-35]

While many walking tests and assessments tools are currently available, only some of them have been assessed for their measurement properties. To the best of our knowledge, no review on this subject has been published so far. However, a complete overview of walking outcome measurement properties in NMDs would be useful for clinicians and researchers when choosing the most appropriate walking test for their specific objective (*e.g.*, to screen pertinent walking outcomes for future clinical trials, or as robust easy-to-use clinical measures). The aim of the present review was to provide an overview of walking test outcomes in adults with genetic NMDs and to provide information concerning the measurement properties of the walking tests in terms of validity, reliability, measurement error and responsiveness.

Our hypotheses were that: 1) the measurement properties of the studied walking tests might depend on the subtype of NMD; and 2) the 6MWT, which measures the total distance walked during the test, would be the most studied

test in terms of measurement properties considering that it is frequently used in clinical trials for patients with NMDs.

Evidence acquisition

This systematic review was conducted according to PRISMA guidelines (*i.e*., Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA Statement).[36] The protocol was submitted *a priori* to the PROSPERO registry (registration number CRD422022366521) on October 11, 2022. Each step was carried out independently by two investigators (N.H., M.G.) and then compared. In case of a disagreement, a third investigator settled the case (P.O.).

Search strategy

The MEDLINE (*via* PubMed), EMBASE, Science direct, Google Scholar and Cochrane Central Register of Controlled Trials databases were queried until November 30th, 2022 with MeSH terms and other free words divided into three components:

- studieddiseases(as("NeuromuscularDiseases"[Mesh]) *OR* ("Muscular Dystrophy, Duchenne"[Mesh]) *OR* ("Charcot-Marie-Tooth Disease"[Mesh]) *etc.),*
- the studied outcome ("gait"[MeSH] and synonyms such as "walking"[MeSH]),
- a search for measurement properties (with terms such as "Validation Study" [Publication Type] *OR* "Reproducibility of Results"[Mesh] etc.). The complete search formulation is available in Supplementary Digital Material 1 (Supplementary Text File 1).

The search was limited to studies published in French or English to prevent translation errors. The bibliographies of the included studies were also checked for additional eligible studies. There were no restrictions for the publication date.

Selection of studies

We excluded non-original articles or articles not published in peer-reviewed journals, such as theses, protocol studies, conference proceedings, letters to the editor, case reports, editorials and systematic reviews with or without meta-analysis. Only studies with full text available were included. Simulation studies were excluded.

We included studies that evaluated walking capacities with measurement of their properties and quality of outcome measures (with a focus on validity, reliability, responsiveness, measurement error studies because these are the main metrological properties studied) in adults with genetic NMDs. Studies with subjective evaluation of walking capacities (*i.e*., only with the use of questionnaires/self-reporting, clinical walking descriptions etc.) were excluded. All study designs were included (*i.e*., interventional studies, cohort with persons with one type of NMD, persons with NMD *versus* healthy persons, several subtypes of NMDs, etc.).

Only studies in adults with inherited NMDs were included. We excluded studies concerning non-inherited NMDs such as poliomyelitis or inclusion body myositis. Studies conducted only in children or in animal models were also excluded.

After removing duplicates, the title and then abstracts of the articles identified in the database search were analyzed for eligibility based on the previously cited inclusion criteria.

Data extraction and quality assessment

Data including study design, sample size, characteristics of participants (type of NMDs, age of participants, presence of a healthy group), the type of walking test and tools, the walking variables assessed, the presence of a physical examination of participants (specifying the type of lower limb physical examination and/or the functional scale used) were extracted by one reviewer (N.H.).

Concerning the data for the main outcome(s) (measurement properties of walking tests), data were extracted in accordance with the methodological guidelines of the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist.[37] The COSMIN checklist consists of ten sub-checklists for different measurement properties (patient-reported outcome measures (PROM) development, internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing for construct validity, cross-cultural validity, criterion validity and responsiveness). In the present COSMIN review, we studied four types of measurement properties for walking outcomes:

- validity of outcomes such as construct validity (*i.e*., if the studied test or measure accurately assesses what it is supposed to), face validity (*i.e*., if the content of the test appears to be suitable to its aims), and criterion validity (*i.e.*, if a gold standard is presented; how a test measures the outcome it was designed to measure);
- reliability (*i.e*., the fact that an instrument gives the same measure each time it is used), including test-retest reliability, intra-rater reliability, inter-rater reliability;
- measurement error;
- responsiveness of walking outcome measures.

If presented in the selected articles, feasibility data were described rather than evaluated, as these are not formal measurement properties according to the step 8 of the COSMIN guidelines (*i.e*., "they do not provide to us something about the quality of an instrument, but they are important aspects that have to be considered in the selection of the most suitable instrument for a specific aim").[38]

The COSMIN checklist was specifically designed to determine methodological quality scores and to assess the risk of bias for the result of each study through the evaluation of design aspects and statistical methods.[38] It was adapted for use on "walking tests and tools" by substituting "health-related patient-reported outcomes" (PROM) by "measurement instruments"; this was done in previous COSMIN reviews concerning the assessment of measurement properties of walking outcomes in other diseases.[39-41] Here, only walking outcome measurement properties were extracted. If there were numerous walking outcomes in a given study, notably for spatiotemporal parameters, we have chosen to only extract variables such as velocity (with or without cadence, stride length), stability parameters in the general walking pattern (*e.g*., double support time, stride width or support base) corresponding to the main gait parameters, but not all spatiotemporal parameters as swing phase. If available in analysis motion studies, we extracted a synthesis of kinematics and kinetics data. Concerning construct validity, we extracted correlations data between walking variables and quantitative physical examination variables (e.g., clinical reference scale, measurement of lower limb muscle strength with isokinetic devices or clinical scale), or/and another walking tests, but not with clinical data such as age or disease duration, or other variables such as quality of life. Firstly, scores (*i.e*., very good, adequate, doubtful, inadequate) were assigned to evaluate the methodological quality of each measurement property using the "worst score counts method."[42] Then, the result of each study for a measurement property was rated against the updated criteria for good measurement properties, based on the adapted quality tool from Prinsen *et al.*[38] and Terwee *et al.*[43] Each result for each measurement property is rated as either positive (+), indeterminate (?), or negative (-), as defined in Supplementary Digital Material 2 (Supplementary Table I). To qualitatively rate the results as sufficient (or insufficient), in principle, 75% of the results should meet the criteria.[38] As recommended by Prinsen *et al.*[38] and Terwee *et al.*,[43] the reviewers made hypotheses to evaluate construct validity and responsiveness based on De Vet *et al.*.[44] Correlations with (changes in) instruments measuring similar constructs should be ≥0.50, and for responsiveness, the area under the curve (AUC) should be ≥0.70. Under coefficient of correlation of 0.50 and AUC <0.70, the results were qualitatively summarized as insufficient. However, if it was impossible to formulate a pertinent hypothesis for a specific analysis, the results were qualitatively summarized as indeterminate. Two independent reviewers (N.H. and M.G.) extracted the outcome data and assessed the methodological quality. If disagreement persisted after discussion, a third reviewer (PO) was consulted.

Evidence synthesis

Study characteristics

Figure 1 shows the methodology of the selected studies. Overall, 46 studies were included for data analysis after the systematic review, including articles identified in the references of the included articles. The main characteristics of the included studies are summarized in Supplementary Digital Material 3 (Supplementary Table II),[24, 45-89] including the various study designs. Concerning the assessed NMDs, eight articles focused on CMT,[46-53] 11 on DM1,[54-64] one on DM1 and DM2,[65] one on DM2,[66] five on FSHD,[24, 67-70] one on FSHD and limb girdle muscle dystrophy (LGMD) type 2,[71] one on HSP,[72] two on LOPD,[73, 74] one on several muscular dystrophies (DMD, BMD, LGMD and FSHD),[75] three on people with several types of inherited NMD,[45, 76, 77] one on primary mitochondrial myopathy,[78] one on spinal bulbar muscular atrophy (SBMA),[79] and 10 on 5q spinal muscular atrophy (5q SMA) types 3 or 4.[80-89] Sample sizes varied from four to 479 participants, and 32 studies did not include a healthy control group.

The 46 included studies used 15 different walking tests. The 6-Minute Walking Test (6MWT) was used in 25 studies.[45, 49, 52-54, 56-58, 61, 63, 64, 66, 68, 73, 74, 76-80, 83-85, 87-89] The included studies reported most walking tests categorized according to the ICF as "walking a short distance" conditions:[6] 1) to measure walking velocity, such as the 10 Meter Walk Test (10MWT)[24, 46, 49, 51, 53, 57-60, 63, 64, 74, 75, 89] or 10 Meter Run Test,[56, 57, 63, 83, 86] and the 30 Foot Go Test;[68] 2) to evaluate walking around obstacles or dynamic balance stance, such as the Timed Up and Go Test (TUG),[56, 58-62, 64, 68-70, 74, 78, 83, 86] Step Test,[58] Flight of Eight (Fo8),[59] or Stance Tandem (ST);[60, 77] 3) to assess walking distance with walking tests or protocols using more time to measure walking distance, such as the 2-Minute Walk Test (2MWT),[45, 77] 6MWT,[45, 49, 52-54, 56-58, 61, 63, 64, 66, 68, 73, 74, 76-80, 83-85, 87-89] or Endurance Shuttle Walking test (ESWT);[81, 82] and 4) to measure walking without a particular distance or

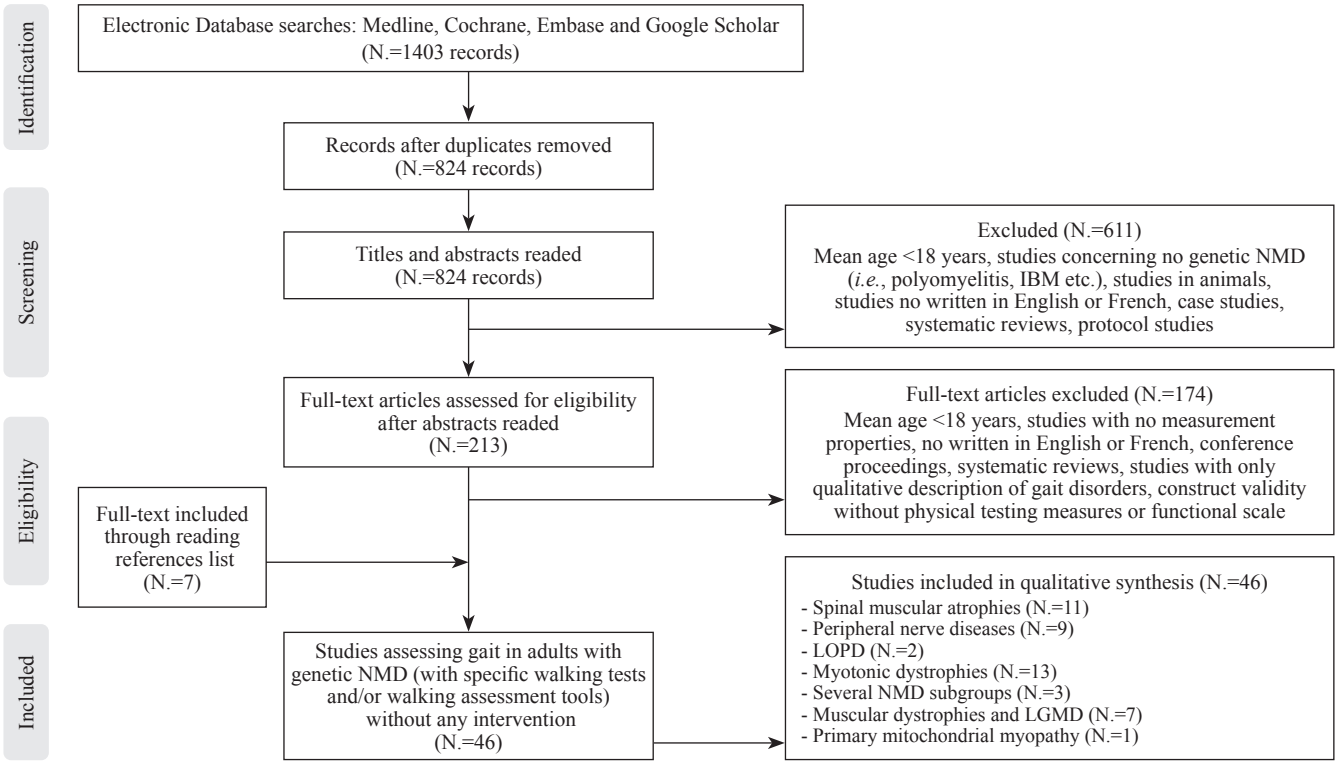

Figure 1.—Flowchart of the present systematic review.

time to accomplish,[47, 48] such as walking assessments with motion analysis system[47, 51, 55, 67, 72] or with a portable pressure sensitive walkway.[48, 65, 83, 88, 89] Only one study could be categorized as "walking a long distance" assessment conditions according to the ICF,[6] which included real-life assessment conditions using wearable magneto-inertial sensors.[71] Several types of walking assessment tools are listed, including easy-to-use tools such as a stopwatch,[45, 46, 52, 54, 56-60, 62, 64, 66, 73-82, 84-87] and other more technical tools such as a GaitRite walkway,[48, 65, 83, 87] activity monitoring watches,[49, 53] a motion analysis system,[47, 50, 51, 55, 67, 72] a baropodometric platform,[24] or wearable magneto-inertial sensors.[69, 71, 89]

Measurement properties

For clarity and readability, the methodological design, statistical methods and statistical outcomes for the variables of interest in the 46 included studies have been presented separately for CMT (Supplementary Digital Material 4: Supplementary Table III, IV, V),[46-53] DM1 and DM2 (Supplementary Digital Material 5: Supplementary Table VI, VII, VIII, IX),[54-66] and grouped together for the other NMDs (Supplementary Digital Material 6: Supplementary Table X, XI, XII, XIII).[24, 45, 67-89] The two reviewers had *moderate* agreement (absolute agreement = 0.69 and $\kappa = 0.59$, 95% CI 0.49-0.68) according the assessment of methodological quality using the COSMIN checklist. After discussion, the third reviewer was not consulted for any study. Construct validity and reliability (and especially test-retest reliability) were the most evaluated measurement properties in the included studies. Specifically, reliability was evaluated in 19 studies, measurement error in eight studies, responsiveness in 16 studies, construct validity in 35 studies, criterion validity in three studies, content validity in seven studies, convergent validity in one study and discriminative validity in one study.

CMT

In CMT (Supplementary Digital Material 4), the methodological quality of the studies examining construct validity were "very good" for all walking tests, and, according to our hypotheses, the hypotheses testing for construct validity were mostly "sufficient," except especially correlations between CMT neuropathy score (CMTNS) and 6MWD, 10MWT time and several outputs during the five-day monitoring as an activity index with $\rho<0.5$ (Supplementa-

ry Digital Material 4). Construct validity was particularly good for velocity assessed during the 10MWT, which was inversely strongly correlated with CMTNS (correlation coefficient ρ=-0.783).

The methodological quality of the three studies examining reliability were "adequate" (*i.e.*, for example when the model or formula of the ICC was not described) to "very good" for the 10MWT,[46] 6MWT[49] and kinetic variables[47] with "positive" quality criteria ratings for all these studies with excellent intraclass correlation coefficients (ICCs) (>0.9).

Only one study analyzed measurement error for walking assessment (concerning kinetic variables with a standard error measurement (SEM) of velocity and double support time of respectively 3.67 s and 1.11% of gait cycle, and SEM of Mean ankle angle Toe Walking - mean ankle angle High Walking of 2.7°) with "very good" methodological quality, but the quality criteria for hypothesis testing was "indeterminate."[47]

Two studies evaluated the responsiveness of the 6MWT, the 10MWT, several outputs from a StepWatch Activity Monitor,[53] and velocity acquired with motion analysis system.[50] Only one study had "very good" methodological quality with the calculation of the SRM of velocity (assessed with motion analysis system) of -0.55% body height/s and an "positive" quality criteria rating[50] according to COSMIN guidelines.

DM1 and DM2

In DM1 (Supplementary Digital Material 5), the methodological quality of the studies examining content and construct validity of the 6MWT was "very good" for all walking tests, and according to our hypotheses, the hypotheses testing for construct validity were mostly "sufficient." The best construct validity was observed for the 6-Minute Walking Distance (6MWD) and the 10MWT in running conditions (10mW/RT), which were moderately correlated with the Scale for the Assessment and Rating of Ataxia (SARA) rating (respectively ρ=0.65 and ρ=0.55). The 10MWT in comfortable conditions was strongly inversely correlated with total lower limb muscle strength (ρ=-0.705), the TUG was moderately inversely correlated with the MMT of the trunk muscle group (ρ=-0.58), and ST and velocity were strongly inversely correlated with total lower limb muscle strength (ρ=-0.705).

The methodological quality of the five studies examining test-retest reliability was "very good" for the 10MWT,[59] 6MWT,[54] TUG,[59] ST[59] and Fo8.[59] The studies which assessed intra-rater reliability of the 10MWT,[56-58] 6MWT,[56] TUG,[56, 58] ST[58] had "adequate" to "very good" methodological quality. All these reliability studies had a "positive" quality criteria rating and excellent ICCs (>0.9), except for the intra-rater reliability used during the 10MWT (ICC=0.86 [95% CI 0.74-0.93]) and the intra-rater reliability (1 to 2 weeks) of the TUG time (ICC=0.68 [0.54-0.79]).

Three studies analyzed measurement error of the 6MWT,[54] the 10MWT,[59] the TUG,[58, 59] the Fo8 and ST[59] with "very good" methodological quality (with the calculation of SEM with or without measurement error (ME)). However, the quality criteria for hypotheses testing was "indeterminate."

Three studies evaluated the responsiveness of velocity during the 10MWT and 10MW/RT,[57, 59, 60] 6MWT,[63] TUG,[62, 64] and ST[64] with methodological quality according to COSMIN guidelines "inadequate" except for 2 studies that assessed the area under curve (AUC),[62, 64] and only 1 study had a "positive" quality criteria rating in the TUG with a criterion approach walking AUC>0.7 according to the COSMIN guidelines (AUC=0.8 [95% CI 0.7-0.9]).[62] According to COSMIN guidelines, the responsiveness of the 10MWT in comfortable velocity, the TUG and ST were considered "very good" in terms of methodological quality (Supplementary Digital Material 6).

Two studies mentioned some feasibility data for all participants using the 6MWT, 10MWT, and 10MW/RT.[54, 57] All participants managed to complete at least one 6MWT as a whole or one trial of each walking test cited. However, this was not the case if the instruction was to perform two different 6MWT or to repeat several walking tests, with the main reported limitation being the generation of fatigue.[51, 54]

In DM2 (Supplementary Digital Material 5), there was only one specific study (concerning 6MWT measurement properties) with "very good" methodological quality according to the COSMIN guidelines concerning construct validity (weak correlation [ρ] between 6MWD and the sum of lower limb MMT ρ=0.492), and, according to our hypotheses, the hypotheses testing for construct validity were mostly "sufficient." In this study, the methodological quality of the responsiveness evaluation of the 6MWD was also "inadequate" according to the COSMIN guidelines.[66]

Other NMDs

Concerning the other NMDs (Supplementary Digital Material 6): in FSHD, the methodological quality of the studies examining construct validity were "adequate" to "very good", and according to our hypotheses, the hypotheses

testing for construct validity were mostly "sufficient." The best construct validity was observed for the 95th centile length assessed in real life by a wearable device which was strongly correlated with lower limb MMT (ρ=0.915, P<0.05).

The methodological quality of the two studies examining test-retest reliability were "adequate" to "very good" for the 6MWT[68] and instrumented TUG (velocity and double support)[69] with ICCs >0.9. We observed "very good" methodological quality for the study of responsiveness for the 6MWT, with the calculation of the minimal detectable change (MDC) with 95% confidence (MDC95=34.3 m).[68] The methodological quality of the responsiveness of spatiotemporal variables in real life according to the COSMIN guidelines was considered "adequate"[71] (Supplementary Digital Material 6).

In HSP, only one study, notably with a velocity assessment, was found to have "very good" methodological quality for the construct validity assessment, but a weak correlation (ρ=0.38) between walking velocity and Spastic Paraplegia Rating Scale, so the hypothesis testing for the construct validity was "indeterminate"[72].

In LOPD, two studies with only people with LOPD were found to have "inadequate" methodological quality for the responsiveness assessment of the 6MWT, 10MWT and TUG,[73, 74] and the quality criteria rating was "indeterminate."

In SBMA, only one study using the 6MWT was found to have "very good" methodological quality for the test-retest reliability assessment (with ICC=0.982), with a "positive" quality criteria rating, and a "very good" methodological quality for the construct validity assessment. The best construct validity was observed for the 6MWD, which was moderately correlated with Limb Norris Score (ρ=0.632; P<0.001), but hypothesis testing for construct validity was "indeterminate."[79]

In 5q SMA, the methodological quality of the study examining the criteria validity of several spatiotemporal parameters assessed by Solesound instrumented footwear (the gold standard was the assessment of these same spatiotemporal parameters by a GAITRite system) were "very good" with all correlation coefficients ρ > 0.9, and the criterion quality ratings were considered "positive" according to the COSMIN guidelines.[89] The methodological quality of the studies examining construct validity was "very good." The best construct validity was observed for the 6MWD, which was strongly correlated with the total leg strength measured with a manual muscle test (ρ=0.733). However, the hypothesis testing for construct validity was "indeterminate." The methodological quality of the studies examining convergent, discriminative and face validity were "very good," but the hypothesis testing for construct validity was "indeterminate" (Supplementary Digital Material 6). The methodological quality of the study examining the criterion validity (*e.g.*, VO_{2peak}, which was the gold standard and measured during 6MWT, TUG and 10-meter walk/run) were "very good" with the calculation of a correlation coefficient, but the quality of the criteria was considered "negative" according to the COSMIN guidelines (ρ=0.558, ICC<0.70).[83]

The methodological quality of the 4 studies examining test-retest reliability were "very good" for the 6MWT,[83, 85] TUG[86] and ESWT,[81] with excellent ICCs from 0.85 to 0.992 and a "positive"[81, 83, 85, 86] quality criteria rating. Two studies analyzed the responsiveness of the 6MWT with "very good" methodological quality (including calculation of MDC90 and minimal clinically important difference (MCID) with or without SEM), but the quality criteria ratings for hypothesis testing were "indeterminate"[80, 83]. Concerning the ESWT, some feasibility data were available, such as the drop-out rate of the ESWT (73.3% in SMA patients *versus* 0% in healthy controls) or the measurement completion of 100% for the ESWT.[81, 82]

In primary mitochondrial myopathy, one psychometric study had "*very good*" methodological quality for the responsiveness assessment (MCID of the 6MWD = 33.3 m), but the quality criteria rating was "indeterminate"[78].

Concerning studies with several subtypes of NMDs, the methodological quality of the construct validity assessment was found to be "very good*"* in different studies in several NMDs with calculation of correlation coefficients. There was a strong correlation between real-life velocity and lower limb MMT (ρ=0.842)[71] or a weak correlation between velocity in the 10MWT and knee extension isometric maximal voluntary contraction (ρ=0.484 with a "indeterminate" quality criteria rating), and even a very strong correlation between the velocity in the 2MWT and the 6MWT (ρ=0.99, P=0.001).[75] The hypothesis testing for the construct validity was "positive" for the construct validity of the 2MWT compared to the 6MWT with "very good" methodological quality.[45] One study analyzed the test-retest reliability properties of the 2MWT and 6MWT.[77] It had "very good" methodological quality with the calculation of ICCs, and a "positive" criterion quality rating (ICCs of 0.99 in both the 2MWT and 6MWT and ICCs>0.90 in all NMD subgroups).[77] This study also analyzed measurement error for the 2MWT and 6MWT with "very good" methodological quality (calculation of SEM,

MDD95 or definition of Limits of Agreement), and the quality criteria for hypotheses testing was "sufficient"[77]. One study analyzed the responsiveness properties of the spatiotemporal parameters acquired in FSHD and LGMD2 in real life with "adequate" methodological quality according to the COSMIN guidelines.[71]

Discussion

In this systematic review of 46 studies, we present an overview of the properties for the measurement of walking tests and tools used in adults diagnosed with genetic NMDs, evaluated according to the COSMIN checklist.

There was a particular focus on three subtypes of NMD, namely myotonic dystrophies (DM1 and DM2) (13 studies),[54-66, 90] 5q SMA (10 studies)[80-89] and CMT (eight studies).[46-53] We can hypothesize that these diseases have been explored more than the others because of their greater prevalence. This is in accordance with the greater number of therapeutic trials compared to the other NMDs, which has resulted in advanced therapeutic strategies in 5Q SMA for instance.[91] In the present review, only five studies out of 46 were conducted in several NMD subtypes.[45, 71, 75-77] There are probably more studies that assess the psychometric properties of walking tests and tools for a single NMD subtype because of the heterogeneous clinical presentation of NMDs.

Concerning the walking tests and tools studied in terms of measurement properties, our first hypothesis was that they might depend on the subtype of NMD. However, we observed a wide variety of measurement properties between the different studies for the same NMD but also between the different subtypes (Supplementary Table II). The same observation could be made for the use of clinical scales for impairment assessments or of the type of the lower limb physical examination, which were varied and not systematic (*i.e.*, five of 44 studies without physical examination). To illustrate this, we can take the example of the myotonic dystrophies, which had the most articles in this review.[54-66, 90] Six different walking tests were used (*i.e.*, 6MWT, 10MWT with comfortable or rapid walking velocity, 10MW/RT, TUG, Step test and Fo8) with, for the most part, the use of a stopwatch (*i.e.*, only one article with motion analysis system and one article with accelerometer; Supplementary Table I, VI, VII, VIII, IX). In these studies, physical examinations (*i.e.*, of lower limb isometric muscle force or MRC lower limbs) with or without the use of a functional assessment scale (MIRS or SARA) were performed. In therapeutic trials on improving walking abilities in myotonic dystrophies, several walking tests and tools were used, such as the 6MWT,[15, 17, 92] 10MWT at a comfortable velocity[93] or TUG, but there was no real justification for the choice of a specific test. Unfortunately, it is extremely difficult to compare the findings due to the discrepancies between studies. Depending on the subtype of NMD, the objective and the endpoint of a study, evaluation protocols should be more homogeneous and standardized.

Our second hypothesis was that the 6MWT would be the test for which measurement properties were the most studied. Indeed, the 6MWT was used in 26 of the 44 studies in our COSMIN review. Moreover, the 6MWT was the only walking test that was studied in all NMDs. The distance traveled during the 6MWT was the single variable assessed in terms of measurement properties. We found "*very good*" methodological quality and "*positive*" quality criteria rating for the reliability assessment of the 6MWT in several NMDs. Specific measurement error assessments of the 6MWT in the studies on DM1 were "very good" in terms of methodological quality, and "very good" in the methodological quality of validity (face, construct, criterion etc.). The responsiveness properties of the 6MWT were the least studied, with "very good" methodological quality of assessments in primary mitochondrial myopathy, 5q SMA and FSHD, but "inadequate" methodological quality in DM1, DM2, CMT and LOPD, which may limit its use, especially in clinical trials. Initially, the 6MWT was used in cardiorespiratory diseases[94] and then gradually for the study of walking and endurance in neurological diseases including NMDs. According to the recommendations of the American Thoracic Society,[11] the 6MWT consists of walking the greatest possible distance over a period of 6 minutes. The 6MWT seems to be the most evaluated walking test in terms of metrological properties and in clinical trials in NMDs, probably due to the muscle fatigue generated, cardiorespiratory and metabolic disabilities in this assessment of aerobic and exercise capacities, making it possible to better understand the impact of the disease on the walking abilities of people with NMDs. However, the length of this test and the generated fatigue could limit the ability of some participants to complete the assessment,[95] hence the need to develop walking tests that can be completed by all people with NMDs.

Furthermore, the same findings were obtained when we looked at walking tests other than the 6MWT, *i.e.*, many studies evaluating reliability and validity measurement properties had "very good" methodological qualities. However, concerning reliability, only few studies examined inter or intra-rater reliability compared to the

number of test–retest reliability evaluations. In clinical practice, these aspects are important because patients are often evaluated by different therapists or physicians over time, so the reliability of a test or tool is essential. Therefore, further studies on inter or intra-rater reliability are necessary. The evaluation of validity was also contrasted because construct validity was the most studied type of validity. For example, it was compared using clinical scales or measurements of muscle strength by physical examination or by isokinetic measures (Supplementary Digital Material 4, 5, 6). As expected, the validity assessment criterion was not very present in our review because it seems to be difficult to find a gold standard for a 10-meter test or even a TUG. Content validity has been tested many times for several of the walking tests and tools in our review with very good results. This is encouraging for the use of these tests in clinical trials and in current practice, though they are not systematically used in the clinical follow-up of people with NMDs.

On the other hand, there were fewer studies focused on the measurement error or responsiveness assessment, which is often associated with inadequate methodological quality according to the COSMIN checklist for responsiveness properties. However, without prior adequate evaluation of a walking test's sensitivity to change and, in particular, the determination of the MCID, it is difficult to conclude on the effectiveness or ineffectiveness of a treatment or of the low sensitivity to change of a walking test or variable in clinical trials (in NMDs or other diseases).[96-98] In this promising new era of therapeutic development in neuromuscular diseases, it seems essential to conduct psychometric studies on the sensitivity to change of walking tests and measurement tools.

Overall, and even if further metrological studies are needed to define relevant outcomes for future clinical trials, the results of this COSMIN review have allowed us to draw up some recommendations for the use of tests and walking tools in the most frequent NMDs.

Concerning CMT disease, the 6MWT and 10MWT had a good inverse correlation for time. The correlation of the 6MWT with CMTNS differed between studies, but the correlation appeared to be greater when the CMTNS was higher.[49, 52] Correlations between lower limb muscle strength and 6MWD and 10MWT time were generally weak (Supplementary Digital Material 4). The second version of the CMTNS (CMTNS2), published by Murphy *et al.* in 2011,[99] had better correlations with spatio-temporal parameters and kinematic data in the 10MWT, especially velocity and knee flexion–extension, with good construct validity properties.[51] Ferrarin *et al.* found excellent reliability for spatio-temporal parameters and kinematic data during motion analysis and Lencioni *et al.* indicated that there was higher responsiveness for kinetic and kinematic measures (according to disease severity) compared with the CMT Examination Score.[47, 50] The 6MWT exhibited excellent reliability but there was inadequate responsiveness data according to the COSMIN guidelines. To the best of our knowledge, and according to this summary, the selection of CMT patients according to disease severity is crucial for selecting the walking test and assessment tool. These results suggested that motion analysis is the most appropriate tool and that the 6MWT is the most appropriate test for walking assessment in CMT, but further feasibility studies for motion analysis and responsiveness studies for the 6MWT are still necessary.

Concerning DM1, several walking tests were studied. Very good construct validity was found for the 6MWT, 10MWT, 10MW/RT, TUG, Step Test and variables assessed by motion analysis in "basic walk" or "dual task walk". High correlations were generally observed between walking tests or motion analysis and lower limb muscle strength/MRC scores (Supplementary Digital Material 5), in contrast to the results in CMT patients. All of these walking tests and tools were reliable, and the studies had very good methodological quality and several types of tested reliability. High responsiveness with AUC>0.7[61] was found for the 10MWT in one study with adequate methodology and sufficient quality criteria rating. For the TUG, one study found AUC<0.7[64] and another >0.7.[62] In DM1, the 10MWT was the walking test that had the most verified metrological properties, which could explain its preferred use in research and clinical practice for walking assessments.

Concerning 5q SMA, several walking tests and tools were studied, including the 6MWT, 10MWT, TUG, ESWT and assessments of spatiotemporal parameters by instrumented footwear and GaitRite. The 6MWT was the most studied test. It was found to have a robust construct validity mainly with high correlations with for lower limb strength, and it was the only test with a study of criterion validity compared to $V0_2$ peak.[83] The calculated MDC90 was 24 meters,[80] and, according to one study, the 6MWT was feasible for 97% of people with 5q SMA.[85] These results could justify the preferential use of the 6MWT to assess walking in research and clinical practice.

Concerning the FSHD, several walking tests and tools were studied, such as the 6MWT, iTUG, motion analysis and the use of wearable magneto-inertial sensors for

real-life walking assessment. All of these walking tests and tools were valid and reliable (Supplementary Digital Material 6). Only the 6MWT and the use of wearable magneto-inertial sensors were studied properly in terms of responsiveness according to the COSMIN guidelines.[68, 71] Their use seems to be preferable. Furthermore, the use of wearable magneto-inertial sensors for walking assessment in real life demonstrated the best construct validity, and the correlations with lower limb MMT were high, and it was a reliable tool with high ICC in inter-session reliability and, according to an early study, it seemed to be a sensitive tool with adequate study methodology according to the COSMIN guidelines.[71] To the best of our knowledge, the metrological properties of this tool were studied among NMDs only in FSHD, which open the door to the assessment of walking capability in these patients. Inertial sensors have the advantage of being used in a patient's everyday environment and was developed in particular due to the COVID-19 pandemic.[100] However, the sample size of the single study was small, limiting the conclusions and suggesting the need for further studies in FSHD. In other NMDs, the lack of metrological studies may limit their use in research and clinical applications.[100] Further studies using these wearable magneto-inertial sensors will be needed for the development of novel walking outcomes in real-life environments.

Strengths and limitations of the study

To the best of our knowledge, this is the first COSMIN review on the measurement properties of walking outcome assessments in adults with genetic or inherited NMDs. Considering the functional and locomotor deficits caused by NMDs, and the growing number of therapeutic trials in NMD patients, it seemed relevant to synthesize the current data concerning the measurements properties of walking tests and tools in order to identify those that might be more relevant for specific research projects or for clinical follow-up.

Herein, we reviewed data from a large number of studies, even though the inclusion criteria were limited to genetic neuromuscular diseases in adults. Consequently, this review provides a synthesis and new insights concerning the measurement properties of walking tests and tools in adults with NMDs. The major strength of this review was the use of the COSMIN checklist and of a quality assessment for the statistical outcomes. The COSMIN checklist was used to assess the quality of the methodological design and statistical methods of the psychometric studies. Moreover, this checklist has been successfully used in previous systematic reviews for the evaluation of walking disorders in other diseases.[39-41] The level of agreement between the two reviewers (NH and MG) before discussion was moderate (κ=0.59) for this COSMIN review. This was in accordance with the agreement level of previous studies using COSMIN guidelines for the assessment of walking psychometric studies in other diseases.[39, 101]

Despite the many strengths of this review, is also presents some limitations. Firstly, it was not initially possible to conduct a relevant meta-analysis because of the high level of heterogeneity among the studies (metric properties, physical examination protocols and type of walking tests and related variables within same and different NMD).

Secondly, most of the studies included in this review had small sample sizes, and often with no justification, which limits the interpretability of our results in relation to the COSMIN checklist.

The last limitation of this paper is related to the selection criteria, which excluded non-hereditary or non-genetic NMDs or pediatric populations. We therefore omitted important NMDs such as Duchenne muscular dystrophy, or children with 5q SMA. We chose these restrictive inclusion criteria in order to obtain the most homogeneous systematic review possible, despite the broad clinical spectrum of NMDs. It is worth noting that there is a recent narrative review of the literature that provides an overview of outcome walking measures in children with NMDs.[102]

Conclusions

To conclude, this review provides an overview of the measurement properties for the walking tests and tools that have been used to evaluate adults with inherited or genetic NMDs.

This information will help researchers and clinicians to choose the most appropriate tests and tools depending on their aim and the studied NMD. There is currently a large panel of potential walking tests and tools to evaluate adults with NMDs, and most were found to be valid and reliable. However, important psychometric studies such as inter- or intra-rater reliability and responsiveness assessments are missing or were found to have inadequate methodological qualities. Future studies are warranted to better clarify these elements due to the current expansion of clinical trials that include medical or physical therapies. This is also of major importance considering the emergence of gene therapies and their potential positive impact on walking disorders in NMDs.

References

1. Bhatt JM. The Epidemiology of Neuromuscular Diseases. Neurol Clin 2016;34:999–1021.

2. Benarroch L, Bonne G, Rivier F, Hamroun D. The 2021 version of the gene table of neuromuscular disorders (nuclear genome). Neuromuscul Disord 2020;30:1008–48.

3. Barreto LC, Oliveira FS, Nunes PS, de França Costa IM, Garcez CA, Goes GM, *et al.* Epidemiologic Study of Charcot-Marie-Tooth Disease: A Systematic Review. Neuroepidemiology 2016;46:157–65.

4. Johnson NE, Butterfield RJ, Mayne K, Newcomb T, Imburgia C, Dunn D, *et al.* Population-Based Prevalence of Myotonic Dystrophy Type 1 Using Genetic Analysis of Statewide Blood Screening Program. Neurology 2021;96:e1045–53.

5. Preston MK, Tawil R, Wang LH. Facioscapulohumeral Muscular Dystrophy. GeneReviews. Seattle, WA: University of Washington; 1993.

6. World Health Organization. IFC: International Classification of Functioning, Disability and Health; 2001 [Internet]. Available from: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [cited 2024, Jan 22].

7. Katirji B. Electrodiagnosis of Neuromuscular Junction Disorders. In: Kaminski HJ, editor. Myasthenia Gravis and Related Disorders. Totowa, NJ: Humana Press; 2003. p. 149–75.

8. Garcia CA. A Clinical Review of Charcot-Marie-Tooth. Ann N Y Acad Sci 1999;883:69–76.

9. McDonald CM. Physical activity, health impairments, and disability in neuromuscular disease. Am J Phys Med Rehabil 2002;81(Suppl):S108–20.

10. Thompson R, Spendiff S, Roos A, Bourque PR, Warman Chardon J, Kirschner J, *et al.* Advances in the diagnosis of inherited neuromuscular diseases and implications for therapy development. Lancet Neurol 2020;19:522–32.

11. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111–7.

12. Lilien C, Gasnier E, Gidaro T, Seferian A, Grelet M, Vissière D, *et al.* Home-Based Monitor for Gait and Activity Analysis. J Vis Exp 2019;150:e59668.

13. Pazzaglia C, Camerota F, Germanotta M, Di Sipio E, Celletti C, Padua L. Efficacy of focal mechanic vibration treatment on balance in Charcot-Marie-Tooth 1A disease: a pilot study. J Neurol 2016;263:1434–41.

14. Ramdharry GM, Pollard AJ, Marsden JF, Reilly MM. Comparing gait performance of people with Charcot-Marie-Tooth disease who do and do not wear ankle foot orthoses. Physiother Res Int 2012;17:191–9.

15. Bassez G, Audureau E, Hogrel JY, Arrouasse R, Baghdoyan S, Bhugaloo H, *et al.* Improved mobility with metformin in patients with myotonic dystrophy type 1: a randomized controlled trial. Brain 2018;141:2855–65.

16. Kierkegaard M, Harms-Ringdahl K, Edström L, Widén Holmqvist L, Tollbäck A. Feasibility and effects of a physical exercise programme in adults with myotonic dystrophy type 1: a randomized controlled pilot study. J Rehabil Med 2011;43:695–702.

17. Kontou E, Papadopoulos C, Papadimas G, Toubekis A, Bogdanis G, Xirou S, *et al.* Effect of exercise training on functional capacity and body composition in myotonic dystrophy type 2 patients. Muscle Nerve 2021;63:477–83.

18. Vielhaber S, Brejova A, Debska-Vielhaber G, Kaufmann J, Feistner H, Schoenfeld MA, *et al.* 24-months results in two adults with Pompe disease on enzyme replacement therapy. Clin Neurol Neurosurg 2011;113:350–7.

19. Angelini C, Semplicini C, Ravaglia S, Moggio M, Comi GP, Musumeci O, *et al.*; Italian Group on GSDII. New motor outcome function measures in evaluation of late-onset Pompe disease before and after enzyme replacement therapy. Muscle Nerve 2012;45:831–4.

20. Andersen G, Prahm KP, Dahlqvist JR, Citirak G, Vissing J. Aerobic training and postexercise protein in facioscapulohumeral muscular dystrophy: RCT study. Neurology 2015;85:396–403.

21. Colson SS, Benchortane M, Tanant V, Faghan JP, Fournier-Mehouas M, Benaïm C, *et al.* Neuromuscular electrical stimulation training: a safe and effective treatment for facioscapulohumeral muscular dystrophy patients. Arch Phys Med Rehabil 2010;91:697–702.

22. Bankolé LC, Millet GY, Temesi J, Bachasson D, Ravelojaona M, Wuyam B, *et al.* Safety and efficacy of a 6-month home-based exercise program in patients with facioscapulohumeral muscular dystrophy: A randomized controlled trial. Medicine (Baltimore) 2016;95:e4497.

23. Berthelsen MP, Husu E, Christensen SB, Prahm KP, Vissing J, Jensen BR. Anti-gravity training improves walking capacity and postural balance in patients with muscular dystrophy. Neuromuscul Disord 2014;24:492–8.

24. Aprile I, Bordieri C, Gilardi A, Lainieri Milazzo M, Russo G, De Santis F, *et al.* Balance and walking involvement in facioscapulohumeral dystrophy: a pilot study on the effects of custom lower limb orthoses. Eur J Phys Rehabil Med 2013;49:169–78.

25. Chisari C, Bertolucci F, Dalise S, Rossi B. Chronic muscle stimulation improves muscle function and reverts the abnormal surface EMG pattern in myotonic dystrophy: a pilot study. J Neuroeng Rehabil 2013;10:94.

26. van de Venis L, van de Warrenburg BP, Weerdesteyn V, van Lith BJ, Geurts AC, Nonnekes J. Improving gait adaptability in patients with hereditary spastic paraplegia (Move-HSP): study protocol for a randomized controlled trial. Trials 2021;22:32.

27. Micallef J, Attarian S, Dubourg O, Gonnaud PM, Hogrel JY, Stojkovic T, *et al.* Effect of ascorbic acid in patients with Charcot-Marie-Tooth disease type 1A: a multicentre, randomised, double-blind, placebo-controlled trial. Lancet Neurol 2009;8:1103–10.

28. Rousseaux M, Launay MJ, Kozlowski O, Daveluy W. Botulinum toxin injection in patients with hereditary spastic paraparesis. Eur J Neurol 2007;14:206–12.

29. Passerieux E, Hayot M, Jaussent A, Carnac G, Gouzi F, Pillard F, *et al.* Effects of vitamin C, vitamin E, zinc gluconate, and selenomethionine supplementation on muscle function and oxidative stress biomarkers in patients with facioscapulohumeral dystrophy: a double-blind randomized controlled clinical trial. Free Radic Biol Med 2015;81:158–69.

30. Menotti F, Laudani L, Damiani A, Mignogna T, Macaluso A. An anterior ankle-foot orthosis improves walking economy in Charcot-Marie-Tooth type 1A patients. Prosthet Orthot Int 2014;38:387–92.

31. Wegener C, Wegener K, Smith R, Schott KH, Burns J. Biomechanical effects of sensorimotor orthoses in adults with Charcot-Marie-Tooth disease. Prosthet Orthot Int 2016;40:436–46.

32. Phillips MF, Robertson Z, Killen B, White B. A pilot study of a crossover trial with randomized use of ankle-foot orthoses for people with Charcot-Marie-tooth disease. Clin Rehabil 2012;26:534–44.

33. Dufek JS, Neumann ES, Hawkins MC, O'Toole B. Functional and dynamic response characteristics of a custom composite ankle foot orthosis for Charcot-Marie-Tooth patients. Gait Posture 2014;39:308–13.

34. Zhang Y, Roxburgh R, Huang L, Parsons J, Davies TC. The effect of hydrotherapy treatment on gait characteristics of hereditary spastic paraparesis patients. Gait Posture 2014;39:1074–9.

35. Marsden J, Stevenson V, McFadden C, Swain I, Taylor P. The effects of functional electrical stimulation on walking in hereditary and spontaneous spastic paraparesis. Neuromodulation 2013;16:256–60, discussion 260.

36. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.

37. Mokkink LB, de Vet HC, Prinsen CA, Patrick DL, Alonso J, Bouter LM, *et al.* COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res 2018;27:1171–9.

38. Prinsen CA, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HC, *et al.* COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147–57.

39. van Bloemendaal M, van de Water AT, van de Port IG. Walking tests for stroke survivors: a systematic review of their measurement properties. Disabil Rehabil 2012;34:2207–21.

40. Zanudin A, Mercer TH, Jagadamma KC, van der Linden ML. Psychometric properties of measures of gait quality and walking performance in young people with Cerebral Palsy: A systematic review. Gait Posture 2017;58:30–40.

41. Anderson DB, Mathieson S, Eyles J, Maher CG, Van Gelder JM, Tomkins-Lane CC, *et al.* Measurement properties of walking outcome measures for neurogenic claudication: a systematic review and meta analysis. Spine J 2019;19:1378–96.

42. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651–7.

43. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.

44. de Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: A Practical Guide; Practical Guides to Biostatistics and Epidemiology. Cambridge: Cambridge University Press; 2011.

45. Andersen LK, Knak KL, Witting N, Vissing J. Two- and 6-minute walk tests assess walking capability equally in neuromuscular diseases. Neurology 2016;86:442–5.

46. Solari A, Laurà M, Salsano E, Radice D, Pareyson D; CMT-TRIAAL Study Group. Reliability of clinical outcome measures in Charcot-Marie-Tooth disease. Neuromuscul Disord 2008;18:19–26.

47. Ferrarin M, Bovi G, Rabuffetti M, Mazzoleni P, Montesano A, Moroni I, *et al.* Reliability of instrumented movement analysis as outcome measure in Charcot-Marie-Tooth disease: results from a multitask locomotor protocol. Gait Posture 2011;34:36–43.

48. Guillebastre B, Calmels P, Rougier P. Effects of muscular deficiency on postural and gait capacities in patients with Charcot-Marie-Tooth disease. J Rehabil Med 2013;45:314–7.

49. Padua L, Pazzaglia C, Pareyson D, Schenone A, Aiello A, Fabrizi GM, *et al.*; CMT-TRIAAL Group. Novel outcome measures for Charcot-Marie-Tooth disease: validation and reliability of the 6-min walk test and StepWatch(™) Activity Monitor and identification of the walking features related to higher quality of life. Eur J Neurol 2016;23:1343–50.

50. Lencioni T, Piscosquito G, Rabuffetti M, Bovi G, Di Sipio E, Diverio M, *et al.* Responsiveness of gait analysis parameters in a cohort of 71 CMT subjects. Neuromuscul Disord 2017;27:1029–37.

51. Coghe G, Pau M, Mamusa E, Pisano C, Corona F, Pilloni G, *et al.* Quantifying gait impairment in individuals affected by Charcot-Marie-Tooth disease: the usefulness of gait profile score and gait variable score. Disabil Rehabil 2020;42:737–42.

52. Mori L, Prada V, Signori A, Pareyson D, Piscosquito G, Padua L, *et al.*; TreSPE Study Group. Outcome measures in the clinical evaluation of ambulatory Charcot-Marie-Tooth 1A subjects. Eur J Phys Rehabil Med 2019;55:47–55.

53. Pazzaglia C, Padua L, Pareyson D, Schenone A, Aiello A, Fabrizi GM, *et al.*; CMT-TRIAAL Group. Are novel outcome measures for Charcot-Marie-Tooth disease sensitive to change? The 6-minute walk test and StepWatch™ Activity Monitor in a 12-month longitudinal study. Neuromuscul Disord 2019;29:310–6.

54. Kierkegaard M, Tollbäck A. Reliability and feasibility of the six minute walk test in subjects with myotonic dystrophy. Neuromuscul Disord 2007;17:943–9.

55. Galli M, Cimolin V, Crugnola V, Priano L, Menegoni F, Trotti C, *et al.* Gait pattern in myotonic dystrophy (Steinert disease): a kinematic, kinetic and EMG evaluation using 3D gait analysis. J Neurol Sci 2012;314:83–7.

56. Kierkegaard M, Petitclerc E, Hébert LJ, Gagnon C. Is one trial enough for repeated testing? Same-day assessments of walking, mobility and fine hand use in people with myotonic dystrophy type 1. Neuromuscul Disord 2017;27:153–8.

57. Jimenez-Moreno AC, Charman SJ, Nikolenko N, Larweh M, Turner C, Gorman G, *et al.* Analyzing walking speeds with ankle and wrist worn accelerometers in a cohort with myotonic dystrophy. Disabil Rehabil 2019;41:2972–8.

58. Knak KL, Sheikh AM, Andersen H, Witting N, Vissing J. Intrarater reliability and validity of outcome measures in myotonic dystrophy type 1. Neurology 2020;94:e2508–20.

59. 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;14:48–54.

60. 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;24:207–15.

61. Solbakken G, Ørstavik K, Hagen T, Dietrichs E, Naerland T. Major involvement of trunk muscles in myotonic dystrophy type 1. Acta Neurol Scand 2016;134:467–73.

62. Kierkegaard M, Petitclerc É, Hébert LJ, Mathieu J, Gagnon C. 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;50:269–77.

63. Jimenez-Moreno AC, Nikolenko N, Kierkegaard M, Blain AP, Newman J, Massey C, *et al.* Analysis of the functional capacity outcome measures for myotonic dystrophy. Ann Clin Transl Neurol 2019;6:1487–97.

64. Knak KL, Sheikh AM, Witting N, Vissing J. Responsiveness of outcome measures in myotonic dystrophy type 1. Ann Clin Transl Neurol 2020;7:1382–91.

65. Radovanović S, Perić S, Savić-Pavićević D, Dobričić V, Pešović J, Kostić V, *et al.* Comparison of temporal and stride characteristics in myotonic dystrophies type 1 and 2 during dual-task walking. Gait Posture 2016;44:194–9.

66. Montagnese F, Rastelli E, Khizanishvili N, Massa R, Stahl K, Schoser B. Validation of Motor Outcome Measures in Myotonic Dystrophy Type 2. Front Neurol 2020;11:306.

67. Iosa M, Mazzà C, Frusciante R, Zok M, Aprile I, Ricci E, *et al.* Mobility assessment of patients with facioscapulohumeral dystrophy. Clin Biomech (Bristol, Avon) 2007;22:1074–82.

68. Eichinger K, Heatwole C, Heininger S, Stinson N, Matichak Stock C, Grosmann C, *et al.*; FSHD Clinical Trials Research Network. Validity of the 6 minute walk test in facioscapulohumeral muscular dystrophy. Muscle Nerve 2017;55:333–7.

69. Huisinga J, Bruetsch A, Mccalley A, Currence M, Herbelin L, Jawdat O, *et al.* An instrumented timed up and go in facioscapulohumeral muscular dystrophy. Muscle Nerve 2018;57:503–6.

70. Statland JM, Karanevich A, Bruetsch A, Huisinga J. A pilot study of the responsiveness of wireless motion analysis in facioscapulohumeral muscular dystrophy. Muscle Nerve 2019;60:590–4.

71. Gidaro T, Gasnier E, Annoussamy M, Vissing J, Attarian S, Mozaffar T, *et al.* Home-based gait analysis as an exploratory endpoint during a multicenter phase 1 trial in limb girdle muscular dystrophy type R2 and facioscapulohumeral muscular dystrophy. Muscle Nerve 2022;65:237–42.

72. Martino G, Ivanenko Y, Serrao M, Ranavolo A, Draicchio F, Casali C, *et al.* Locomotor coordination in patients with Hereditary Spastic Paraplegia. J Electromyogr Kinesiol 2019;45:61–9.

73. Vanherpe P, Fieuws S, D'Hondt A, Bleyenheuft C, Demaerel P, De Bleecker J, *et al.* Late-onset Pompe disease (LOPD) in Belgium: clinical characteristics and outcome measures. Orphanet J Rare Dis 2020;15:83.

74. Claeys KG, D'Hondt A, Fache L, Peers K, Depuydt CE. Six-Minute Walk Distance Is a Useful Outcome Measure to Detect Motor Decline in Treated Late-Onset Pompe Disease Patients. Cells 2022;11:334.

75. Jacques MF, Onambele-Pearson GL, Reeves ND, Stebbings GK, Smith J, Morse CI. Relationships between muscle size, strength, and physical activity in adults with muscular dystrophy. J Cachexia Sarcopenia Muscle 2018;9:1042–52.

76. Prahm KP, Witting N, Vissing J. Decreased variability of the 6-minute

walk test by heart rate correction in patients with neuromuscular disease. PLoS One 2014;9:e114273.

77. Knak KL, Andersen LK, Witting N, Vissing J. Reliability of the 2- and 6-minute walk tests in neuromuscular diseases. J Rehabil Med 2017;49:362–6.

78. Montano V, Lopriore P, Gruosso F, Carelli V, Comi GP, Filosto M, *et al.* Primary mitochondrial myopathy: 12-month follow-up results of an Italian cohort. J Neurol 2022;269:6555–65.

79. Takeuchi Y, Katsuno M, Banno H, Suzuki K, Kawashima M, Atsuta N, *et al.* Walking capacity evaluated by the 6-minute walk test in spinal and bulbar muscular atrophy. Muscle Nerve 2008;38:964–71.

80. Stolte B, Bois JM, Bolz S, Kizina K, Totzeck A, Schlag M, *et al.* Minimal clinically important differences in functional motor scores in adults with spinal muscular atrophy. Eur J Neurol 2020;27:2586–94.

81. Bartels B, de Groot JF, Habets LE, Wijngaarde CA, Vink W, Stam M, *et al.* Fatigability in spinal muscular atrophy: validity and reliability of endurance shuttle tests. Orphanet J Rare Dis 2020;15:75.

82. Bartels B, Habets LE, Stam M, Wadman RI, Wijngaarde CA, Schoenmakers MA, *et al.* Assessment of fatigability in patients with spinal muscular atrophy: development and content validity of a set of endurance tests. BMC Neurol 2019;19:21.

83. Dunaway Young S, Montes J, Kramer SS, Marra J, Salazar R, Cruz R, *et al.* Six-minute walk test is reliable and valid in spinal muscular atrophy. Muscle Nerve 2016;54:836–42.

84. Rodriguez-Torres R, Fabiano J, Goodwin A, Rao AK, Kinirons S, De Vivo D, *et al.* Neuroanatomical Models of Muscle Strength and Relationship to Ambulatory Function in Spinal Muscular Atrophy. J Neuromuscul Dis 2020;7:459–66.

85. Elsheikh B, King W, Peng J, Swoboda KJ, Reyna SP, LaSalle B, *et al.* Outcome measures in a cohort of ambulatory adults with spinal muscular atrophy. Muscle Nerve 2020;61:187–91.

86. Dunaway S, Montes J, Garber CE, Carr B, Kramer SS, Kamil-Rosenberg S, *et al.* Performance of the timed "up & go" test in spinal muscular atrophy. Muscle Nerve 2014;50:273–7.

87. Montes J, Dunaway S, Garber CE, Chiriboga CA, De Vivo DC, Rao AK. Leg muscle function and fatigue during walking in spinal muscular atrophy type 3. Muscle Nerve 2014;50:34–9.

88. Montes J, McDermott MP, Martens WB, Dunaway S, Glanzman AM, Riley S, *et al.*; Muscle Study Group and the Pediatric Neuromuscular Clinical Research Network. Six-Minute Walk Test demonstrates motor fatigue in spinal muscular atrophy. Neurology 2010;74:833–8.

89. Montes J, Zanotto D, Dunaway Young S, Salazar R, De Vivo DC, Agrawal S. Gait assessment with solesound instrumented footwear in spinal muscular atrophy. Muscle Nerve 2017;56:230–6.

90. Hammarén E. Reliability of Static and Dynamic Balance Tests in Subjects with Myotonic Dystrophy Type 1. Adv Physiother 2012;14:48.

91. Hwu WL, Muramatsu SI, Chien YH, Byrne BJ. Advanced therapeutic strategy for hereditary neuromuscular diseases. Mol Ther 2022;30:12–3.

92. Heatwole C, Luebbe E, Rosero S, Eichinger K, Martens W, Hilbert J, *et al.* Mexiletine in Myotonic Dystrophy Type 1: A Randomized, Double-Blind, Placebo-Controlled Trial. Neurology 2021;96:e228–40.

93. Roussel MP, Hébert LJ, Duchesne E. Strength-training effectively alleviates skeletal muscle impairments in myotonic dystrophy type 1. Neuromuscul Disord 2020;30:283–93.

94. Butland RJ, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six-, and 12-minute walking tests in respiratory disease. Br Med J (Clin Res Ed) 1982;284:1607–8.

95. Witherspoon JW, Vasavada R, Logaraj RH, Waite M, Collins J, Shieh C, *et al.* Two-minute versus 6-minute walk distances during 6-minute walk test in neuromuscular disease: is the 2-minute walk test an effective alternative to a 6-minute walk test? Eur J Paediatr Neurol 2019;23:165–70.

96. Bohannon RW, Crouch R. Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: a systematic review. J Eval Clin Pract 2017;23:377–81.

97. Lachmann R, Schoser B. The clinical relevance of outcomes used in late-onset Pompe disease: can we do better? Orphanet J Rare Dis 2013;8:160.

98. Bohannon RW, Glenney SS. Minimal clinically important difference for change in comfortable gait speed of adults with pathology: a systematic review. J Eval Clin Pract 2014;20:295–300.

99. Murphy SM, Herrmann DN, McDermott MP, Scherer SS, Shy ME, Reilly MM, *et al.* Reliability of the CMT neuropathy score (second version) in Charcot-Marie-Tooth disease. J Peripher Nerv Syst 2011;16:191–8.

100. Bortolani S, Brusa C, Rolle E, Monforte M, De Arcangelis V, Ricci E, *et al.* Technology outcome measures in neuromuscular disorders: A systematic review. Eur J Neurol 2022;29:1266–78.

101. Himuro N, Abe H, Nishibu H, Seino T, Mori M. Easy-to-use clinical measures of walking ability in children and adolescents with cerebral palsy: a systematic review. Disabil Rehabil 2017;39:957–68.

102. Kennedy RA, Carroll K, McGinley JL, Paterson KL. Walking and weakness in children: a narrative review of gait and functional ambulation in paediatric neuromuscular disease. J Foot Ankle Res 2020;13:10.

103. Mänttäri A, Suni J, Sievänen H, Husu P, Vähä-Ypyä H, Valkeinen H, *et al.* Six-minute walk test: a tool for predicting maximal aerobic power (VO2 max) in healthy adults. Clin Physiol Funct Imaging 2018. [Epub ahead of print]

Conflicts of interest

The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Authors' contributions

Nawale Hadouiri: conception and design, data extraction, writing, revision, approval and submission of the article. Isabelle Fournel: conception and design, writing, revision and approval of the article. Christel Thauvin-Robinet: conception and design, writing, revision and approval of the article: Agnès Jacquin-Piques: writing, revision and approval of the article. Paul Ornetti: conception and design, writing, revision and approval of the article. Mathieu Gueugnon: conception and design, data extraction, writing, revision and approval of the article.

Acknowledgements

The authors are grateful to Suzanne Rankin for proofreading the manuscript.

History

Article first published online: February 1, 2024. - Manuscript accepted: January 19, 2024. - Manuscript revised: December 11, 2023. - Manuscript received: June 16, 2023.

SUPPLEMENTARY DIGITAL MATERIAL 1

The search formulation of this COSMIN review

1) (“gait”[MeSH] OR gait[tiab] OR walk*[tiab] OR “walking”[MeSH] OR walking[tiab] OR ambulation[tiab])

AND

2) ("Neuromuscular Diseases"[Mesh]) *OR* ("Muscular Dystrophy, Duchenne"[Mesh]) *OR* ("Charcot-Marie-Tooth Disease"[Mesh]) *OR* ("Muscular Atrophy, Spinal"[Mesh]) *OR ("Muscular Dystrophies"[Mesh]) OR* ("Muscular Dystrophies, Limb-Girdle"[Mesh]) OR "Distal Myopathies"[Mesh] OR "Myopathies, Structural, Congenital"[Mesh] *OR* ("Myotonic Dystrophy"[Mesh]) *OR* ("Muscular Dystrophy, Facioscapulohumeral"[Mesh]) *OR* ("Glycogen Storage Disease Type II"[Mesh] OR "Glycogen Phosphorylase, Muscle Form"[Mesh] OR "GAA protein, human" [Supplementary Concept] OR "Spastic Paraplegia, Hereditary"[Mesh])

AND

3) ("Validation Study" [Publication Type] OR "Reproducibility of Results"[Mesh] OR "observer variation"[MeSH] OR reliability[tiab] OR validity[tiab] OR ((generaliza*[tiab] OR reliab*[tiab] OR "intraclass correlation"[tiab]) AND coefficient*[tiab]) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tiab] OR precision[title] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR ((interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intrarater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intraexaminer[tiab] OR inter-individual[tiab] OR intraparticipant[tiab]) AND reliab*[tiab]) OR repeatab*[title] OR "interscale correlation"[tiab] OR "interscale correlations"[tiab] OR ((error[tiab] OR errors[tiab]) AND (measure*[tiab] OR correlat*[tiab] OR evaluat*[tiab] OR accuracy[tiab] OR accurate[tiab] OR precision[tiab] OR mean[tiab])) OR "individual variability"[title] OR "interval variability"[title] OR "rate variability"[title] OR "variability analysis"[tiab] OR (uncertainty[title] AND (measurement[title] OR measuring[title])) OR "standard error of measurement"[tiab] OR sensitiv*[title] OR sensitivity[tiab] OR responsive*[title] OR responsiveness[tiab] OR "minimal detectable concentration"[tiab] OR "minimal important change"[tiab] OR "minimal important difference"[tiab] OR "minimally important change"[tiab] OR "minimally important difference"[tiab] OR (minimal*[tiab] AND "detectable change"[tiab]) OR "minimal detectable difference"[tiab] OR "minimally detectable difference"[tiab])

SUPPLEMENTARY DIGITAL MATERIAL 2

Supplementary Table I.—Criteria to rate each measurement property according to Prinsen *et al.*[38] and Terwee *et al.*[43]

Measurement property	**Rating[1]**	**Criteria**
Structural validity	+	CTT: CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.08[2] IRT/Rasch: No violation of unidimensionality[3]: CFI or TLI or comparable measure >0.95 OR RMSEA < 0.06 OR SRMR<0.08 AND no violation of local independence: residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 AND no violation of monotonicity: adequate looking graphs OR item scalability >0.30 AND adequate model fit: IRT: $\chi 2$ >0.01 Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z- standardized values > -2 and <2
	?	CTT: Not all information for « +» reported IRT/Rasch: Model fit not reported
	-	Criteria for « + » not met
Internal consistency	+	At least low evidence[4] for sufficient structural validity[5] AND Cronbach's alpha(s) ≥ 0.70 for each unidimensional scale or subscale[6]
	?	Criteria for "At least low evidence4 for sufficient structural validity[5]" not met
	-	At least low evidence[4] for sufficient structural validity[5] AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale[6]
Reliability	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	-	ICC or weighted Kappa < 0.70
Measurement error	+	SDC or LoA < MIC[5]
	?	MIC not defined
	-	SDC or LoA > MIC[5]
Hypotheses testing for construct validity	+	The result is in accordance with the hypothesis
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis
Cross-cultural validity\measurement invariance	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's R^2 < 0.02)
	?	No multiple group factor analysis OR DIF analysis performed
	-	Important differences between group factors OR DIF was found
Criterion validity	+	Correlation with gold standard ≥ 0.70 OR AUC ≥ 0.70
	?	Not all information for « + » reported
	-	Correlation with gold standard < 0.70 OR AUC < 0.70
Responsiveness	+	The result is in accordance with the hypothesis[7] OR AUC ≥ 0.70
	?	No hypothesis defined
	-	The result is not in accordance with the hypothesis[7] OR AUC < 0.70

AUC, area under the curve ; CFA, confirmatory factor analysis ; CFI, comparative fit index ; CTT, classical test theory ; DIF, differential item functioning ; ICC, intraclass correlation coefficient ; IRT, item response theory ; LoA, limits of agreement ; MIC, minimal important change ; RMSEA, Root Mean Square Error of Approximation ; SEM, Standard Error of Measurement ; SDC, smallest detectable change ; SRMR, Standardized Root Mean Residuals ; TLI, Tucker-Lewis index ; [1] "+" sufficient ; " –" insufficient ; "?" indeterminate ; [2] To rate the quality of the summary score, the factor structures should be equal across studies ; [3] unidimensionality refers to a factor analysis per subscale, while structural validity refers to a factor analysis of a (multidimensional) patient-reported outcome measure ; [4] As defined by grading the evidence according to the GRADE approach ; [5] This evidence may come from different studies ; [6] The criteria 'Cronbach alpha < 0.95' was deleted, as this is relevant in the development phase of a PROM and not when evaluating an existing PROM ; [7] The results of all studies should be taken together and it should then be decided if 75% of the results are in accordance with the hypotheses.

Supplementary Table II.—Characteristics of the selected studies.

Articles	Type of NMD	Study design	Sample size	Age of participants	Healthy group presence (N/No)	Walking capacities test	Type of measures assessment	Walking assessments tools	Physical examination (Yes/No)	Lower limbs physical examination	Impairment scale
Solari *et al.*, 2007[46]	CMT	cross-sectional study monocentric	40	42.4 (12.6)	No	10MWT	10MWT time	Stop watch	Yes	Lower limbs maximal voluntary isometric contraction	ONLS
Ferrarin *et al.*, 2011[47]		cross-sectional study monocentric	20	24.6 (17.1)	No	NA	kinematics, EMG and SPT measures	motion analysis system	Yes	NA	CMTES
Guillebastre *et al.*, 2013[48]		prospective monocentric study	26	50.7 (16.0)	19	12-m distance at a self-selected velocity	SPT measures	GAITRite electronic walkway of 8.3m active area	Yes	dorsal and plantar-flexor muscles with MRC	NA
Padua *et al.*, 2016[49]		prospective multicenter study	168	44.4 (13.7)	No	6MWT and 10MWT	6MWD	Activity Monitoring Watch	Yes	NA	CMTNS
Lencioni *et al.*, 2017[50]		prospective unicentric study	71	31.5 (17.6)	No	NA	kinematics and SPT measures	motion analysis system	Yes	NA	CMTES
Coghe *et al.* 2020[51]		cross sectional study	20	48.9 (15.5)	20	10MWT	SPT and kinematics measures	motion analysis system	Yes	NA	CMTNS (v2)
Mori *et al.*, 2018[52]		multicentric longitudinal prospective study	53	52.1 (11.9)	No	6MWT and 10MWT	6MWD and 10MWT time	Stop watch	Yes	dorsal and plantar-flexor muscles strength with dynamometer	CMTNS
Pazzaglia *et al.*, 2019[53]		longitudinal and prospective multicenter study	149	42.5 (12.5)	No	6MWT, 10MWT and monitoring during 5 days	6MWD, 10MWT time and several outputs during the five days monitoring (as activity index)	Stop watch and accelerometer (StepWatch Activity Monitor)	Yes	NA	CMTNS (v2)
Hammaren *et al.*, 2012[59]	DM1	cross-sectional study monocentric	10	42.7 (10.7)	No	10mWT, Fo8, ST and TUG	10MWT and TUG times (and number of steps of ST and Fo8)	Stop watch	No	NA	NA
Galli *et al.*, 2012[55]		cross-sectional study monocentric	10	1.5 (7.6)	20	NA	kinematics, EMG and SPT measures	motion analysis system	Yes	MRC lower limbs	NA
Hammaren *et al.*, 2014[60]		cross-sectional study monocentric	51	41.3 (9.7)	No	10MWT, TUG and ST	10MWT and TUG times,	Stop watch	Yes	Lower limb isometric muscle force (with a	NA

						number of steps of ST			handheld dynamometer)	
Kierkegaard *et al.*, 2017[56]		prospective unicentric study	70	45(13)	No	10mW/RT, 6MWT and TUG	6MWD; 10mW/RT and TUG times	Stop watch	Yes	NA	MIRS
Solbakken *et al.*, 2016[61]		cross sectional study	38	39 (12.4)	No	6MWT, TUG	6MWD and TUG time	corridor of 20m, stopwatch	Yes	MMT of upper, lower limbs and anterior flexors/back extensors trunk	NA
Kierkegaard *et al.*, 2018[62]		cross-sectional study monocentric	11	52 (range 29–85)	No	TUG	TUG time	Stop watch	Yes	NA	MIRS
Kierkegaard *et al.*, 2007[54]		cross-sectional study monocentric	12 for test retest reliability (i) part and 64 for the feasibility part (ii)	(i) 44 (12) and (ii) 43 (13)	No	6MWT	6MWT distance	Stop watch	Yes	NA	MIRS
Jimenez-Moreno *et al.*, 2019[57]		from observational natural history PHENODM1 study	30	48 (25–72)	14	10MWT, 10mW/RT and 6MWT	Mean acceleration units during the 6MWD, 10MWT and 10mW/RT	Accelerometer	Yes	Lower limb isometric muscle force (with a myometer): ankle dorsiflexors, knee extensors and hip flexors	No
Jimenez-Moreno *et al.*, 2019[63]		from observational natural history PHENODM1 study	213	45.2 (14.5)	No	10MWT, 10mW/RT and 6MWT	6MWD, 10MWT, 10mW/RT times	Stop watch	Yes	QMT	SARA and MIRS
Knak *et al.*, 2020[58]		prospective study and bicentric	78	40 (10)	No	10MWT, TUG	10MWT and TUG times	Stop watch	Yes	hip extensor, knee extensor, ankle plantar and dorsal flexor muscles strength with dynamometer	MIRS
Knak *et al.*, 2020[64]		bicentric longitudinal prospective study	63	41 (10)	No	6MWT, 10MWT, TUG and ST	6MWD, 10MWT and TUG times, number of steps of ST	Stop watch	Yes	Lower limbs muscle strength with dynamometer	MIRS
Radovanovic *et al.*, 2016[65]	DM1 and DM2	prospective monocentric study	37 (20 DM1 and 17 DM2)	38.6 (10.9)	48	NA	SPT measures	GAITRite electronic walkway of 5.5 m active area	Yes	Lower limbs muscle strength with MRC	NA
Montagnese *et al.*, 2020[66]	DM2	monocentric longitudinal prospective study	66	54.8 (12.4)	No	6MWT	6MWD	Stop watch	Yes	MMT lower limbs	NA

Iosa *et al.*, 2007[67]	FSHD	cross-sectional study monocentric	12	40,4 (10.8)	12	NA	kinematics and SPT measures	motion analysis system	Yes	MMT lower limb	FCS
Aprile *et al.*, 2013[24]		cross-sectional study monocentric	16	46.5 (16.4)	16	10mWT, 2MWT and walking on baropodometric plateform	2MWD, 10MWT time and SPT measures	Stop watch and baropodometric plateform	Yes	MMT with MRC of lower limbs	CSS
Eichinger *et al.*, 2017[68]		cross sectional study bicentric	86	49.1 (15.2)	No	6MWT, TUG, 30 foot Go and 10MWT	6MWD and TUG, 30 foot Go, 10MWT times	corridor of 40 m in one center and a corrider of 50m in the other, stopwatch	Yes	MMT of upper and lower limbs	FCS
Huisinga *et al.*, 2018[69]		prospective cohort unicentric study	17	53.7 [32-67]	No	TUG	SPT parameters during TUG	wearable magneto-inertial sensors	Yes	MMT of lower limbs	FCS
Statland *et al.*, 2019[70]		prospective cohort study	10	54 (8.2)	No	iTUG	SPT parameters during TUG	wireless inertial sensors (at each wrist, 1 at the sternum, 1 at the lumbar area, and 1 at each ankle)	Yes	NA	FCS
Gidaro *et al.* 2022[71]	FSHD and LGMD2	exploratory endpoint in multicenter phase 1 trial	10	aged 18 to 75 years	No	NA	SPT parameters in real life conditions	wearable magneto-inertial sensors	Yes	MMT of lower limbs	NA
Martino *et al.* 2019[72]	HSP	cross sectional study	21	48.4 (10.9)	20	15x7 m walkway	kinematics, SPT parameters and EMG measures	motion analysis system and wireless EMG	Yes	NA	SPRS
Claeys *et al.*, 2022[74]	LOPD	cross sectional study	12	51.3 (range 22–67)	12	6MWT, 10MWT and TUG	6MWD, 10MWT and TUG times	Stop watch	Yes	MRC sum score and lower limb isometric muscle force (with Biodex dynamometer)	NA
Vanherpe *et al.*, 2020[73]		retrospective multicentric study	52	47.9 (15.2)	No	6MWT	6MWD	Stop watch	Yes	MRC sum score	NA
Jacques *et al.* 2018[75]	Muscular Dystrophies	cross sectional study	24 (only ambulant people out of 60 included with NMD)	BMD: 42.4 (13.5) LGMD: 43.1 (12.4) FSHD: 47.7 (11.1)	16	10MWT	10MWT time	Stop watch	Yes	Isometric PFMVC and KEMVC force	NA
Prahm *et al.*, 2014[76]	Several NMD	cross sectional study monocentric	16	47.4 (14.4)	12	6MWT	6MWD	Stop watch and pulse watch for HR measurement	Yes	MRC Hip,knee and Anfle (F/E)	NA
Andersen *et al.*, 2016[45]		prospective monocentric study	115	52.6 (22–83)	38	2MWT and 6MWT	2MWD and 6MWD and velocity during 2MWT and at 1st	Stop watch	Yes	Lower limbs muscle strength with MRC	NA

							and last minute of 6MWT				
Knak *et al.*, 2017[77]		monocentric prospective study	93	53 (17)	No	2MWT and 6MWT	2MWD and 6MWD	Stop watch	Yes	MRC lower limbs	NA
Takeuchi *et al.*,2008[79]	SBMA	cross-sectional study monocentric	35	55.8 (11.2)	29	6MWT	6MWT distance	Stop watch	Yes	Limb Norris Score	NA
Montes *et al.*, 2010[88]	SMA	cross-sectional study monocentric	9	22 [4-49]	9	6MWT	SPT measures and 6MWD	Stop watch and GAITRite electronic walkway of 4.6m active area	No	NA	NA
Montes *et al.*, 2014[87]		cross-sectional study monocentric	10	31.2 (9-49)	No	6MWT	6MWD	Stop watch	Yes	Total leg strength measured by MMT	NA
Dunaway *et al.*, 2014[86]		monocentric longitudinal prospective study	15	28.73 (4.17)	No	6MWT, TUG, 10 meter walk/run	6MWD, TUG time, 10 meter walk/run time	Stop watch	Yes	MMT lower limbs	HFMSE
Dunaway *et al.*, 2016[83]		Retrospective study	30	23.7 (16.4)	No	6MWT, TUG, 10 meter walk/run	6MWD, SPT measures, TUG time, 10 meter walk/run time	GAITRite electronic walkway of 4.25m active area for SPT measures and instrumented footwear (SoleSound) and stop watch	Yes	MMT lower limbs	NA
Rodriguez-Torres *et al.*, 2020[84]		sub study from 2 clinical trial studies	23	28.0 (16.5)	No	6MWT	6MWD	Stop watch	Yes	MRC lower limbs	NA
Bartels *et al.*, 2019[82]		Cross sectional moncentric study (Pilot test sample)	4	26.2 (10–37)	No	ESWT	ESWT time	Stop Watch	No	NA	NA
Bartels *et al.*, 2020[81]		monocentric longitudinal prospective study	15	28.4 (12.4)	25	ESWT	ESWT time	Stop Watch	Yes	Lower limbs MRC	NA
Elsheikh *et al.*, 2020[85]		inside a monocentric double-blind, placebo-controlled, cross-over trial	33	37.2 (9.1)	No	6MWT	6MWD	Stop watch	Yes	Lower limbs maximal voluntary isometric contraction	NA
Stolte *et al.*, 2020[80]		monocentric longitudinal prospective study	51	35.8 (12.7)	No	6MWT	6MWD	Stop watch	Yes	NA	HFMSE

Montes *et al.*, 2017[89]		cross-sectional monocentric study	9	27.9 (range 11.0–51.8)	No	6MWT and 10MWT	(i) velocity (ii) initial double support	Instrumented footwear (SoleSound) and GaitRite	No	NA	NA
Montano *et al.*, 2022[78]	Primary mitochondrial myopathy	monocentric longitudinal prospective study	117	NA	No	6MWT and TUG	6MWD and TUG time	Stop watch	No	NA	NA

ALS-FRS-R: amyotrophic lateral sclerosis functional rating scale; BMD: Becker Muscular Dystrophy; CMT: Charcot Marie Tooth; CMTES: CMT Examination Score; CMTNS: CMT Neuropathy Score ; CSS: Clinical Severity Score; DM1: Dystrophy myotonic type 1; DM2: Dystrophy Myotonic type 2;DMD: Duchenne Muscular Dystrophy; FCS: FSHD Clinical Score; FSHD: facioscapulohumeral muscular dystrophy; F/E: Flexion/Extension; Neuropathy Score; CS: Comfortable speed; EMG: Electromyographic; ESWT: Endurance Shuttle Walk Test; HFMSE: Hammersmith Functional Motor Scale Expanded; HR: Heart Rate; HSP: Hereditary Spastic Paraplegia; KEMVC: Knee Extensor Muscular; Voluntary Contraction; LGMD: Limb Girdle Muscle Dystrophy; LOPD: Late onset Pompe; NA: Non Applicable; NMD: Neuromuscular diseases; MMT: Manual Muscle Testing; MIRS: Muscular Impairment Rating Scale; MRC: Medical Research Council Scale; ONLS: Overall Neuropathy Limitations Scale; PFMVC: Plantar Flexor Muscular Voluntary Contraction; SARA: Scale for the Assessment and Rating of Ataxia; SPT: spatiotemporal; SRPS: Spastic Paraplegia Rating Scale; TUG: Times up and Go test; 2MWD: 2-minute walking distance; 2MWT: 2-minute walking Test; 6MWD: 6- minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test; 10MW/RT: 10-meter walk/run test; ST: step test; Fo8,Walking in a Figure of eight.

SUPPLEMENTARY DIGITAL MATERIAL 4

Supplementary Table III.—Measurement properties of the included studies in Charcot Marie Tooth disease: validity.

			Validity					
Articles	Walking test	Type of variables studied	Type of validity	Compared to	Results (95% CI)	COSMIN	Criterion quality and/or criteria rating	Hypotheses testing for construct validity
Padua *et al.*, 2016[49]	6MWT	6MWD	Construct	(i) 10MWT (ii) CMTNS (iii) plantar flexion as detected by a myometer (iv) dorsi flexion as detected by a myometer	(i) ρ = - 0.54 (P<0.001) (ii) ρ = - 0.55 (P<0.001) (iii) ρ = 0.45 (P<0.001) (iv) ρ = 0.49 (P<0.001)	Very good	NA	?
Mori *et al.*, 2018[52]			Construct	(1) CMTNS (2) dorsal and (3) plantar-flexor muscles strength	(1) ρ=0.11 ; P=0.52 (2) ρ=0.32; P=0.051 (3) ρ=0.004; P=0.98 and correlation between 6MWT and 10MWT: ρ=-0.63; P<0,001	Very good	NA	?
Pazzaglia *et al.*, 2019[53]			NA	NA	NA	NA	NA	NA
Mori *et al.*, 2018[52]	10MWT	10MWT time	Construct	(1) CMTNS (2)dorsal and (3) plantar-flexor muscles strength	(1) 0.26; P=0.084 (2) ρ= - 0.50; P=0.0017 (3) ρ=-0.34; P=0.027	Very good	NA	?
Solari *et al.*, 2007[46]			NA	NA	NA	NA	NA	NA
Pazzaglia *et al.*, 2019[53]			NA	NA	NA	NA	NA	NA
Coghe *et al.* 2020[51]		(i) velocity (ii) step width (iii) stride length (iv) DST (v) kinematic data	Construct	CMTNS2	(i) ρ = - 0.783 (P=0.001) (ii) ρ = - 0.248 (P=0.306) (iii) ρ=-0.776 (P<0.001) (iv) ρ=0.523 (P=0.022) (v) ρ from - 0.004 (p=0.986) for Hip Abduction–adduction to 0.832 (P<0.001) for Knee Flexion–extension	Very good	NA	+
Pazzaglia *et al.*, 2019[52]	Monitoring during 5 days	Several outputs during the five days monitoring (as activity index)	NA	NA	NA	NA	NA	NA
Padua *et al.*, 2016[49]	Monitoring during 5 days	Several outputs during the five days monitoring (as activity index)	Construct	(i) 10MWT (ii) CMTNS (iii) plantar flexion as detected by a myometer	(i) ρ from -0.08 to 0.10 (P>0.05) (ii) ρ from -0.31 to 0.07 (the most of correlation	Very good	NA	?

				(iv) dorsi flexion as detected by a myometer	with P>0.05) (iii) ρ from -0.20 to 0.40 (the most of correlation with P>0.05) (iv) ρ from -0.23 to 0.35(the most of correlation with P>0.05)			
Ferrarin *et al.*, 2011[47]	NA (Motion analysis)	(i) CS (1) velocity and (2) double support (3) Mean ankle angle TW– mean ankle angle HW (4) cadence (5)stride length normalized to body height (6) kinematics data (7) kinetics data (ii) RS (1) velocity and (2) double support (3) cadence (4) stride length normalized to body height (5) kinematics data (6) kinetics data	NA	NA	NA	NA	NA	NA
Guillebastre *et al.*, 2013[48]	NA (GAITRite of 8.3m)	(1) Velocity (2) double support	Construct	MRC scores (i) dorsal (ii) plantar-flexor muscles	(1) (i) ρ = 0.49, P<0.05 (ii) ρ = 0.50, P<0.05 (2) (i) ρ = -0.19, P>0.05 (ii) ρ = -0.04, P>0.05	Very good	NA	+
Lencioni *et al.*, 2017[50]	NA (Motion analysis)	(1) Velocity normalized to body height (2) stride length normalized to body height (3) kinematics data (4) kinetics data	NA	NA	NA	NA	NA	NA

CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; HW: Heel Walking: GC: Gait Cycle; ICC: Intercorrelation coefficient; NA: Non Applicable; RS: rapid speed; RSME: Root-mean-square error; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRM: Standardised Response Mean; ST: Step Test; SWT: Swing Time; TW: Toe Walking; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table IV.—Measurement properties of the included studies in Charcot Marie Tooth disease: reliability.

			Reliability			
Articles	Walking test	Type of variables studied	Design	Results (95% CI)	COSMIN*	Quality criteria rating
Padua *et al.*, 2016[49]	6MWT	6MWD	Test-retest	ICCs >0.9	adequate	+
Mori *et al.*, 2018[52]			NA	NA	NA	NA
Pazzaglia *et al.*, 2019[53]			NA	NA	NA	NA
Mori *et al.*, 2018[52]	10MWT	10MWT time	NA	NA	NA	NA
Solari *et al.*, 2007[46]			(i)Inter-rater (ii) Intra-rater	10MWT (i)ICC=0.97, 95% CI [0.88–0.99]; (ii)ICC=0.96, 95% CI [0.87–0.99]	very good	+
Pazzaglia *et al.*, 2019[53]			NA	NA	NA	NA
Coghe *et al.* 2020[51]		(i) velocity (ii) step width (iii) stride length (iv) DST (v) kinematic data	NA	NA	NA	NA
Pazzaglia *et al.*, 2019[53]	Monitoring during 5 days	Several outputs during the five days monitoring (as activity index)	NA	NA	NA	NA
Padua *et al.*, 2016[49]	Monitoring during 5 days	Several outputs during the five days monitoring (as activity index)		ICCs >0.9	NA	NA
Ferrarin *et al.*, 2011[47]	NA (Motion analysis)	(i) CS (1) velocity and (2) double support (3) Mean ankle angle TW– mean ankle angle HW (4) cadence (5)stride length normalized to body height (6) kinematics data (7) kinetics data (ii) RS (1) velocity and (2) double support (3) cadence (4) stride length normalized to body height (5) kinematics data (6) kinetics data	Test-restest (4–6 weeks apart)	(i) (1) ICC=0.95, (2)ICC=0.93 (3) ICC = 0.95 (4) ICC=0.91 (5) ICC=0.96 (6) All variables have ICC > 0.7 except Trunk ROM in sagittal plane (ICC=0.22) and in transverse plane (ICC=0.34) (7) All variables have ICC > 0.8 (ii) (1) ICC=0.78 (2) ICC = 0.72 (3) ICC=0.71 (4) ICC = 0.78 (5) All variables have ICC > 0.7 except Trunk ROM in sagittal plane (ICC=0.27) and in frontal plane (ICC=-0.04) and transverse plane (ICC=0.49)(6) All variables have ICC > 0.8 except hip positive mechanical work (ICC=0.74)	very good	+

Guillebastre *et al.*, 2013[48]	NA (GAITRite of 8.3m)	(1) Velocity (2) double support	NA	NA	NA	NA
Lencioni *et al.*, 2017[50]	NA (Motion analysis)	(1) Velocity normalized to body height (2) stride length normalized to body height (3) kinematics data (4) kinetics data	NA	NA	NA	NA

CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; HW: Heel Walking: GC: Gait Cycle; ICC: Intercorrelation coefficient; NA: Non Applicable; RS: rapid speed; RSME: Root-mean-square error; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRM: Standardised Response Mean; ST: Step Test; SWT: Swing Time; TW: Toe Walking; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table V.—Measurement properties of the included studies in Charcot Marie Tooth disease: measurement error, responsiveness, and feasibility.

			Measurement error			Responsiveness			Feasibility	
Articles	Walking test	Type of variables studied	Results (95% CI)	COSMIN	Quality criteria rating	Results (95% CI)	COSMIN	Quality criteria rating	Outcomes	Results
Padua *et al.*, 2016[49]	6MWT	6MWD	NA	NA	NA	NA	NA	NA		NA
Mori *et al.*, 2018[52]			NA	NA	NA	NA	NA	NA	NA	NA
Pazzaglia *et al.*, 2019[53]			NA	NA	NA	baseline vs. 12 months: 6MWT p=0.71	inadequate	?	NA	NA
Mori *et al.*, 2018[52]	10MWT	10MWT time	NA	NA	NA	NA	NA	NA	NA	NA
Solari *et al.*, 2007[46]			NA	NA	NA	NA	NA	NA	NA	NA
Pazzaglia *et al.*, 2019[53]			NA	NA	NA	baseline vs. 12 months: 10 MWT p=0.21	inadequate	?	NA	NA
Coghe *et al.* 2020[51]		(i) velocity (ii) step width (iii) stride length (iv) DST (v) kinematic data	NA	NA	NA	NA	NA	NA	NA	NA
Pazzaglia *et al.*, 2019[53]	Monitoring during 5 days	Several outputs during the five days monitoring (as activity index)	NA	NA	NA	baseline vs. 12 months: several SAM outputs demonstrated worsening (P<0.05) and CMTNS too (P<0.001)	inadequate	?	NA	NA
Padua *et al.*, 2016[49]	Monitoring during 5 days	Several outputs during the five days monitoring (as activity index)	NA	NA	NA	NA	NA	NA	NA	NA
Ferrarin *et al.*, 2011[47]	NA (Motion analysis)	(i) CS (1) velocity and (2) double support (3) Mean ankle angle TW– mean ankle angle HW (4) cadence (5)stride length normalized to body height (6) kinematics data (7) kinetics data (ii) RS (1) velocity and (2) double support (3) cadence (4) stride length normalized to body height (5) kinematics data (6) kinetics data	(i) (1) SEM=3.67m/s (2) SEM=1.11%GC (3) SEM=2.7° (4) SEM=1.99 steps/min (5) ICC=2.16% (6) SEM from 2.31° to 9.36° according to the kinematics variables (ii) (1) SEM = 10.12 m/s (2) SEM = 2.19%GC (3) SEM=5.77 steps/min (4) SEM=6.09% (5) (6) SEM from 2.54° to 6.03°	very good	?	NA	NA	NA	NA	NA
Guillebastre *et al.*, 2013[48]	NA (GAITRite of 8.3m)	(1) Velocity (2) double support	NA	NA	NA	NA	NA	NA	NA	NA
Lencioni *et al.*, 2017[50]	NA (Motion analysis)	(1) Velocity normalized to body height (2) stride length normalized to body	NA	NA	NA	(1) SRM=-0.55 [%body height/s] for step ascending and -0.43 for step descending (2)	very good	+	NA	NA

	height (3) kinematics data (4) kinetics data		SRM=-0.54 [%body height] for step ascending and -0.56 for step descending (3) from -0.56° to -0.25° according to variables in CS and from -0.53° to -0.25° in step descending (4) for step descending and ascending: from -0.53J to 0.35J

CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; HW: Heel Walking: GC: Gait Cycle; ICC: Intercorrelation coefficient; NA: Non Applicable; RS: rapid speed; RSME: Root-mean-square error; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRM: Standardised Response Mean; ST: Step Test; SWT: Swing Time; TW: Toe Walking; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

SUPPLEMENTARY DIGITAL MATERIAL 5

Supplementary Table VI.—Measurement properties of the included studies in myotonic dystrophies: validity.

				Validity					
Articles	Subtype of DM	Walking test	Type of variables studied	Type of validity	Compared to	Results (95% CI)	COSMIN	Criterion quality and/or criteria rating	Hypotheses testing for construct validity
Kierkegaard *et al.*, 2007[54]	DM1	6MWT	6MWD	NA	NA	NA	NA	NA	NA
Kierkegaard *et al.*, 2017[56]				NA	NA	NA	NA	NA	NA
Solbakken *et al.*, 2016[61]				Construct	(i)MMT of DEmg (ii) PEmg (iii)Tmg	(i) ρ = 0.45, P=0.005 (ii) ρ = 0.62, P<0.001 (iii) ρ = 0.67, P<0.001	Very good	NA	+
Jimenez-Moreno *et al.*, 2019[57]				construct	(i) SARA (ii) Hip flexors QMT (iii) ankle dorsiflexors QMT (iv) knee extensors QMT	(i) ρ=-0.65; (ii) ρ=0.51; (iii) ρ=0.45;(iv) ρ=0.47	Very good	NA	+
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 6MWD	NA	NA	NA	NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	10mWT (CS and RS)	10MWT time	NA	NA	NA	NA	NA	
Hammaren *et al.*, 2014[60]				Construct	Total lower limb muscle strength	ρ=-0.705; P<0.001 (CS) and ρ=-0.665; P<0.001 (RS)	Very good	NA	
Kierkegaard *et al.*, 2017[56]		10MWT (CS)		NA	NA	NA	NA	NA	
Jimenez-Moreno *et al.*, 2019[57]				construct	(i) SARA (ii) Hip flexors QMT (iii) ankle dorsiflexors QMT (iv) knee extensors QMT	(a) 100% (b) (i) ρ=0.65, (ii) ρ=-0.45, (iii) ρ=-0.43, (iv) ρ=-0.36	Very good	NA	
Knak *et al.*, 2020[64]				NA	NA	NA	NA	NA	

Knak *et al.*, 2020[58]				Construct	Muscle strength (hip extension, knee extension, and ankle plantar and dorsal flexion)	ρ = −0.488, 95%CI −0.719 to −0.256 (P<0.0005)	Very good	NA
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10MWT	NA	NA	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]	DM1	10mW/RT	10mW/RT time	construct	(i) SARA (ii) Hip flexors QMT (iii) ankle dorsiflexors QMT (iv) knee extensors QMT	(i) ρ=0.55 (ii) ρ=-0.51 (iii) ρ=-0.47 (iv) ρ=-0.32	Very good	NA
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10mW/RT	Construct	muscle strength of (i) knee extensors and (ii) hip flexors	(i) ρ= 0.4, P=0.05 (ii) ρ=0.43, p=0.02	NA	NA
Hammaren *et al.*, 2012[59]	DM1	TUG	TUG time	NA	NA	NA	NA	NA
Hammaren *et al.*, 2014[60]				Construct	Total lower limb muscle strength	ρ=-0.585; P<0.001	Very good	NA
Kierkegaard *et al.*, 2017[56]				NA	NA	NA	NA	NA
Solbakken *et al.*, 2016[61]				Construct	(i)MMT of dorsal extension wrist, dorsal extension ankle, (ii) MMT of Hip flexion, knee flexion, knee extension, shoulder abduction, elbow flexion, elbow extension, (iii) MMT of trunk muscle group	(i) ρ = -0.43, P<0.001 (ii) ρ =-0.3, P=0.08 (iii) ρ = -0.58, P<0.001	Very good	NA
Knak *et al.*, 2020[64]				NA	NA	NA	NA	NA
Knak *et al.*, 2020[58]				Construct	Muscle strength (hip extension, knee	ρ = −0.313, 95%CI −0.687 to −0.286 (P<0.0005)	Very good	NA

					extension, and ankle dorsal flexion)			
Kierkegaard *et al.*, 2018[62]				NA	NA	NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	(i)Fo8 and (ii)ST	Fo8 time and number of steps of ST	NA	NA	NA	NA	NA
Hammaren *et al.*, 2014[60]	DM1	ST	number of steps of ST	Construct	Total lower limb muscle strength	ρ=0.610; P<0.001	Very good	NA
Knak *et al.*, 2020[58]				Construct	Muscle strength (ankle plantar and dorsal flexion)	ρ=0.546, P<0.0005 95%IC[0.327–0.764]	Very good	NA
Knak *et al.*, 2020[64]				NA	NA	NA	NA	NA
Galli *et al.*, 2012[55]	DM1	velocity (CS)	NA	Construct	lower limb muscle strength (MRC rectus femoris)	ρ= 0.62, P<0.05	Very good	NA
Radovanovic *et al.*, 2016[65]	(i)DM1 and (ii)DM2	NA (Motion analysis)	CT, SL, SWT	Construct	MRC scores	(i)with CT during basic walk (ρ =-0.658), with CT during dual task (ρ = -0.579) and SL in all gait conditions (ρ from 0.625 to 0.726) (ii) with SL during basic walk (ρ = 0.692), SL during dual task (ρ = 0.627)	Very good	NA
Montagnese *et al.*, 2020[66]	DM2	6MWT	6MWD	Construct	lower limb MMT sum	ρ=0.492 P<0.001	Very good	NA

CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; HW: Heel Walking: GC: Gait Cycle; ICC: Intercorrelation coefficient;NA: Non Applicable; RS: rapid speed; RSME: Root-mean-square error; SD: Standard Deviation; SEM: Standard Error Measurement;SL: Stride Length; SRM: Standardised Response Mean; ST: Step Test; SWT,Swing Time; TW: Toe Walking; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table VII.—Measurement properties of the included studies in myotonic dystrophies: reliability.

				Reliability			
Articles	Subtype of DM	Walking test	Type of variables studied	Design	Results (95% CI)	COSMIN	Quality criteria rating
Kierkegaard *et al.*, 2007[54]	DM1	6MWT	6MWD	Test-retest (one week apart)	ICC=0,99 (95%CI 0.97-1.0)	very good	+
Kierkegaard *et al.*, 2017[56]				Intra-rater (Best trial–average of trials)	ICC=1.00 (0.99-1.00)	very good	+
Solbakken *et al.*, 2016[61]				NA	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]				NA	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 6MWD	intra-accelerometer reliability	ICC= 0.97 (95% CI 0.95–0.99; P<0.001)	very good	+
Hammaren *et al.*, 2012[59]	DM1	10mWT (CS and RS)	10MWT time	Test-retest (1 week)	ICC=0.91(CS) ICC=0.94(RS)	very good	+
Hammaren *et al.*, 2014[60]				NA	NA	NA	NA
Kierkegaard *et al.*, 2017[56]		10MWT (CS)		Intra-rater (Best trial–average of trials)	ICC=0.99 (0.92-1.00)	very good	+
Jimenez-Moreno *et al.*, 2019[57]				Intra-rater (3 measures)	ICC = 0.99, 95% CI [0.99–0.99]	very good	+
Knak *et al.*, 2020[64]				NA	NA	NA	NA
Knak *et al.*, 2020[58]				Intrarater reliability	ICC= 0.96 (0.94–0.98)	very good	
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10MWT	intra-accelerometer reliability	ICC=0.86 (95% CI 0.74–0.93; p=0.003)	very good	+
Jimenez-Moreno *et al.*, 2019[57]	DM1	10mW/RT	10mW/RT time	Intra-rater (3 measures)	ICC=0.99 (0.98–0.99)	adequate	+
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10mW/RT	intra-accelerometer reliability	ICC=0.96 (95% CI 0.93–0.98; P<0.001)	very good	+
Hammaren *et al.*, 2012[59]	DM1	TUG	TUG time	Test-retest (1 week)	ICC=0.83	very good	+
Hammaren *et al.*, 2014[60]				NA	NA	NA	NA
Kierkegaard *et al.*, 2017[56]				Intra-rater (Best trial–average of trials)	TUG=0.98 (0.92-0.99)	very good	+

Solbakken *et al.*, 2016[61]				NA	NA	NA	NA
Knak *et al.*, 2020[64]				NA	NA	NA	NA
Knak *et al.*, 2020[58]				Intrarater reliability (1 to 2 weeks)	ICC= 0.68 (0.54–0.79)	very good	-
Kierkegaard *et al.*, 2018[62]				NA	NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	(i)Fo8 and (ii)ST	Fo8 time and number of steps of ST	Test-retest (1 week)	(i)ICC=0.96, (ii)ICC=0.94	very good	+
Hammaren *et al.*, 2014[60]	DM1	ST	number of steps of ST	NA	NA	NA	NA
Knak *et al.*, 2020[58]				Intrarater reliability (1 to 2 weeks)	0.90 IC95% [0.82–0.94]	very good	+
Knak *et al.*, 2020[64]				NA	NA	NA	NA
Galli *et al.*, 2012[55]	DM1	velocity (CS)	NA	NA	NA	NA	NA
Radovanovic *et al.*, 2016[65]	(i)DM1 and (ii)DM2	NA (Motion analysis)	CT, SL, SWT	NA	NA	NA	NA
Montagnese *et al.*, 2020[66]	DM2	6MWT	6MWD	NA	NA	NA	NA

CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; HW: Heel Walking: GC: Gait Cycle; ICC: Intercorrelation coefficient;NA: Non Applicable; RS: rapid speed; RSME: Root-mean-square error; SD: Standard Deviation; SEM: Standard Error Measurement;SL: Stride Length; SRM: Standardised Response Mean; ST: Step Test; SWT,Swing Time; TW: Toe Walking; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table VIII.—Measurement properties of the included studies in myotonic dystrophies: measurement error.

				Measurement error		
Articles	Subtype of DM	Walking test	Type of variables studied	Results (95% CI)	COSMIN	Quality criteria rating
Kierkegaard *et al.*, 2007[54]	DM1	6MWT	6MWD	SEM= 12m (Better of 2 6MWTs); SEM=11m (Better of 3 6MWTs); SEM=9m (Mean of 3 6MWTs)	very good	?
Kierkegaard *et al.*, 2017[56]				NA	NA	NA
Solbakken *et al.*, 2016[61]				NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]				NA	NA	NA
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 6MWD	NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	10mWT (CS and RS)	10MWT time	(i) SEM=0.6s and ME=1.3(CS) SEM=0.4s and ME=0.7(RS)	very good	?
Hammaren *et al.*, 2014[60]				NA	NA	NA
Kierkegaard *et al.*, 2017[56]		10MWT (CS)		NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]				NA	NA	NA
Knak *et al.*, 2020[64]				NA	NA	NA
Knak *et al.*, 2020[58]				SEM=0.26s	very good	?
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10MWT	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]	DM1	10mW/RT	10mW/RT time	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10mW/RT	NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	TUG	TUG time	SEM=0.7s and ME=1.4	very good	?
Hammaren *et al.*, 2014[60]				NA	NA	NA

Kierkegaard *et al.*, 2017[56]				NA	NA	NA
Solbakken *et al.*, 2016[61]				NA	NA	NA
Knak *et al.*, 2020[64]				NA	NA	NA
Knak *et al.*, 2020[58]				SEM=1.09s	very good	?
Kierkegaard *et al.*, 2018[62]				NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	(i)Fo8 and (ii)ST	Fo8 time and number of steps of ST	(i)SEM=1.7steps and ME=14.4 (ii)SEM=1.5 steps and and ME=2.9	very good	?
Hammaren *et al.*, 2014[60]	DM1	ST	number of steps of ST	NA	NA	NA
Knak *et al.*, 2020[58]				SEM=1.7steps ((±9%)	very good	?
Knak *et al.*, 2020[64]				NA	NA	NA
Galli *et al.*, 2012[55]	DM1	velocity (CS)	NA	NA	NA	NA
Radovanovic *et al.*, 2016[65]	(i)DM1 and (ii)DM2	NA (Motion analysis)	CT, SL, SWT	NA	NA	NA
Montagnese *et al.*, 2020[66]	DM2	6MWT	6MWD	NA	NA	NA

CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; HW: Heel Walking: GC: Gait Cycle; ICC: Intercorrelation coefficient;NA: Non Applicable; RS: rapid speed; RSME: Root-mean-square error; SD: Standard Deviation; SEM: Standard Error Measurement;SL: Stride Length; SRM: Standardised Response Mean; ST: Step Test; SWT,Swing Time; TW: Toe Walking; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table IX.—Measurement properties of the included studies in myotonic dystrophies: responsiveness and feasibility.

				Responsiveness			Feasibility	
Articles	Subtype of DM	Walking test	Type of variables studied	Results (95% CI)	COSMIN	Quality criteria rating	Outcomes	Results
Kierkegaard *et al.*, 2007[54]	DM1	6MWT	6MWD	NA	NA	NA	(i) to be able to perform two 6MWTs on the same day (1 hour apart) (ii) Median Borg RPE-score	(i) 52/64 participants (ii) 13 (6–19)
Kierkegaard *et al.*, 2017[56]				NA	NA	NA	NA	NA
Solbakken *et al.*, 2016[61]				NA	NA	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]				baseline vs. one year visit (1) 1.3s [2.2–0.3] p= 0.009 (2) 0.9s [1.5–0.2] p=0.01 and (3) -36.2 [-19.3– -53.1) P<0.001	inadequate	?	% of participants made 1 trial of 6MWT, 10mWT, 10mW/RT and 30 seconds sit and stand test	100% performed one trial, 80% performed at least a second trial of each test and over 50% completed three trials in those tests required
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 6MWD	NA	NA	NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	10mWT (CS and RS)	10MWT time	NA	NA	NA	NA	NA
Hammaren *et al.*, 2014[60]				NA	NA	NA	NA	NA
Kierkegaard *et al.*, 2017[56]		10MWT (CS)		NA	NA	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]				baseline vs. one year visit: 1.3s [2.2–0.3] p= 0.009	inadequate	?	% of participants made 1 trial of 6MWT, 10mWT, 10mW/RT and 30 seconds sit and stand test	100% performed one trial, 80% performed at least a second trial of each test and over 50% completed three trials in those tests required
Knak *et al.*, 2020[64]				change between baseline and 1 year: −0.009 (−0.15; 0.13) p=0,90; AUC between GRS and 10MWT: AUC >0.7	very good	+	NA	NA
Knak *et al.*, 2020[58]				MCID = 0.69 (13%) with MDD95% = ±0.72 (±12%) [±12%]	very good	?	NA	NA

Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10MWT	NA	NA	NA	NA	NA
Jimenez-Moreno *et al.*, 2019[57]	DM1	10mW/RT	10mW/RT time	0.9s [1.5–0.2] p=0.01	inadequate	?	% of participants made 1 trial of 6MWT, 10mWT, 10mW/RT and 30 seconds sit and stand test	100% performed one trial, 80% performed at least a second trial of each test and over 50% completed three trials in those tests required
Jimenez-Moreno *et al.*, 2019[63]			Mean acceleration units during the 10mW/RT	NA	NA	NA	NA	NA
Hammaren *et al.*, 2012[59]	DM1	TUG	TUG time	NA	NA	NA	NA	NA
Hammaren *et al.*, 2014[60]				NA	NA	NA	NA	NA
Kierkegaard *et al.*, 2017[56]				NA	NA	NA	NA	NA
Solbakken *et al.*, 2016[61]				NA	NA	NA	NA	NA
Knak *et al.*, 2020[64]				change between baseline and 1 year: 0.35 (0.17; 0.53) p=0,0003; AUC between 0.6 and 0.7	very good	-	NA	NA
Knak *et al.*, 2020[58]				MCID = 0.92 (11%) with MDD95% = ±1.26 (±26%) [±19%]	NA	NA	NA	NA
Kierkegaard *et al.*, 2018[62]				Criterion approach AUC walking = 0.8; 95CI% [0.7–0.9]	very good	+	NA	NA
Hammaren *et al.*, 2012[59]	DM1	(i)Fo8 and (ii)ST	Fo8 time and number of steps of ST	NA	NA	NA	NA	NA
Hammaren *et al.*, 2014[60]	DM1	ST	number of steps of ST	NA	NA	NA	NA	NA
Knak *et al.*, 2020[58]				MCID = 2.87 (16%) with MDD95% = ±4.70 (±26%) [±19%]	very good	?	NA	NA
Knak *et al.*, 2020[64]				change between baseline and 1 year: −0.06 (−0.75;	very good	-	NA	NA

				0.63) p=0.88; AUC between GRS and ST: AUC<0.7				
Galli *et al.*, 2012[55]	DM1	velocity (CS)	NA	NA	NA	NA	NA	NA
Radovanovic *et al.*, 2016[65]	(i)DM1 and (ii)DM2	NA (Motion analysis)	CT, SL, SWT	NA	NA	NA	NA	NA
Montagnese *et al.*, 2020[66]	DM2	6MWT	6MWD	decrease of 34,16 m (p=0.003) between visit 1 and 2 (one year)	inadequate	?	NA	NA

CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; HW: Heel Walking: GC: Gait Cycle; ICC: Intercorrelation coefficient; NA: Non Applicable; RS: rapid speed; RSME: Root-mean-square error; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRM: Standardised Response Mean; ST: Step Test; SWT: Swing Time; TW: Toe Walking; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

SUPPLEMENTARY DIGITAL MATERIAL 6

Supplementary Table X.—Measurement properties of the included studies in other NMDs: validity.

				Validity					
Articles	Subtype of NMD	Walking test	Type of variables studied	Type of validity	Compared to	Results (95% CI)	COSMIN	Criterion quality and/or criteria rating	Hypotheses testing for construct validity
Gidaro *et al.* 2022[71]	FSH and LGMD2	NA	(i) Median Velocity (ii) distance walked per hour (iii) 95th centile length (iv) 95th centile velocity	Construct	Lower limb MMT	(i) ρ=0.842; P<0.05 (ii) ρ=0.333; P>0.05 (iii) ρ=0.915; P<0.05 (iv) ρ=0.866; P<0.05	Very good	NA	+
Iosa *et al.*, 2007[67]	FSHD	NA (Motion analysis in CS)	velocity	Construct	FCS	ρ = - 0.51, P=0.09	Very good	NA	+
Aprile *et al.*, 2013[24]		NA (Motion analysis in CS)	(1)Velocity (2) Step length (3)Stride width	Construct	(i)CSS (ii)10MWT time (iii)2MWD	(1) (i) ρ = 0.7, P<0.002 (ii) ρ = 0.8, P<0.0003 (iii)ρ=0.7 P<0.01 (2)(i) ρ=0.7 P<0.003 (ii) ρ=0.9 P<0.00001 (iii)ρ=0.6 P<0.002 (3) (i) ρ = 0.6, P< 0.002 (ii) ρ = 0.7, P<0.005 (iii) ρ = 0.7, P<0.005	Very good	NA	+
Eichinger *et al.*, 2017[68]		6MWT	6MWD	Construct	(i) FCS (ii) MMT LEXT	(i) ρ=−0.57; p <0.0001 (ii) ρ= 0.79; P <0.0001	Very good	NA	+
Huisinga *et al.*, 2018[69]		iTUG	(1) velocity (2) double support	Construct	(i) FCS (ii) MMT LEXT	(1) (i) ρ=-0.65 (p=0.005) (ii) ρ=0.52 (p=0.03)	adequate	NA	+
Statland *et al.*, 2019[70]		TUG	duration, cadence, DS	NA	NA	NA	NA	NA	NA
Martino *et al.* 2019[72]	HSP	15x7 m walkway (CS)	Velocity	Construct	SPRS	ρ= 0.38 (P=0.09)	Very good	NA	?
Claeys *et al.*, 2022[74]	LOPD	6MWT, 10MWT and TUG	6MWD, 10MWT and TUG times	NA	NA	NA	NA	NA	NA

Vanherpe *et al.*, 2020[73]		6MWT	6MWD	NA	NA	NA	NA	NA	NA
Jacques *et al.* 2018[75]	Muscular Dystrophies	10MWT (CS)	Velocity	Construct	knee extension isometric maximal voluntary contraction	ρ = 0.484, P=0.030	Very good	NA	?
Prahm *et al.*, 2014[76]	NMD	6MWT	6MWD	Construct	HR	ρ= 0.731 [0.573–0.886], P<0.001	Very good	NA	+
Andersen *et al.*, 2016[45]		2MWT and 6MWT	2MWT velocity	Criterion	6MWT velocity	ρ = 0.99, p=0.001	Very good	+	+
Knak *et al.*, 2017[77]		(i)2MWT and (ii)6MWT	2MWD and 6MWD	NA	NA	NA	NA	NA	NA
Takeuchi *et al.*,2008[79]	SBMA	6MWT	6MWD	Construct	(i)Limb Norris Score (ii)Norris Bulbar score (iii) and ALSFRS-R	(i) ρ=0.632; P <0.001 (ii) ρ= 0.510; P <0.002 (iii) ρ= 0.557; P <0.001	Very good	NA	+
Montes *et al.*, 2010[88]	SMA	NA (Motion analysis in CS)	(1) Velocity (2) double support (3) support base	Construct	6MWD	1) first ρ= 0.966 P<0.01; last ρ= 0.982 P<0.01 2) first ρ = -0.357 p=0.145; last ρ= -0.293 p=0.238 3) first ρ= -0.602 p=0.008; last ρ= -0.631 p= 0.005	Very good	NA	+
Elsheikh *et al.*, 2020[85]		6MWT	6MWD	(i) Content (i) construct	Lower limbs maximal voluntary isometric contraction	(i)97% of participants realized the 6MWT (i) ρ=0.83 (P<0.0001)	Very good	NA	+
Stolte *et al.*, 2020[80]			6MWD	NA	NA	NA	NA	NA	NA
Rodriguez-Torres *et al.*, 2020[84]			6MWD	Construct	6MWD	Model with hip extensors and abductors, knee flexors, ankle dorsiflexors and plantar flexors strength scores explained 67% of the variability observed in 6MWT (beta = 0.670, P=0.003), correlation between strength of hip abductors and knee flexors strength and 6MWD (ρ = 0.62, P=0.001)	Very good	NA	+

Montes *et al.*, 2014[87]			6MWD	Construct	Total leg strength measured by MMT	ρ=0.733; p=0.016	Very good	NA	+
Dunaway *et al.*, 2016[83]			6MWD	(1)Convergent (2) Discriminative validity of 6MWT between types 3a and 3b (3)Criterion	(1) MMT lower limbs (2) Discriminative validity of 6MWT between types 3a and 3b (3) VO2 peak	(1) ρ=0.676 p=0.002 (2) 6MWT between types 3a and 3b disease severity was established (F=5.707; P=0.024) (3)ρ=0.558; P=0.038	Very good	-	+
Montes *et al.*, 2017[89]		10MWT	(i) velocity (ii) stride length (iii) initial double support assesed by instrumented footwear (SoleSound)	(1) criterion validity	(i)velocity, (ii)stride length and (iii)initial double support assessed with GaitRite	(1) (i) RMSE (SD) = 1.74 (0.83) and ρ=1.00 (ii) RMSE (SD) = 1,83 (0,80) and ρ=0,99 (iii) RMSE (SD) = 0.015 (0.004) and ρ=0.94	Very good	+	NA
Dunaway *et al.*, 2014[86]		TUG	TUG time	Convergent	(i) 10MWT (ii) 6MWT (iii)HFMSE (iv)MMT lower limbs	(i) ρ=0.691; p=0.009 (ii) ρ=-0.514;p=0.050 (iii) ρ=-0.717; p=0.003 (iv) ρ=0.783; p=0.001	Very good	NA	+
Bartels *et al.*, 2019[82]		ESWT	ESWT time	NA	NA	NA	NA	NA	?
Bartels *et al.*, 2020[81]			ESWT time	Convergent validity	MRC knee flexion	MRC knee flexion= - 8.9 after ESWT(P=0.011)	Very good	NA	?
Montano *et al.*, 2022[78]	Primary mitochondrial myopathy	6MWT	6MWD	NA	NA	NA	NA	NA	NA

ALSFRS-R: Revised Amyotrophic Lateral Sclerosis Functional Rating Scale; Borg RPE-score: Borg rating perceived exertion score; CL: Confidence Limit; CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; CS: Comfortable speed; CT: cycle time; DM1: Dystrophy Myotonic Type 1; DM2: Dystrophy Myotonic Type 2; EWT: Endurance Walking Test; FCS: FSHD Clinical Severity Score; HR: Heart rate; HSP: Hereditary Spastic Paraplegia; ICC: Intercorrelation coefficient; LGMD2: Limb Girdle Muscular Dystrophy type 2; MMT LEXT: average lower extremity manual muscle testing score; NA: Non Applicable; NMD: Neuromuscular diseases; RSME: Root-mean-square error; SBMA: Spinal Bulbar and Muscular Amyotrophy; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRPS: Spastic Paraplegia Rating Scale; SWT: swing time; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table XI.—Measurement properties of the included studies in other NMDs: reliability.

				Reliability			
Articles	Subtype of NMD	Walking test	Type of variables studied	Design	Results (95% CI)	COSMIN	Quality criteria rating
Gidaro *et al.* 2022[71]	FSH and LGMD2	NA	(i) Median Velocity (ii) distance walked per hour (iii) 95th centile length (iv) 95th centile velocity	Inter-session (1 month)	ICC > 0.9 for all variables	adequate	+
Iosa *et al.*, 2007[67]	FSHD	NA (Motion analysis in CS)	velocity	NA	NA	NA	NA
Aprile *et al.*, 2013[24]		NA (Motion analysis in CS)	(1)Velocity (2) Step length (3)Stride width	NA	NA	NA	NA
Eichinger *et al.*, 2017[68]		6MWT	6MWD	Test-retest	ICC= 0.99 (lower confidence limit 0.98)	adequate	+
Huisinga *et al.*, 2018[69]		iTUG	(1) velocity (2) double support	Test-retest	(1) ICC=0.99 (2) ICC=0.99	very good	+
Statland *et al.*, 2019[70]		iTUG	TUG duration, cadence, DS	NA	NA	NA	NA
Martino *et al.* 2019[72]	HSP	15x7 m walkway (CS)	Velocity	NA	NA	NA	NA
Claeys *et al.*, 2022[74]	LOPD	6MWT, 10MWT and TUG	6MWD, 10MWT and TUG times	NA	NA	NA	NA
Vanherpe *et al.*, 2020[73]		6MWT	6MWD	NA	NA	NA	NA
Jacques *et al.* 2018[75]	Muscular Dystrophies	10MWT (CS)	Velocity	NA	NA	NA	NA
Prahm *et al.*, 2014[76]	NMD	6MWT	6MWD	NA	NA	NA	NA
Andersen *et al.*, 2016[45]		2MWT and 6MWT	2MWT velocity	NA	NA	NA	NA
Knak *et al.*, 2017[77]		(i)2MWT and (ii)6MWT	2MWD and 6MWD	Test-retest (1-2 weeks)	ICC=0.99, P<0.001, 95% CI [0.98–1.00] for 2MWT and 6MWT	very good	+
Takeuchi *et al.*,2008[79]	SBMA	6MWT	6MWD	Test-retest (1 month)	ICC= 0.982 (P<0.001)	adequate	+

Montes *et al.*, 2010[88]	SMA	NA (Motion analysis in CS)	(1) Velocity (2) double support (3) support base	NA	NA	NA	NA
Elsheikh *et al.*, 2020[85]		6MWT	6MWD	Test-restest (6 weeks apart)	ICC=0.85	adequate	+
Stolte *et al.*, 2020[80]			6MWD	NA	NA	NA	NA
Rodriguez-Torres *et al.*, 2020[84]			6MWD	NA	NA	NA	NA
Montes *et al.*, 2014[87]			6MWD	NA	NA	NA	NA
Dunaway *et al.*, 2016[83]			6MWD	Test-restest (4 weeks apart)	ICC: 0.992; 95% CI, 0.979–0.997	very good	+
Montes *et al.*, 2017[89]		10MWT	(i) velocity (ii) stride length (iii) initial double support assesed by instrumented footwear (SoleSound)	NA	NA	NA	NA
Dunaway *et al.*, 2014[86]		TUG	TUG time	Test-restest (4 weeks apart)	ICC=0.948 95%CI [0.838–0.985]	very good	+
Bartels *et al.*, 2019[82]		ESWT	ESWT time	NA	NA	NA	NA
Bartels *et al.*, 2020[81]			ESWT time	Test-restest (4 weeks apart)	ICC=0.91 95%CI [0.77–0.97]	very good	+
Montano *et al.*, 2022[78]	Primary mitochondrial myopathy	6MWT	6MWD	NA	NA	NA	NA

ALSFRS-R: Revised Amyotrophic Lateral Sclerosis Functional Rating Scale; Borg RPE-score: Borg rating perceived exertion score; CL: Confidence Limit; CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; CS: Comfortable speed; CT: cycle time; DM1: Dystrophy Myotonic Type 1; DM2: Dystrophy Myotonic Type 2; EWT: Endurance Walking Test; FCS: FSHD Clinical Severity Score; HR: Heart rate; HSP: Hereditary Spastic Paraplegia; ICC: Intercorrelation coefficient; LGMD2: Limb Girdle Muscular Dystrophy type 2; MMT LEXT: average lower extremity manual muscle testing score; NA: Non Applicable; NMD: Neuromuscular diseases; RSME: Root-mean-square error; SBMA: Spinal Bulbar and Muscular Amyotrophy; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRPS: Spastic Paraplegia Rating Scale; SWT: swing time; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table XII.—Measurement properties of the included studies in other NMDs: measurement error.

				Measurement error		
Articles	Subtype of NMD	Walking test	Type of variables studied	Results (95% CI)	COSMIN	Quality criteria rating
Gidaro *et al.* 2022[71]	FSH and LGMD2	NA	(i) Median Velocity (ii) distance walked per hour (iii) 95th centile length (iv) 95th centile velocity	NA	NA	NA
Iosa *et al.*, 2007[67]	FSHD	NA (Motion analysis in CS)	velocity	NA	NA	NA
Aprile *et al.*, 2013[24]		NA (Motion analysis in CS)	(1)Velocity (2) Step length (3)Stride width	NA	NA	NA
Eichinger *et al.*, 2017[68]		6MWT	6MWD	NA	NA	NA
Huisinga *et al.*, 2018[69]		iTUG	(1) velocity (2) double support	NA	NA	NA
Statland *et al.*, 2019[70]			TUG duration, cadence, DS	NA	NA	NA
Martino *et al.* 2019[72]	HSP	15x7 m walkway (CS)	Velocity	NA	NA	NA
Claeys *et al.*, 2022[74]	LOPD	6MWT, 10MWT and TUG	6MWD, 10MWT and TUG times	NA	NA	NA
Vanherpe *et al.*, 2020[73]		6MWT	6MWD	NA	NA	NA
Jacques *et al.* 2018[75]	Muscular Dystrophies	10MWT (CS)	Velocity	NA	NA	NA
Prahm *et al.*, 2014[76]	NMD	6MWT	6MWD	NA	NA	NA
Andersen *et al.*, 2016[45]		2MWT and 6MWT	2MWT velocity	NA	NA	NA
Knak *et al.*, 2017[77]		(i)2MWT and (ii)6MWT	2MWD and 6MWD	(i) SEM= 4.9 m (3.4%); MDD95= 13.7 (9.3%); (ii) SEM=14.0 m (3.4%) MDD95= 38.8 m (9.3%); LoA 95% CI (i) –13.9 to +22.5 m (ii) –40.8 to+63.3 m	very good	?

Takeuchi *et al.*,2008[79]	SBMA	6MWT	6MWD	NA	NA	NA
Montes *et al.*, 2010[88]	SMA	NA (Motion analysis in CS)	(1) Velocity (2) double support (3) support base	NA	NA	NA
Elsheikh *et al.*, 2020[85]		6MWT	6MWD	NA	NA	NA
Stolte *et al.*, 2020[80]			6MWD	SEM=55.5m	very good	?
Rodriguez-Torres *et al.*, 2020[84]			6MWD	NA	NA	NA
Montes *et al.*, 2014[87]			6MWD	NA	NA	NA
Dunaway *et al.*, 2016[83]			6MWD	NA	NA	NA
Montes *et al.*, 2017[89]		10MWT	(i) velocity (ii) stride length (iii) initial double support assesed by instrumented footwear (SoleSound)	NA	NA	NA
Dunaway *et al.*, 2014[86]		TUG	TUG time	NA	NA	NA
Bartels *et al.*, 2019[82]		ESWT	ESWT time	NA	NA	NA
Bartels *et al.*, 2020[81]			ESWT time	NA	NA	NA
Montano *et al.*, 2022[78]	Primary mitochondrial myopathy	6MWT	6MWD	NA	NA	NA

ALSFRS-R: Revised Amyotrophic Lateral Sclerosis Functional Rating Scale; Borg RPE-score: Borg rating perceived exertion score; CL: Confidence Limit; CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; CS: Comfortable speed; CT: cycle time; DM1: Dystrophy Myotonic Type 1; DM2: Dystrophy Myotonic Type 2; EWT: Endurance Walking Test; FCS: FSHD Clinical Severity Score; HR: Heart rate; HSP: Hereditary Spastic Paraplegia; ICC: Intercorrelation coefficient; LGMD2: Limb Girdle Muscular Dystrophy type 2; MMT LEXT: average lower extremity manual muscle testing score; NA: Non Applicable; NMD: Neuromuscular diseases; RSME: Root-mean-square error; SBMA: Spinal Bulbar and Muscular Amyotrophy; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRPS: Spastic Paraplegia Rating Scale; SWT: swing time; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.

Supplementary Table XIII.—Measurement properties of the included studies in other NMDs: responsiveness and feasibility.

				Responsiveness			Feasibility	
Articles	Subtype of NMD	Walking test	Type of variables studied	Results (95% CI)	COSMIN	Quality criteria rating	Outcomes	Results
Gidaro *et al.* 2022[71]	FSH and LGMD2	NA	(i) Median Velocity (ii) distance walked per hour (iii) 95th centile length (iv) 95th centile velocity	(i) change from baseline at 3 months = -2% p=0.02; SRM=0.904; p=0.017 (iii) SRM = 1.254; P=0.025	Adequate	+	NA	NA
Iosa *et al.*, 2007[67]	FSHD	NA (Motion analysis in CS)	velocity	NA	NA	NA	NA	NA
Aprile *et al.*, 2013[24]		NA (Motion analysis in CS)	(1)Velocity (2) Step length (3)Stride width	NA	NA	NA	NA	NA
Eichinger *et al.*, 2017[68]		6MWT	6MWD	MDC95=34.3 m	very good	?	NA	NA
Huisinga *et al.*, 2018[69]		iTUG	(1) velocity (2) double support	NA	NA	NA	NA	NA
Statland *et al.*, 2019[70]			TUG duration, cadence, DS	TUG duration -0.6% 90%CL [-5.2, 4.1], cadence -0.2% 90%CL[-2.0, 1.6] and Double support 1.3% 90%CL[-6.7, 9.1]	inadequate	?	NA	NA
Martino *et al.* 2019[72]	HSP	15x7 m walkway (CS)	Velocity	NA	NA	NA	NA	NA
Claeys *et al.*, 2022[74]	LOPD	6MWT, 10MWT and TUG	6MWD, 10MWT and TUG times	6WMT: decrease of 83,8m at 24 months of follow-up (P<0,003) no difference in 10MWT and TUG (P>0,005)	inadequate	?	NA	NA

Vanherpe *et al.*, 2020[73]		6MWT	6MWD	6MWD: significant decrease over years since onset (P=0.0002)	inadequate	?	NA	NA
Jacques *et al.* 2018[75]	Muscular Dystrophies	10MWT (CS)	Velocity	NA	NA	NA	NA	NA
Prahm *et al.*, 2014[76]	NMD	6MWT	6MWD	NA	NA	NA	NA	NA
Andersen *et al.*, 2016[45]		2MWT and 6MWT	2MWT velocity	NA	NA	NA	NA	NA
Knak *et al.*, 2017[77]		(i)2MWT and (ii)6MWT	2MWD and 6MWD	NA	NA	NA	NA	NA
Takeuchi *et al.*,2008[79]	SBMA	6MWT	6MWD	NA	NA	NA	NA	NA
Montes *et al.*, 2010[88]	SMA	NA (Motion analysis in CS)	(1) Velocity (2) double support (3) support base	NA	NA	NA	NA	NA
Elsheikh *et al.*, 2020[85]		6MWT	6MWD	NA	NA	NA	% of SMA people to be able to realize a 6MWT	97%
Stolte *et al.*, 2020[80]			6MWD	SMA type 2 MCID=71.7m, SMA type 3 MCID=47.8m	very good	+	NA	NA
Rodriguez-Torres *et al.*, 2020[84]			6MWD	NA	NA	NA	NA	NA
Montes *et al.*, 2014[87]			6MWD	NA	NA	NA	NA	NA
Dunaway *et al.*, 2016[83]			6MWD	MDC90=24.0	very good	+	NA	NA
Montes *et al.*, 2017[89]		10MWT	(i) velocity (ii) stride length (iii) initial double support assesed by instrumented footwear (SoleSound)	NA	NA	NA	% of SMA people to be able to walk safely with SoleSound	100%
Dunaway *et al.*, 2014[86]		TUG	TUG time	NA	NA	NA	NA	NA

Bartels *et al.*, 2019[82]		ESWT	ESWT time	NA	NA	NA	(i) Reduced time to limitation (ii) Measurement completion (iii) Comprehensibility (iv) Acceptability	(i) 50% Yes (ii)100% Yes (iii)100% Yes (iv) 9.2 (7.4–10) min and muscle fatigue 7(6-9)
Bartels *et al.*, 2020[81]			ESWT time	NA	NA	NA	(1)Time to limitation in SMA people (Mdn (s)) (2) SMA versus Healthy Controls drop out (%)	(1) 861, 95% CI[218–1200] (2) SMA: 73.3% and Healthy controls: 0%
Montano *et al.*, 2022[78]	Primary mitochondrial myopathy	6MWT	6MWD	MCID = 33.3 m	very good	?	NA	NA

ALSFRS-R: Revised Amyotrophic Lateral Sclerosis Functional Rating Scale; Borg RPE-score: Borg rating perceived exertion score; CL: Confidence Limit; CMT: Charcot Marie Tooth; CMTNS: Charcot-Marie-Tooth Neuropathy Score; CS: Comfortable speed; CT: cycle time; DM1: Dystrophy Myotonic Type 1; DM2: Dystrophy Myotonic Type 2; EWT: Endurance Walking Test; FCS: FSHD Clinical Severity Score; HR: Heart rate; HSP: Hereditary Spastic Paraplegia; ICC: Intercorrelation coefficient; LGMD2: Limb Girdle Muscular Dystrophy type 2; MMT LEXT: average lower extremity manual muscle testing score; NA: Non Applicable; NMD: Neuromuscular diseases; RSME: Root-mean-square error; SBMA: Spinal Bulbar and Muscular Amyotrophy; SD: Standard Deviation; SEM: Standard Error Measurement; SL: Stride Length; SRPS: Spastic Paraplegia Rating Scale; SWT: swing time; 6MWD: 6-minute walking distance; 6MWT: 6-minute walk test; 10MWT: 10-minute walk test.