JRM

Journal of Rehabilitation Medicine

REVIEW ARTICLE

MEASUREMENT PROPERTIES OF AEROBIC CAPACITY MEASURES IN NEUROMUSCULAR DISEASES: A SYSTEMATIC REVIEW

Tim VENEMAN[1], MSC, Fieke Sophia KOOPMAN[1], MD, PHD, Joost DAAMS[2], MSC, Frans NOLLET[1], MD, PHD and Eric Lukas VOORN[1], PHD

From the [1]Amsterdam Universitair Medische Centra (UMC), University of Amsterdam, Department of Rehabilitation Medicine, Amsterdam Movement Sciences and [2]Amsterdam UMC, University of Amsterdam, Medical Library, Meibergdreef 9, Amsterdam, The Netherlands

Objective: To systematically evaluate the measurement properties of aerobic capacity measures in neuromuscular diseases.

Data sources: MEDLINE, EMBASE, SportDiscus and Web of Science Conference Proceedings Citation Index – Science were systematically searched from inception until 30 June 2021.

Study selection and data extraction: Screening, data extraction, risk of bias assessment and quality assessment were performed by 2 independent researchers. Studies were included if they evaluated measurement properties of aerobic capacity measures in adults with neuromuscular diseases. Risk of bias was assessed using the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist. Results were pooled and the quality of the evidence was determined using a modified Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach.

Data synthesis: Nine studies including 187 participants were included in this review. Low quality of evidence was found for sufficient content validity of peak oxygen consumption through maximal exercise testing. Criterion validity of 4 out of 7 different measures to predict peak oxygen consumption was sufficient; however, quality of evidence was low or very low for all measures. No studies were found evaluating reliability or responsiveness.

Conclusion: There was a lack of high-quality studies with sufficiently large sample sizes that evaluated the measurement properties of aerobic capacity measures in neuromuscular diseases.

LAY ABSTRACT

Aerobic capacity (or cardiovascular endurance) is an important outcome measure in exercise intervention studies and pharmacological trials in neuromuscular diseases. To establish the effects of these interventions it is important to use outcome measures with good measurement properties. This means that outcome measures are accurate (valid), repeatable (reliable) and able to detect change over time (responsive). The aim of this study was to review the scientific literature regarding the measurement properties of aerobic capacity measures in neuromuscular diseases. Nine small studies (4–44 participants) reporting on the validity of 8 aerobic capacity measures were found. Five of these measures were judged as valid, but the quality of evidence was low. There were no studies evaluating reliability and responsiveness. Taken together, these results were considered insufficient to make recommendations. High-quality studies, with more participants and a focus on reliability and responsiveness, are required.

Neuromuscular diseases (NMD) encompass over 600 different disorders that affect muscle and nerve function, with varying degrees of severity, rate of progression and prevalence (1). People with NMD often experience muscle weakness, cramps, pain and fatigue. These symptoms can hamper physical activity and an active lifestyle, resulting in low physical fitness (2).

Therefore, improving or maintaining physical fitness, in particular aerobic capacity, is an important component of clinical management in NMD (3). Aerobic capacity is often used as a clinical endpoint in exercise intervention studies and pharmacological trials, and is defined as the ability of the respiratory and cardiovascular system to deliver oxygen to the muscles and to utilize it to generate energy during exercise (4). It reflects the capacity of the aerobic system, which is the main provider of oxygen to the working muscles during exercises lasting longer than 75 s (5). Aerobic capacity is an important health mar-

ker (6, 7) and is strongly associated with functional performance in daily living and independent living at an older age (8–10). Low aerobic capacity is a substantial risk factor for diseases such as cancer (11) and cardiovascular diseases (11, 12) and is one of the most powerful predictors of overall mortality, both in healthy people and patients (13).

In addition to being an important clinical endpoint, aerobic exercise measures are also frequently used to guide intensity prescription for aerobic exercise programmes, in healthy people and chronic diseases, such as NMD (14). To accurately prescribe exercise and to evaluate the effects of interventions in people with NMD, it is important to use aerobic capacity measures with adequate measurement properties, i.e. validity, reliability and responsiveness.

A wide variety of aerobic capacity measures have been reported in NMD studies. The peak oxygen consumption (VO_{2peak}), measured through a maximal effort graded exercise test with respiratory gas exchange measurements, is considered the gold-standard measure to assess maximal aerobic capacity (15). VO_{2peak} is often used as primary outcome in aerobic exercise studies in NMD (16, 17). However, there is debate regarding whether VO_{2peak} reflects true maximal aerobic capacity in individuals with NMD (18, 19), because exercise performance may be determined primarily by the extent of upper or lower extremities muscle weakness. Moreover, assessing VO_{2peak} requires expensive equipment and an extensive exercise protocol until exhaustion, which could lead to overburdening of already weakened muscles (20). Therefore, other aerobic capacity measures are being used that require submaximal exercise testing or that can be assessed without respiratory gas exchange measurements, such as the anaerobic threshold (AT) (21). Alternatively, VO_{2peak} can be predicted through submaximal exercise tests, such as the Åstrand test (22) and field tests, such as the shuttle run test (23).

To our knowledge, an overview of the measurement properties of aerobic capacity measures in NMD is currently missing. Since this concerns a very heterogeneous group, where measurement properties may vary between different (types of) NMD, such an overview is highly needed. Therefore, the aim of this study was to systematically review the scientific literature, based on the following research question: What are the measurement properties of aerobic capacity measures used in individuals with NMD? We aimed to identify the measurement properties of aerobic capacity measures and assessed the quality of the evidence. The outcomes of this review may help to formulate guidelines for the application of aerobic capacity measures in NMD.

METHODS

This systematic review was performed according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (24). The study was registered in the International Prospective Register of Systematic Reviews (PROSPERO) on 10 September 2020 (CRD42020200372).

Study selection: criteria

Studies were considered for inclusion in this systematic review if they: *(i)* evaluated at least one of the measurement properties of aerobic capacity measures, as defined by COSMIN (**Appendix SI**): Reliability, validity and responsiveness (25), *(ii)* examined a study population diagnosed with any type of NMD, *(iii)* included adults (≥ 18 years); and *(iv)* were published in English, German or Dutch. Reviews, single case studies or commentaries were excluded.

Test protocols were considered if: *(i)* the reported goal of the test was to determine the aerobic capacity; and *(ii)* the duration of the exercise test was ≥ 3 min, since beyond this the relative contribution of the aerobic system is $\geq 73\%$) (5). For articles on the content validity of VO_{2peak}, additional inclusion criteria were used. The content validity of VO_{2peak} could be assessed by evaluating if a priori criteria for maximal aerobic exercise were achieved. Studies on content validity were included if they: *(i)* used a maximal exercise test protocol, *(ii)* reported at least 1 a priori maximal aerobic exercise criterion, and *(iii)* evaluated the a priori stated criteria, or presented data by which the criteria could be evaluated. Examples of recommended criteria for achieving maximal aerobic capacity are: *(i)* respiratory gas exchange ratio (RER) ≥ 1.1, *(ii)* heart rate (HR) >90% predicted maximum, *(iii)* patient exhaustion/Borg scale ≥ 9 (range 1–10), ≥ 17 (range 6–20), and *(iv)* a plateau in VO_2 despite increasing workload (26, 27).

Various exercise test protocols were allowed, that could differ in intensity (maximal or submaximal) and in workload (e.g. incremental, even-paced or self-paced) (28). Exercise tests could be laboratory- or field-based. Laboratory-based assessments of aerobic capacity use standardized exercise protocols and equipment in a controlled environment (i.e. a laboratory setting). Field-based assessments of aerobic capacity are performed outside a controlled environment (i.e. outside a laboratory setting) using standardized

protocols without the need for (expensive) laboratory equipment.

Search strategy

A systematic search was performed of the following computerized databases through 10 July 2020 and updated on 30 June 2021: MEDLINE, EMBASE, SPORTDiscus and Web of Science Conference Proceedings Citation Index – Science. The literature search was supplemented by searching for trial protocols through ClinicalTrials.gov, ISRCTN clinical trial registry and the Netherlands Trial Register (NTR). Conceptually, the systematic database searches can be described as follows: ([aerobic capacity] AND [neuromuscular diseases]) OR [relevant trials/studies]. The aerobic capacity concept includes search terms for eligible outcome measures and aerobic exercise tests. The scoping search for trial protocols yielded trial names and registry numbers. These were included in the search strategy in order to identify trial updates or information possibly not retrieved by the main subject search. See **Appendix SII** for full search details. In addition, reference lists and citations of key articles were manually checked for relevant additional studies after the search was performed.

Study selection: procedures

The selection of studies from the database was conducted in 2 stages. During the first stage, titles and abstracts of the retrieved searches were screened on eligibility by 2 raters (TV and EV) independently. As a calibration exercise, the 2 raters compared eligibility assessments after the first 100 abstracts and discussed their choices and considerations. After screening all the titles and abstracts, disagreement between raters was resolved by joint review of the studies to reach consensus, and if consensus was not reached, by a third rater (FK). After consensus was reached on studies meeting the inclusion criteria, or if the decision could not be made based on the title and abstract alone, full reports were obtained. In the second selection stage, full reports were screened on eligibility following the same procedure as used for the selection of titles and abstracts. Authors were contacted if full reports could not be obtained, or if the information in the full report was insufficient to make a decision about eligibility.

Data extraction: procedures

From the included articles 2 independent raters (TV and EV) extracted data in terms of study characteristics and measurement properties. In addition, the 2 raters assessed the risk of bias and the rating of the measurement properties (e.g. for criterion validity; correlation between aerobic capacity measure and criterion measure). As a pilot exercise, the raters' extracted data, risk of bias scores and rating of the measurement properties of the first 2 articles were discussed to make sure that procedures were clear and interpreted correctly by the raters. After the data extraction of all studies was completed, disagreement between raters was resolved by joint review of the studies to reach consensus, and if consensus was not reached, by a third rater (FK). Finally, the available data for each measurement property of each aerobic capacity measure was quantitatively pooled and the overall quality of the evidence for a measurement property was graded.

Data extraction: risk of bias assessment

The risk of bias of included studies was assessed with the COSMIN Risk of Bias checklist for systematic reviews of patient-reported outcome measures (PROMs) (29). This checklist was originally developed for PROMs, but its use is also recommended to evaluate the risk of bias of aerobic fitness outcomes (30, 31). The original checklist was adjusted to its use in the current study; i.e. for the risk of bias assessment of aerobic capacity measures. The sections of the checklist that were used in this review are reported in **Appendices SIII** and **SIV**. Items on the checklist were rated as very good, adequate, doubtful or inadequate. In line with the COSMIN guideline, the lowest rating of any item in a box determined the overall score for methodological quality (32).

Data extraction: rating of measurement properties

Measurement properties reported in the individual studies were rated using the criteria proposed by COSMIN (33). An overview of the criteria by which the measurement properties were rated is shown in **Appendix SI**. Measurement properties were rated "sufficient" (+) or "insufficient" (−).

Data synthesis

First, it was decided if the results of the studies for a particular measurement property could be quantitatively pooled. Results that were consistent across the studies were pooled. If the results were inconsistent, 3 strategies could be used: *(i)* if an explanation could be found for the inconsistent results (e.g. [type of] NMD or clear differences in physical functioning), results could be pooled per subgroup, *(ii)* if no explanation could be found for the inconsistent results, results could not be pooled and the overall quality of the measurement property could be rated as "inconsistent"' without grading the evidence, or *(iii)* the

conclusion could be based on the majority of consistent results, and downgraded for inconsistency. The choice for 1 of these strategies was made based on recommendations by COSMIN (33). Pooled results per measurement property per outcome measure were rated against the same criteria as for individual studies. The overall rating for the pooled result were rated sufficient ($+$) or insufficient ($-$).

Using a modified Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach (34), the quality of the evidence of the pooled results was graded. The following 4 factors were taken into account: *(i)* risk of bias of the included studies, *(ii)* inconsistency (i.e. unexplained inconsistency of results across studies), *(iii)* imprecision (i.e. total sample size of the available studies), and *(iv)* indirectness (i.e. evidence from different populations than the population of interest in the review). Based on these factors, the quality of the evidence was graded as high, moderate, low, or very low (**Appendix SV**). The quality of the evidence refers to the confidence that the pooled rating of a measurement property is trustworthy.

RESULTS

Study selection and characteristics

The database search resulted in 3,663 abstracts, of which 3,447 abstracts were excluded based on screening of title and abstract. Full texts of the remaining 216 articles were screened, of which 207 articles did not meet the inclusion criteria. Most studies were excluded because no measurement properties were reported. The remaining 9 articles were included in this review (**Fig. 1**). There were 6 cross-sectional studies, 2 uncontrolled intervention studies, and 1 randomized controlled trial.

The studies included in this review are summarized in **Table I**. Sample sizes ranged from 4 to 44 participants, and included a diversity of NMD. Maximal aerobic capacity of the study populations were reported as VO_{2peak} in mL/kg/min or L/min. Five studies reported data on the content validity of VO_{2peak} assessed through maximal graded exercise testing (35–39). One of these 5 studies specifically aimed to study the content validity of VO_{2peak} as an outcome measure for aerobic capacity (38). Other studies on content validity aimed to evaluate the effectiveness of aerobic interventions (35, 36, 39), or to determine pathophysiological mechanisms of exercise intolerance (37).

The other 4 studies investigated the criterion validity of aerobic capacity measures (40–43), of which 3 studies reported on more than 1 aerobic capacity measure (40, 41, 43). None of the included studies investigated reliability or responsiveness.

Content validity

The exercise test protocols and rating of measurement properties of the 5 included studies reporting on the content validity of VO_{2peak} are described in **Table II**.
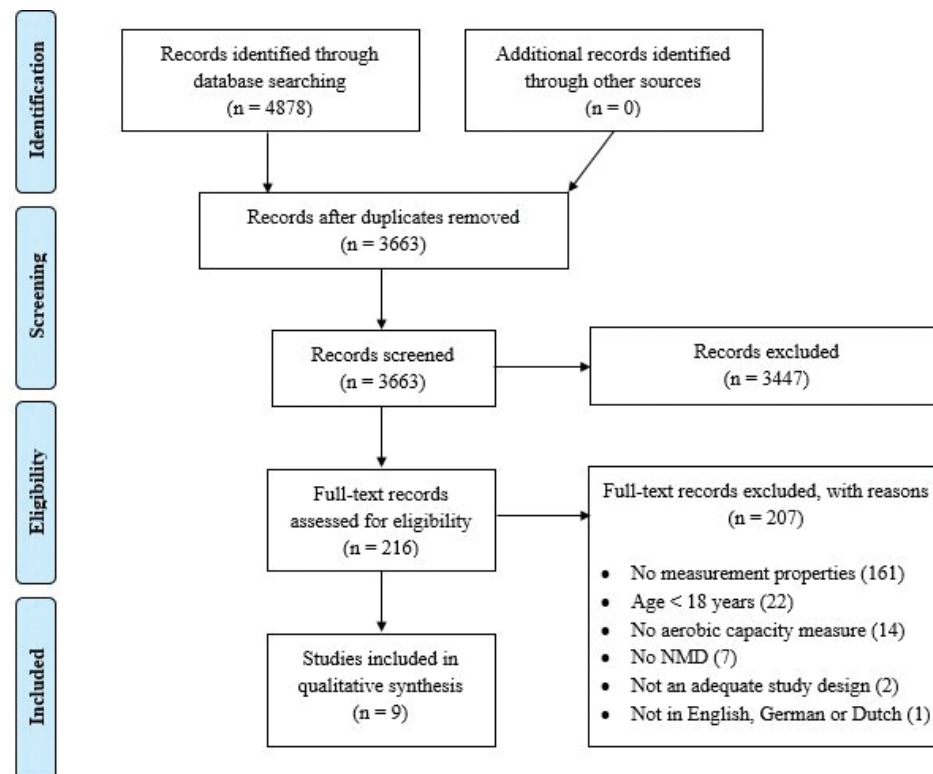

Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart for search outcomes and included studies.

Table I. Study characteristics of the included studies

Author, year of publication	Sample size	Diagnosis	Age, years	Males	VO_{2peak}	Disease characteristics	Measurement property
Al-Rahamneh, 2011 (40)	15	Poliomyelitis	35.0 ± 4.0	8 (53)	21.40 ± 4.50 mL/kg/min	NR	Criterion validity
Cade, 2016 (35)	4	Barth syndrome	23 ± 5	4 (100)	0.73 ± 0.18 L/min	NR	Content validity
Crescimanno, 2015 (41)	8	Pompe disease	49.1 ± 12.6	5 (63)	20.50 [15.10–26.40] mL/kg/min	WGMS: 1.8 ± 0.9	Criterion validity
Gimenes, 2011 (42)	14	Mitochondrial myopathy	35.4 ± 10.8	7 (50)	13.11 ± 3.40 mL/kg/min	NR	Criterion validity
Jones, 1989 (36)	37	Postpoliomyelitis Sequelae	NR	NR	1.53 ± 0.52 L/min	NR	Content validity
McCoy, 2017 (37)	28	Mitochondrial disease	48 ± 9	20 (71)	19.5 ± 5.4 mL/kg/min	NR	Content validity
Montes, 2021 (43)	14	Spinal muscular atrophy	37 [19–56]	9 (64.2)	12.35 [7.90–25.60] mL/kg/min	NR	Criterion validity
Rapin, 2013 (38)	44	Muscular dystrophies, metabolic myopathies and Charcot-Marie-Tooth	43 [21–69]	24 (55)	1.37 L/min (SD not reported)	NR	Content validity
Van den Berg, 2015 (39)	23	Pompe disease	46.0 [20–71]	12 (52)	22.1 ± 7.0 mL/kg/min	QMFT score: 51 ± 8	Content validity

Values are mean ± standard deviation, median [range], or number (percentage).
NR: not reported; WGMS: Walton Gardener Medwin Score; QMFT: Quantitative Motor Function Test.

In the study of McCoy et al. (37), the exercise test was conducted on a bicycle ergometer, with no further protocol description. In the other 4 studies, exercise testing was conducted on a bicycle ergometer with workloads, starting at 0 or 10 Watts and increasing by 5–20 Watts per min, using either ramping increments or graded increments with 1-min stages. Three studies used a fixed workload increment size for all participants, while 2 studies determined the increment size individually based on an estimation of the patients' physical capacities. No information was provided on the level of experience of the test supervisors, except for the study by Rapin et al. (38), in which exercise tests were supervised by an experienced physician. Criteria for the achievement of maximal aerobic exercise were different across studies. RER was used as a criterion in all 5 studies, with threshold values ranging from an RER of 1.0 to 1.15. Furthermore, a plateau in VO_2 during exercise, peak heart rate (HR_{peak}), and the Borg score were used as criteria for achieving maximal aerobic exercise in 2, 3 and 1 studies, respectively.

Scoring of risk of bias assessment items for each study can be found in **Table III**. Risk of bias was rated as doubtful in 4 out of 5 studies (35–37, 39), mainly due to small sample size and lack of reporting on the experience of the test supervisor. Risk of bias in the study of Rapin et al. (38) was scored as adequate. The percentage of participants meeting the a priori criteria for maximal exercise testing ranged between 78% and 100% in 4 out of the 5 studies ($n=92$) and therefore content validity was rated as sufficient in these studies (35–37, 39). In the study of Rapin (38) ($n=44$), the percentage of participants meeting the criteria for maximal exercise testing was 64%, which was rated as insufficient.

Quality of the evidence synthesis for content validity
We decided to pool the study results of all 5 studies on content validity based on the majority of results; 4 of the 5 studies rated the content validity of VO_{2peak} as sufficient with, in total, 113 out of 136 (83%) participants meeting the a priori criteria for maximal exercise testing. The quality of evidence for content validity was downgraded with 1 level for inconsistency based on the insufficient rating of content validity in the study of Rapin (38). Furthermore, the quality of evidence was downgraded 1 point based on the risk of bias assessment (**Table III** and **IV**). Therefore, low-quality evidence was found for a sufficient content validity of VO_{2peak} to measure maximal aerobic capacity in NMD.

Criterion validity
Four studies evaluated the criterion validity of 7 different aerobic capacity measures that aimed to predict VO_{2peak} (40–43). An overview of the study protocols is shown in **Table V**. Scoring of risk of bias for each aerobic capacity measure is shown in **Table VI**. Scoring of the measurement property ratings is shown in **Table V**.

In the study of Al-Rahamneh et al. (40), values of rating of perceived exertion (RPE) and VO_2 during submaximal exercise were used to predict VO_{2peak}: VO_{2peak} was predicted by extrapolating the submaximal RPE and VO_2 values by linear regression to RPE 20. The criterion validity of this method using 3 different submaximal RPE ranges was determined: the predicted VO_{2peak} from RPEs below and including RPE 13, RPE 15 and RPE 17. Risk of bias of the measurement property was scored as "very good". The validity of the predicted VO_{2peak} was rated as insufficient for submaximal RPEs below and including 13 (ICC=0.61), but rated as sufficient for RPEs below and including 15 (ICC=0.78) and 17 (ICC=0.83).

Two studies evaluated the criterion validity of outcomes for the 6MWT to assess aerobic capacity (41, 43). The studies used similar 6MWT protocols (44). In the study of Crescimanno et al. (41),

Table II. Content validity: risk of bias assessment and rating of measurement properties

Author, year of publication	Protocol	Criteria for maximal exercise	Risk of bias assessment	Achievement maximal aerobic exercise	Rating of measurement property
Cade, 2016 (35)	Modality: Bicycle ergometer Warm-up: 1 min unloaded cycling at 60 rpm. Exercise: work rate started at 10 W and increased 10 W/min. Stop: volitional exhaustion.	One of the criteria below: 1) HR_{peak} ≥85% $HR_{peakpred}$ 2) RER_{peak} ≥1.15	Doubtful	4/4 (100%)	Sufficient (+)
Jones, 1989 (36)	Modality: Bicycle ergometer Warm-up: 1 min unloaded cycling at 50–70 rpm. Exercise: work rate increased 20 W/min using a ramping protocol. Stop: volitional exhaustion.	One of the criteria below: 1) VO_2 plateau 2) RER >1.0	Doubtful	37/37 (100%)	Sufficient (+)
McCoy, 2017 (37)	Modality: Bicycle ergometer Warm-up: NR. Exercise: Graded exercise at 60–70 rpm Stop: volitional exhaustion.	1) RER >1.1	Doubtful	22/28 (78%)	Sufficient (+)
Rapin, 2013 (38)	Modality: Bicycle ergometer Warm-up: 1–3 min unloaded cycling at 50–70 rpm. Exercise: Incrementing workload ranging from 5–20 W/min (adapted to the patient's functional capacities according to the examiner's judgment) on ramps or by successive stages, at 50–70 rpm. Stop: volitional exhaustion.	Three of the criteria below: 1) VO_2 plateau 2) RER >1.1 3) Chronotropic reserve 15% lower than $HR_{peakpred}$ 4) Cadence <50 rpm AND Borg score ≥7	Adequate	28/44 (64%)	Insufficient (−)
Van den Berg, 2015 (39)	Modality: Bicycle ergometer Warm-up: 4 min unloaded cycling. Exercise: work rate increased 5–20 W/min (based on the patient's functional capacities) using a ramping protocol. Test duration ranged between 6 and 12 min. Stop: volitional exhaustion.	One of the criteria below: 1) HR >90% $HR_{peakpred}$ 2) RER >1.11 3) VO_2 plateau	Doubtful	22/23 (96%)	Sufficient (+)

rpm: revolutions per min; W: watts; HR_{peak}: peak heart rate; $HR_{peakpred}$: predicted peak heart rate; RER_{peak}: peak respiratory exchange ratio; VO_2: oxygen uptake; RER: respiratory exchange ratio; NR: not reported; HR: heart rate.

the percentage of predicted distance walked in the 6MWT (%DW6MWT) was associated with the percentage of predicted VO_{2peak}. %DW6MWT was determined using regression equations based on sex, height, weight and age (45). The method for determining predicted VO_{2peak} was not reported. Risk of bias was scored as "doubtful", based on several methodological flaws. First, the percentage of predicted VO_{2peak} was used as criterion instead of absolute VO_{2peak} values, without argumentation and without information on the calculation of the predicted VO_{2peak}. Secondly, the %DW6MWT was based on data of healthy individuals. The criterion validity of %DWpredicted was rated as sufficient (rho=0.85). The association between the absolute VO_{2peak} and the absolute distance walked in a 6MWT

was not reported, but from the data available in the article the correlation coefficient was determined. Risk of bias was scored as "very good". The criterion validity of absolute distance walked in a 6MWT was rated as sufficient (rho=0.72).

In the study of Montes et al. (43), the study population consisted of 14 adults and 5 children. The author provided us with their data, which allowed us to assess the correlation coefficient between distance walked in the 6MWT and VO_{2peak} for the adult population only. Risk of bias was scored as "very good". The criterion validity for absolute distance walked in the 6MWT was rated as insufficient (r=0.58). This is in contrast with the sufficient rating of criterion validity in the study of Crescimanno (41). The inconsistent results may be explained by differences between

Table III. Content validity of VO_{2peak} measured during incremental maximal exercise testing: Risk of Bias assessment

Author, year of publication	Test protocol	Maximal effort criteria	Number of participants	Test supervision	Data analysis	Additional flaws	Lowest score
Cade, 2016 (35)	Adequate	Adequate	Doubtful	Doubtful	Very good	Very good	**Doubtful**
Jones, 1989 (36)	Adequate	Doubtful	Adequate	Doubtful	Very good	Very good	**Doubtful**
McCoy, 2017 (37)	Adequate	Adequate	Doubtful	Doubtful	Very good	Very good	**Doubtful**
Rapin, 2013 (38)	Adequate	Very good	Adequate	Very good	Very good	Very good	**Adequate**
Van den Berg, 2015 (39)	Very good	Adequate	Doubtful	Doubtful	Very good	Very good	**Doubtful**

VO_{2peak}: peak oxygen consumption.

Table IV. Content validity: grading the quality of the evidence

Aerobic capacity measure	Risk of bias	Inconsistency	Imprecision	Indirectness	Quality of the evidence
VO_{2peak} during GXT	Serious (−1)	Serious (−1)	NA*	Not downgraded	Low

*Imprecision is not applicable for content validity according to the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) guidelines (33). The sample size of individual studies is already incorporated in the risk of bias checklist for content validity; VO_{2peak}: peak oxygen consumption; GXT: maximal effort graded exercise testing; NA: not applicable.

study cohorts. First, the study of Crescimanno evaluated Pompe disease, while Montes studied Spinal Muscular Atrophy (SMA). Furthermore, the study cohort of Crescimanno had a higher aerobic capacity (median [range], 20.50 [15.10–26.40] mL/kg/min) and greater distance walked in the 6MWT (400 [380–500] m) compared with the cohort of Montes (12.35 [7.90–25.60] mL/kg/min; 354.0 [137.0–557.0] m).

Table V. Criterion validity: quality of study methodology and measurement property

Author	Aerobic capacity measure: Test protocol	Criterion: Test protocol	Risk of bias assessment	Criterion validity	Rating of measurement property
Al-Rahamneh, 2011 (40)	Aerobic capacity measure: Predicted VO_{2peak} by extrapolating sub-maximal RPE (13, 15 or 17) and VO_2 values to RPE 20. Warm-up: 4 min unloaded cycling. Exercise: Incrementing workload of 9 W/min for men, and 6 W/min increment for women Stop: volitional exhaustion.	Aerobic capacity measure: VO_{2peak} measured through GXT. Warm-up: 4 min unloaded cycling. Exercise: Incrementing workload of 9 W/min for men, and 6 W/min increment for women until volitional exhaustion.	Very good	RPE 13: ICC = 0.61 RPE 15: ICC = 0.78 RPE 17: ICC = 0.83	RPE 13: Insufficient (−) RPE 15: Sufficient (+) RPE 17: Sufficient (+)
Crescimanno, 2015 (41)	Aerobic capacity measure: Distance walked at the 6MWT, expressed as percentage of predicted (%DW6MWT). Warm-up: NR Exercise: The 6MWT was performed in accordance with American Thoracic Society Guidelines (44)	Aerobic capacity measure: VO_{2peak}, expressed as percentage of predicted Warm-up: NR Exercise: Symptom-limited treadmill test until exhaustion. No further test details provided.	Doubtful	%DW6MWT: rho = 0.85	Sufficient (+)
Crescimanno, 2015 (41)	Aerobic capacity measure: Distance walked in the 6MWT, expressed as absolute value (DW6MWT). Warm-up: NR Exercise: The 6MWT was performed in accordance with American Thoracic Society Guidelines (44)	Aerobic capacity measure: VO_{2peak}, expressed as absolute value. Warm-up: NR Exercise: Symptom-limited treadmill test until exhaustion. No further test details provided.	Very good	DW6MWT: rho = 0.72	Sufficient (+)
Gimenes, 2011 (42)	Aerobic capacity measure: $\Delta VO2/\Delta WR$ ratio over entire incremental exercise test. Warm-up: NR Exercise: Increasing workload in a linear ramp pattern of 5–15 W/min until volitional exhaustion. Test duration between 8 and 12 min.	Aerobic capacity measure: VO_{2peak} measured through GXT. Warm-up: NR Exercise: Increasing workload in a linear ramp pattern of 5-15 W/min until volitional exhaustion. Test duration between 8 and 12 min.	Very good	$\Delta VO_2/\Delta WR$ r = 0.88	Sufficient (+)
Montes, 2021 (43)	Aerobic capacity measure: Percentage change in workload from first to last min during submaximal bicycle exercise test (Fatigue$_{SME}$) Warm-up: 1 min at 0 Watt. Exercise: 10-min cycling test with workload corresponding to 3–5 on the OMNI scale of perceived exertion (62). Workload was adjusted during the test to maintain the target intensity.	Aerobic capacity measure: VO_{2peak} measured through GXT. Warm-up: 1 min at 0 Watt. Exercise: Incremental graded workload. No further test details provided.	Very good	Fatigue$_{SME}$ r = 0.18	Insufficient (−)
Montes, 2021 (43)	Aerobic capacity measure: Distance walked at the 6MWT, expressed as absolute value (DW6MWT). Warm-up: NR Exercise: The 6MWT was performed in accordance with American Thoracic Society Guidelines (44)	Aerobic capacity measure: VO_{2peak} measured through GXT. Warm-up: 1 min at 0 Watt. Exercise: Incremental graded workload. No further test details provided.	Very good	DW6MWT r = 0.58	Insufficient (−)

VO_{2peak}: peak oxygen consumption; VO_2: oxygen consumption; W: Watts; RPE: rating of perceived exertion; GXT: maximal effort graded exercise testing; ICC: intraclass correlation coefficient; 6MWT: six-min walk test; NR: not reported; %DW6MWT: percentage of predicted distance walked in the 6MWT; DW: distance walked; VO_2: oxygen consumption; WR: workload; Fatigue$_{SME}$: percentage change in workload from first to last min during submaximal bicycle exercise test.

Table VI. Criterion validity: risk of bias assessment

Author, year of publication	Aerobic capacity measure	Gold standard	Data analysis	Additional flaws	Lowest score
Al-Rahamneh, 2011(40)	RPE 13	Very good	Very good	Very good	**Very good**
	RPE 15	Very good	Very good	Very good	**Very good**
	RPE 17	Very good	Very good	Very good	**Very good**
Crescimanno, 2015 (41)	%DW6MWT	Doubtful	Very good	Doubtful	**Doubtful**
	DW6MWT	Very good	Very good	Very good	**Very good**
Gimenes, 2011 (42)	$\Delta VO_2/\Delta WR$ ratio	Very good	Very good	Very good	**Very good**
Montes, 2021 (43)	$Fatigue_{SME}$	Very good	Very good	Very good	**Very good**
	DW6MWT	Very good	Very good	Very good	**Very good**

RPE: Rating of Perceived Exertion; %DW6MWT: percentage of predicted distance walked in the six-min walk test; DW6MWT: distance walked in the six-min walk test; VO_2: oxygen consumption; WR: workload; $Fatigue_{SME}$: percentage change in workload from first to last min during submaximal bicycle exercise test.

Montes (43) also determined the criterion validity of the fatigue during a submaximal exercise test (Fatigue$_{SME}$) as a measure of aerobic capacity. Fatigue$_{SME}$ was determined as the percentage difference between the workload in the first compared with the last min of a 10 min submaximal exercise test. Risk of bias was scored as "very good". The criterion validity was rated as insufficient (r = 0.18).

Finally, the criterion validity of the ratio between the rate of oxygen consumption (ΔVO_2) and the power output (ΔWR) during incremental exercise for predicting maximal aerobic capacity was determined in the study of Gimenes et al. (42). Risk of bias was scored as "very good". The criterion validity was rated as sufficient (r = 0.88).

Quality of the evidence synthesis for criterion validity

Quality of evidence for distance walked in the 6MWT was determined separately for patients with Pompe disease and patients with SMA. These subgroups were made based on the differences in type of NMD and physical capacity between study cohorts that could have explained the inconsistent criterion validity ratings. For each of the aerobic capacity measures, the quality of evidence was determined separately (**Table VII**). The quality of the evidence for criterion validity of all aerobic capacity measures (RPE 13, RPE 15, RPE 17, %DW6MWT, DW6MWT in Pompe disease and in SMA, $\Delta VO2/\Delta WR$ ratio and Fatigue$_{SME}$) was downgraded 2 points for imprecision (i.e. small sample sizes) and %DW6MWT was downgraded an additional 2 points based on risk of bias.

Therefore, low quality of evidence was found for sufficient criterion validity of RPEs below and including RPE15 and RPE17, DW6MWT in Pompe disease and $\Delta VO_2/\Delta WR$ ratio as a measure of maximal aerobic capacity. Low-quality evidence for insufficient criterion validity was found of RPEs below and including RPE 13, DW6MWT in SMA and Fatigue$_{SME}$. Finally, very low quality evidence was found for sufficient criterion validity of %DW6MWT.

DISCUSSION

This review reveals a lack of high-quality studies investigating the measurement properties of aerobic capacity measures in individuals with NMD. A limited number of studies including small sample sizes reported on content and criterion validity. Low quality of evidence was found for sufficient content validity of VO_{2peak} measured through maximal exercise testing. Criterion validity of 7 different measures to predict VO_{2peak}, based on RPE, (percentage of predicted) distance walked in a 6MWT, the ratio between the rate of oxygen consumption and workload assessed during a maximal exercise test and fatigue during a submaximal exercise test, were determined. Four of these aerobic capacity measures were rated as sufficient (RPE 15, RPE 17, %DW6MWT, $\Delta VO2/\Delta WR$ ratio), 2 as insufficient (RPE 13, Fatigue$_{SME}$). DW6MWT was rated sufficient in Pompe disease, but insufficient in SMA. Low (RPE 13, RPE 15, RPE 17, DW6MWT in Pompe and in SMA, $\Delta VO2/\Delta WR$ ratio, Fatigue$_{SME}$) or very low (%DW6MWT) quality of evidence was found for the criterion validity of these measures. No

Table VII. Criterion validity: Grading the quality of the evidence

Aerobic capacity measure	Risk of bias	Inconsistency	Imprecision	Indirectness	Quality of the evidence
RPE 13 (40)	Not downgraded	Not downgraded	Very serious (−2)	Not downgraded	**Low**
RPE 15 (40)	Not downgraded	Not downgraded	Very serious (−2)	Not downgraded	**Low**
RPE 17 (40)	Not downgraded	Not downgraded	Very serious (−2)	Not downgraded	**Low**
%DW6MWT (41)	Very serious (-2)	Not downgraded	Very serious (−2)	Not downgraded	**Very low**
DW6MWT					
Pompe (41)	Not downgraded	Not downgraded	Very serious (−2)	Not downgraded	**Low**
SMA (43)	Not downgraded	Not downgraded	Very serious (−2)	Not downgraded	**Low**
$\Delta VO_2/\Delta WR$ ratio (42)	Not downgraded	Not downgraded	Very serious (−2)	Not downgraded	**Low**
Fatigue$_{SME}$ (43)	Not downgraded	Not downgraded	Very serious (−2)	Not downgraded	**Low**

RPE: Rating of Perceived Exertion; %DW6MWT: percentage of predicted distance walked in the six-min walk test; DW6MWT: distance walked in the six-min walk test; NA: not applicable; VO_2: oxygen consumption; WR: workload; Fatigue$_{SME}$: percentage change in workload from first to last min during submaximal bicycle exercise test.

studies were identified that reported on the reliability or responsiveness of aerobic capacity measures in NMD.

The limited number of studies reporting on measurement properties of aerobic capacity measures contrasts with the large number of studies using it as a clinical endpoint, or to target intensity of aerobic exercise programmes in NMD. During the selection procedure of studies for this review, we observed a large variety of aerobic capacity measures, such as walking distance in a shuttle walking test (46), time until exhaustion during constant workload endurance cycling tests (46, 47), maximal aerobic power during an incremental cycling test (48), and the anaerobic threshold measured through a submaximal exercise test (21). It is surprising that the measurement properties of these aerobic capacity measures have not been assessed. Moreover, all studies included for qualitative analysis in this review focused on the content or criterion validity, and not on the reliability, construct validity or responsiveness. This lack of information on measurement properties of aerobic capacity measures is in contrast with available information for the healthy population (49, 50), but has also been noted in other patient populations, such as multiple sclerosis (51), cerebral palsy (52) and stroke survivors (30).

This review showed that content validity of VO_{2peak}, as assessed through maximal exercise testing, was evaluated in 5 studies and that the content validity of this measure is not sufficiently assured given the low grade of evidence. In contrast to previous hypotheses that achieving maximal aerobic capacity in NMD may be hindered by muscle weakness of the lower or upper extremities (17), the limited evidence suggests that determination of VO_{2peak} through maximal exercise testing seems possible in most individuals with NMD. An important reason for downgrading the level of evidence of the content validity of VO_{2peak} was the quality of a priori criteria for achieving the maximal aerobic capacity. A variety of physiological and/or perceptual criteria for VO_{2peak} were used across studies. In some studies criteria for HR_{peak} (35, 38) and RER_{peak} (35, 36, 39) did not correspond with the criteria for achieving maximal aerobic capacity generally recommended (26). Moreover, the number of criteria *(i)* used and *(ii)* needed to be met to be labelled as a successful maximal exercise test differed between studies. In the study of Rapin et al. (38), 3 of the a priori criteria had to be met, while only 1 criterion was required in the other studies. Although no gold standard exists for the criteria used and the number of criteria considered, it appears to be more likely that the maximal aerobic capacity has been achieved if more of the recommended criteria are considered (26). In the studies where only 1 criterion was considered, or where the criteria did not correspond with the recommended criteria, the true number of participants for whom the maximal aerobic capacity was determined may have been inaccurate. Furthermore, small sample sizes, lack of reporting

on test supervisor experience and limited description of test protocols caused downgrading of the evidence.

Similar to our findings on content validity, we found a limited number of studies reporting on the criterion validity of aerobic capacity measures in NMD, all of which focused on predicting VO_{2peak}. Criterion validity of RPE 13, RPE 15 and $\Delta VO_2/\Delta WR$ in the studies of Al-Rahamneh et al. (40) and Gimenes et al. (42) were rated as sufficient, but the added value of these measures as an alternative for the direct assessment of VO_{2peak} through maximal exercise testing is questionable. The same protocol and equipment required for the direct assessment of VO_{2peak} is needed to determine the $\Delta VO_2/\Delta WR$ ratio in study of Gimenes. The RPE-method used in the study of Al-Rahamneh requires a similar exercise protocol and equipment, with the only difference that the test was stopped at RPE values of 15 or 17. Submaximal exercise testing can be advantageous compared with maximal exercise testing because it can reduce physical strain, but RPE values of 15–17 approach maximal exercise intensities, which limits the potential benefit of this method compared with the direct assessment of VO_{2peak}.

The 6MWT requires no expensive equipment, is easy to conduct and requires submaximal exercise in most cases. Studies in other populations have shown that the distance walked in a 6MWT can be a predictor for VO_{2peak} (53, 54). However, more research is needed to determine the criterion validity of the distance walked at the 6MWT in the NMD population, given the inconsistent results and low quality of evidence in patients with Pompe disease and SMA. The association between distance walked in a 6MWT and VO_{2peak} in NMD may be impacted by walking impairments in individuals with NMD. To our knowledge, a limited number of studies have evaluated the reliability of the distance walked at the 6MWT in NMD (55–57). However, these studies were not included in this review, since the reported goal of the 6MWT in these studies was not to determine aerobic capacity.

Study limitations

Despite the extensive search supplemented with reference checking and searching study registers, it is possible that some studies have not been identified in the current search. It was impossible to include all NMDs in the search terms, because there are approximately 600 NMD, which are often described in multiple ways. Furthermore, we may have missed studies that were not published in English, German or Dutch.

This review followed the COSMIN guidelines for systematic reviews of PROMs (34). However, not all items of the COSMIN risk of bias checklist were applicable for measurement properties of aerobic capacity measures. Therefore, we had to adjust the risk of bias checklist, which can potentially affect its validity.

Lastly, it is important to note that, although not included in this review, the choice for a measurement instrument is based not only on the measurement properties of the instrument, but also on the feasibility and the interpretability of the instrument (33). During the study selection process for this review we came across several studies that reported on the feasibility of aerobic capacity measures (58–60). Feasibility encompasses the practical considerations of using an instrument, including its ease of use, the costs of an instrument, completion time, ease of administration and training for test supervisors (61). The interpretability refers to the degree to which qualitative meaningful results can be assigned to quantitative scores or changes in scores of outcome measures (33). In order to determine if the use of a certain measurement instrument is appropriate, all these factors have to be taken into account.

Implications for clinical practice and research

Low or very low quality of evidence was found for the measurement properties of aerobic capacity measures included in this review. Furthermore, studies that determine the reliability and responsiveness of aerobic capacity measures in NMD are missing. Therefore, we consider the evidence insufficient to recommend the use of certain outcomes.

For future research, we recommend studying: *(i)* the measurement properties of VO_{2peak} measurement through graded maximal exercise testing, and *(ii)* the measurement properties of submaximal aerobic capacity measures without respiratory gas exchange measurements. Respiratory gas exchange measurements require expensive equipment, which is not available in most rehabilitation centres and physiotherapy practices. Therefore, research to the measurement properties of easy to perform submaximal aerobic capacity measures, like the 6MWT, is recommended, given its potential use in clinical practice. To increase the quality of evidence on the content validity of VO_{2peak} as an aerobic capacity measure in NMD, future studies are needed that apply appropriate criteria (26) for maximal aerobic capacity, with large sample sizes (*n*>50) and experienced test supervisors.

Conclusion

This review demonstrates a lack of high-quality studies with sufficiently large sample sizes regarding the measurement properties of aerobic capacity measures in NMD. From the reported aerobic capacity measures, the content validity of VO_{2peak} measured through a maximal effort graded exercise test is the most extensively studied measurement property, but the evidence was insufficient to assure its validity. No studies were found that reported on the reliability or responsiveness of any aerobic capacity measure. More research into the measurement properties of aerobic capacity measures in NMD is warranted to guide the selection of these measures for future clinical trials as well as for clinicians during their daily practice.

REFERENCES

1. Bhatt JM. The epidemiology of neuromuscular diseases. Neurol 2016; 34: 999–1021.
2. McDonald CM. Physical activity, health impairments, and disability in neuromuscular disease. Am J Phys Med Rehabil 2002; 81: 108–120.
3. Abresch RT, Han JJ, Carter GT. Rehabilitation management of neuromuscular disease: the role of exercise training. J Clin Neuromuscul Dis 2009; 11: 7–21.
4. Armstrong N, Welsman JR. Aerobic fitness: what are we measuring? Med Sport Sci 2007; 50: 5–25.
5. Gastin PB. Energy system interaction and relative contribution during maximal exercise. Sports Med 2001; 31: 725–741.
6. Eriksen L, Curtis T, Grønbæk M, Helge JW, Tolstrup JS. The association between physical activity, cardiorespiratory fitness and self-rated health. Prev Med 2013; 57: 900–902.
7. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. Int J Obes (Lond) 2008; 32: 1–11.
8. Arnett SW, Laity JH, Agrawal SK, Cress ME. Aerobic reserve and physical functional performance in older adults. Age Ageing 2008; 37: 384–389.
9. Cress ME, Meyer M. Maximal voluntary and functional performance needed for independence in adults aged 65 to 97 years. Phys Ther 2003; 83: 37–48.
10. Driehuis ER, Van Den Akker LE, De Groot V, Beckerman H. Aerobic capacity explains physical functioning and participation in patients with multiple sclerosis-related fatigue. J Rehabil Med 2018; 50: 185–192.
11. Antero-Jacquemin J, Pohar-Perme M, Rey G, Toussaint J-F, Latouche A. The heart of the matter: years-saved from cardiovascular and cancer deaths in an elite athlete cohort with over a century of follow-up. Eur J Epidemiol 2018; 33: 531–543.
12. Carnethon MR, Gulati M, Greenland P. Prevalence and cardiovascular disease correlates of low cardiorespiratory fitness in adolescents and adults. JAMA 2005; 294: 2981–2988.
13. Blair SN, Kohl HW, Barlow CE, Paffenbarger RS, Gibbons LW, Macera CA. Changes in physical fitness and all-cause mortality: a prospective study of healthy and unhealthy men. JAMA 1995; 273: 1093–1098.
14. American College of Sports Medicine. Resource manual for guidelines for exercise testing and prescription. Baltimore: Lippincott Williams & Wilkins; 2012.
15. Beltz NM, Gibson AL, Janot JM, Kravitz L, Mermier CM, Dalleck LC. Graded exercise testing protocols for the determination of VO2max: historical perspectives, progress, and future considerations. J Sports Med 2016; 2016: 3968393
16. Stefanetti RJ, Blain A, Jimenez-Moreno C, Errington L, Ng YS, McFarland R, et al. Measuring the effects of exercise in neuromuscular disorders: a systematic review and meta-analyses. Wellcome Open Res 2020; 5: 84
17. Oorschot S, Brehm MA, Daams J, Nollet F, Voorn EL. Efficacy of aerobic exercise on aerobic capacity in slowly

progressive neuromuscular diseases: a systematic review and meta-analysis. Ann Phys Rehabil Med 2022; 101637.

18. Willén C, Grimby G. Pain, physical activity, and disability in individuals with late effects of polio. Arch Phys Med Rehabil 1998; 79: 915–919.

19. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. Phys Ther 2000; 80: 782–807.

20. Birk TJ. Poliomyelitis and the post-polio syndrome: exercise capacities and adaptation--current research, future directions, and widespread applicability. Med Sci Sports Exerc 1993; 25: 466–472.

21. Voorn EL, Gerrits KH, Koopman FS, Nollet F, Beelen A. Determining the anaerobic threshold in postpolio syndrome: comparison with current guidelines for training intensity prescription. Arch Phys Med Rehabil 2014; 95: 935–940.

22. Åstrand P-O, Ryhming I. A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. J Appl Physiol 1954; 7: 218-221.

23. Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci 1988; 6: 93–101.

24. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015; 4: 1–9.

25. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010; 63: 737–745.

26. Stickland MK, Butcher SJ, Marciniuk DD, Bhutani M. Assessing exercise limitation using cardiopulmonary exercise testing. Pulm Med 2012; 2012: 824091.

27. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake. Sports Med 2007; 37: 1019–1028.

28. Carvalho VO, Mezzani A. Aerobic exercise training intensity in patients with chronic heart failure: principles of assessment and prescription. Eur J Cardiovasc Prev Rehabil 2011; 18: 5–14.

29. 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: 1171–1179.

30. Wittink H, Verschuren O, Terwee C, De Groot J, Kwakkel G, Van De Port I. Measurement properties of maximal cardiopulmonary exercise test protocols in persons after stroke: a systematic review. J Rehabil Med 2017; 49: 689–699.

31. Balemans AC, Fragala-Pinkham MA, Lennon N, Thorpe D, Boyd RN, O'Neil ME, et al. Systematic review of the clinimetric properties of laboratory-and field-based aerobic and anaerobic fitness measures in children with cerebral palsy. Arch Phys Med Rehabil 2013; 94: 287–301.

32. 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: 651–657.

33. Mokkink LB, Prinsen C, Patrick DL, Alonso J, Bouter LM, De Vet H, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual. 2018; 78.

34. 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: 1147–1157.

35. Cade WT, Reeds DN, Peterson LR, Bohnert KL, Tinius RA, Benni PB, et al. Endurance exercise training in young adults with barth syndrome: a pilot study. JIMD Rep 2016; 32: 15–24.

36. Jones DR, Speier J, Canine K, Owen R, Stull GA. Cardiorespiratory responses to aerobic training by patients with postpoliomyelitis sequelae. JAMA 1989; 261: 3255–3258.

37. McCoy J, Bates M, Eggett C, Siervo M, Cassidy S, Newman J, et al. Pathophysiology of exercise intolerance in chronic diseases: the role of diminished cardiac performance in mitochondrial and heart failure patients. Open Heart 2017; 4.

38. Rapin A, Etosse A, Tambosco L, Nicomette J, Percebois-Macadre L, Mouret P, et al. Aerobic capacities and exercise tolerance in neuromuscular diseases: a descriptive study. Ann Phys Rehabil Med 2013; 56: 420–433.

39. van den Berg LE, Favejee MM, Wens SC, Kruijshaar ME, Praet SF, Reuser AJ, et al. Safety and efficacy of exercise training in adults with Pompe disease: evalution of endurance, muscle strength and core stability before and after a 12 week training program. Orphanet J Rare Dis 2015; 10: 1–8.

40. Al-Rahamneh H, Faulkner J, Byrne C, Eston R. Prediction of peak oxygen uptake from ratings of perceived exertion during arm exercise in able-bodied and persons with poliomyelitis. Spinal Cord 2011; 49: 131–135.

41. Crescimanno G, Modica R, Mauro RL, Musumeci O, Toscano A, Marrone O. Role of the cardio-pulmonary exercise test and six-minute walking test in the evaluation of exercise performance in patients with late-onset Pompe disease. Neuromuscul Disord 2015; 25: 542–547.

42. Gimenes A, Neder JA, Dal Corso S, Nogueira CR, Nápolis L, Mello MTd, et al. Relationship between work rate and oxygen uptake in mitochondrial myopathy during ramp-incremental exercise. Braz J Med Biol 2011; 44: 354–360.

43. Montes J, Goodwin AM, McDermott MP, Uher D, Hernandez FM, Coutts K, et al. Diminished muscle oxygen uptake and fatigue in spinal muscular atrophy. Ann Clin Transl Neurol 2021; 8: 1086–1095.

44. Laboratories ACoPSfCPF. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002; 166: 111–117.

45. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. Am J Respir Crit Care Med 1998; 158: 1384–1387.

46. Cejudo P, Bautista J, Montemayor T, Villagómez R, Jiménez L, Ortega F, et al. Exercise training in mitochondrial myopathy: a randomized controlled trial. Muscle Nerve 2005; 32: 342–350.

47. Munters LA, Dastmalchi M, Katz A, Esbjörnsson M, Loell I, Hanna B, et al. Improved exercise performance and increased aerobic capacity after endurance training of patients with stable polymyositis and dermatomyositis. Arthritis Res Ther 2013; 15: 1–13.

48. Bankolé L-C, 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 2016; 95.

49. Mayorga-Vega D, Bocanegra-Parrilla R, Ornelas M, Viciana, J. Criterion-related validity of the distance-and time-based walk/run field tests for estimating cardiorespiratory fitness: a systematic review and meta-analysis. PloS ONE 2016; 11: e0151671.

50. Evans HJ, Ferrar KE, Smith AE, Parfitt G, Eston RG. A systematic review of methods to predict maximal oxygen uptake from submaximal, open circuit spirometry in healthy adults. J Sci Med Sport 2015; 18: 183–188.

51. Langeskov-Christensen M, Heine M, Kwakkel G, Dalgas U. Aerobic capacity in persons with multiple sclerosis: a systematic review and meta-analysis. Sports Med 2015; 45: 905–923.

52. Lennon N, Thorpe D, Balemans AC, Fragala-Pinkham M, O'Neil M, Bjornson K, et al. The clinimetric properties of aerobic and anaerobic fitness measures in adults with cerebral palsy: a systematic review of the literature. Res Dev Disabil 2015; 45: 316–328.

53. Ross RM, Murthy JN, Wollak ID, Jackson AS. The six minute walk test accurately estimates mean peak oxygen uptake. BMC Pulm Med 2010; 10: 1–9.

54. Burr JF, Bredin SS, Faktor MD, Warburton DE. The 6-minute walk test as a predictor of objectively measured aerobic fitness in healthy working-aged adults. Phys Sportsmed 2011; 39: 133–139.
55. Brehm MA, Verduijn S, Bon J, Bredt N, Nollet F. Comparison of two 6-minute walk tests to assess walking capacity in polio survivors. J Rehabil Med 2017; 49: 732–737.
56. Skough K, Broman L, Borg K. Test–retest reliability of the 6-min walk test in patients with postpolio syndrome. Int J Rehabil Res 2013; 36: 140–145.
57. Knak, K. L., Andersen, L. K., Witting, N., & Vissing, J. Reliability of the 2-and 6-minute walk tests inneuromuscular diseases. J Rehabil Med 2017; 49(4): 362–366.
58. Astrand P, Rodahl K. Textbook of Work Physiology, Physiological bases of exercise. Singapore: Mc Graw-Hill; 1986, p. 98.
59. van Groenestijn AC, Verschuren O, Schröder CD, van den Berg LH, Visser-Meily J. The Åstrand-Ryhming test is not a feasible measure in ambulatory patients with amyotrophic lateral sclerosis. J Neuromuscul Dis 2016; 3: 539–544.
60. Lygren H, Jones K, Grenstad T, Dreyer V, Farbu E, Rekand T. Perceived disability, fatigue, pain and measured isometric muscle strength in patients with post-polio symptoms. Physiother Res Int 2007; 12: 39–49.
61. Wells G, Beaton DE, Tugwell P, Boers M, Kirwan JR, Bingham CO, et al. Updating the OMERACT filter: discrimination and feasibility. J Rheumatol 2014; 41: 1005–1010.
62. Robertson RJ, Goss FL, Dube J, Rutkowski J, Dupain M, Brennan C, et al. Validation of the adult OMNI scale of perceived exertion for cycle ergometer exercise. Med Sci Sports Exerc 2004; 36: 102–108.