Reliability of the 2- and 6-minute walk tests in neuromuscular diseases

Knak, Kirsten Lykke; Andersen, Linda Kahr; Witting, Nanna; Vissing, John

Published in:
Journal of Rehabilitation Medicine

DOI:
10.2340/16501977-2222

Publication date:
2017

Document version
Publisher's PDF, also known as Version of record

J Rehabil Med 2017; 49: 362–366

SHORT COMMUNICATION

RELIABILITY OF THE 2- AND 6-MINUTE WALK TESTS IN NEUROMUSCULAR DISEASES

Kirsten Lykke KNAK*, PT, MSc, Linda Kahr ANDERSEN*, PT, MSc, Nanna WITTING, MD, PhD and John VISSING, MD, PhD
From the Copenhagen Neuromuscular Center, Department of Neurology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
**These authors contributed equally to this work.*

Objective: **The 2- and 6-minute walk tests are used to evaluate walking capacity, but reliability has been sparsely investigated in patients with neuromuscular diseases. The aim of this study was to investigate the relative and absolute reliability of the 2- and 6-minute walk tests in patients with neuromuscular diseases.**
Design: **Each patient performed a 2- and a 6-minute walk test on 2 test days separated by 1–2 weeks.**
Subjects: **A total of 93 adult patients (mean age 53 years, age range 22–83 years) with 12 different neuromuscular diseases were included.**
Results: **The mean walking distance increased by 4.3 and 11.2 m ($p < 0.001$) in repeated 2- and 6-minute walk tests, respectively. Intraclass correlation coefficient in the 2- and 6-minute walk tests was 0.99 ($p < 0.001$). Standard error of measurement was 4.9 m in the 2-minute walk test and 14.0 m in the 6-minute walk test. Minimal detectable difference was 13.7 m in the 2-minute walk test and 38.8 m in the 6-minute walk test.**
Conclusion: **These findings show good relative reliability of the 2- and 6-minute walk tests in patients with neuromuscular diseases. However, absolute reliability demonstrated variability in neuromuscular diseases. This should be considered when interpreting a change in walking distance.**

Key words: neuromuscular diseases; reproducibility of results; walk test.

Accepted Feb 14, 2017; Epub ahead of print Mar 29, 2017

J Rehabil Med 2017; 49: 362–366

Correspondence address: Kirsten Lykke Knak, Rigshospitalet, Department of Neurology, Copenhagen Neuromuscular Center, 6921NM, Juliane Maries Vej 28, DK-2100 Copenhagen, Denmark. E-mail: kirsten.lykke.knak@regionh.dk

The 2-minute walk test (2MWT) and 6-minute walk test (6MWT) are used to evaluate walking capacity in patients with neuromuscular diseases (NMD) in the clinic and in clinical trials. However, reliability of the 2MWT has been documented only in Duchenne muscular dystrophy (1) and reliability of the 6MWT has been documented only in Duchenne muscular dystrophy (1, 2), myotonic dystrophy (3), and spinal and bulbar muscular atrophy (4). Relative reliability characterizes the trustworthiness of a test in a specific diagnosis. Absolute reliability provides actual values for measurement error that should be exceeded in order to represent a real difference in a patient's clinical condition.

The aim of the present study was to investigate the relative and absolute reliability of the 2MWT and 6MWT in patients with NMD. The walk tests in some of the patients have been published previously in Andersen et al. (5).

METHODS

Recruitment

A total of 93 patients were recruited from March 2014 to January 2015. Patients followed at the Copenhagen Neuromuscular Center were invited to participate in the study. Inclusion criteria were: biopsy or genetically confirmed NMD; age ≥18 years; and capacity to walk further than 60 m. Exclusion criteria were: conditions with a significant impact on walking capacity.

Study design

This is a methodological study. The patients were tested at our Neuromuscular Center on 2 test days at the same time of the day, 1–2 weeks apart. On each test day, a 2MWT and a 6MWT were performed, with 30 min rest between tests. Patients were block randomized (block size 10) using sealed envelopes on the first test day to perform either the 2MWT or the 6MWT first. The same order was used for re-test. The 6MWT was performed according to the American Thoracic Society (ATS) guidelines (6) and the 2MWT was a modified version of the ATS guidelines for the 6MWT. In the 2MWT, patients were informed that the test was a 2MWT, and the encouragement used at 3 min in ATS guidelines was used at 1 min in the 2MWT. Two investigators performed the 2MWT and 6MWT, but each patient was instructed by the same investigator at test and re-test.

Patient characteristics and walking distance in the present study investigating relative and absolute reliability have been reported in a published study assessing validity of the 2MWT compared with the 6MWT (5). The Regional Committee on Health Research Ethics in Denmark approved the study (H-4-2014-FSP) and informed verbal consent was obtained.

Statistical analysis

Analyses have been conducted in SPSS version 22, except Student's paired *t*-tests, which were performed in Microsoft Excel 2010. As subgroup behave differently, we have conducted both analyses for all patients as a group and for each NMD. Data for all patients as a group and for each subgroup were normally distributed tested by Shapiro-Wilk test and Quantile-Quantile Plot in SPSS. Systematic differences in walking distance between test and re-test were tested by the parametric 2-sided Student's paired *t*-test. Patient characteristics

doi: 10.2340/16501977-2222

and change in walking distance between tests are presented as mean and standard deviation (SD). Relative reliability was determined by intraclass correlation (ICC) coefficient (model: 3, 2-way mixed, type: absolute agreement, confidence interval: 95%). Relative reliability indicates the degree of reliability and ranges from 0.00 = no reliability to 1.00 = perfect reliability (7). The higher reliability, the less measurement error in the assessment. Absolute reliability was investigated by standard error of measurement (SEM), minimal detectable difference (MDD) (repeatability) for 95% confidence interval and by Bland-Altman plots, which also tested for heteroscedasticity (if variability is unequal across magnitudes of the mean).

$SEM = SD*\sqrt{1-ICC}$ (7) $\quad MDD_{95} = 1.96*SEM*\sqrt{2}$ (7)

SEM% and MDD% were calculated using the formulas:

$SEM\% = SEM/mean*100$ (8) $MDD_{95}\% = MDD/mean*100$ (8)

The mean for SEM and MDD_{95} is for the test and re-test. SEM and MDD_{95} estimate the smallest change that is beyond measurement error in a group and in a single subject, respectively (3). Bland-Altman plot visualizes agreement between repeated tests. The higher reliability, the closer the dots are to zero. Statistical significance was set at $p \leq 0.05$. Results are shown for the NMD patients all together and for each NMD. *Post hoc* sample size calculation with alpha level = 0.05, 2-tailed test and correlation coefficient = 0.90 required a sample of 7 patients (7).

RESULTS

Patient characteristics

Patient characteristics were as follows: 12 different NMD (Fig. 1, Table I), 32 females, 61 males, age 53 years (SD 17), height 1.74 m (SD 0.10), weight 76.2 kg (SD 16.3), body mass index 25.2 kg/m^2 (SD 5). Mean muscle strength, using the Medical Research Council scale, for ankle dorsal flexion was 4 (SD 1.6), ankle plantar flexion 4.4 (SD 1.1), hip flexion 4.4 (SD 1), and hip extension 4.5 (SD 0.7). Twenty-eight patients used assistive devices in the 2MWT and 6MWT. At re-test of the 2MWT, 60 patients walked longer, 27 walked shorter, and 6 walked the same distance. Similarly, in the 6MWT, 67 patients walked longer at re-test, 24 walked shorter, and 2 walked the same distance.

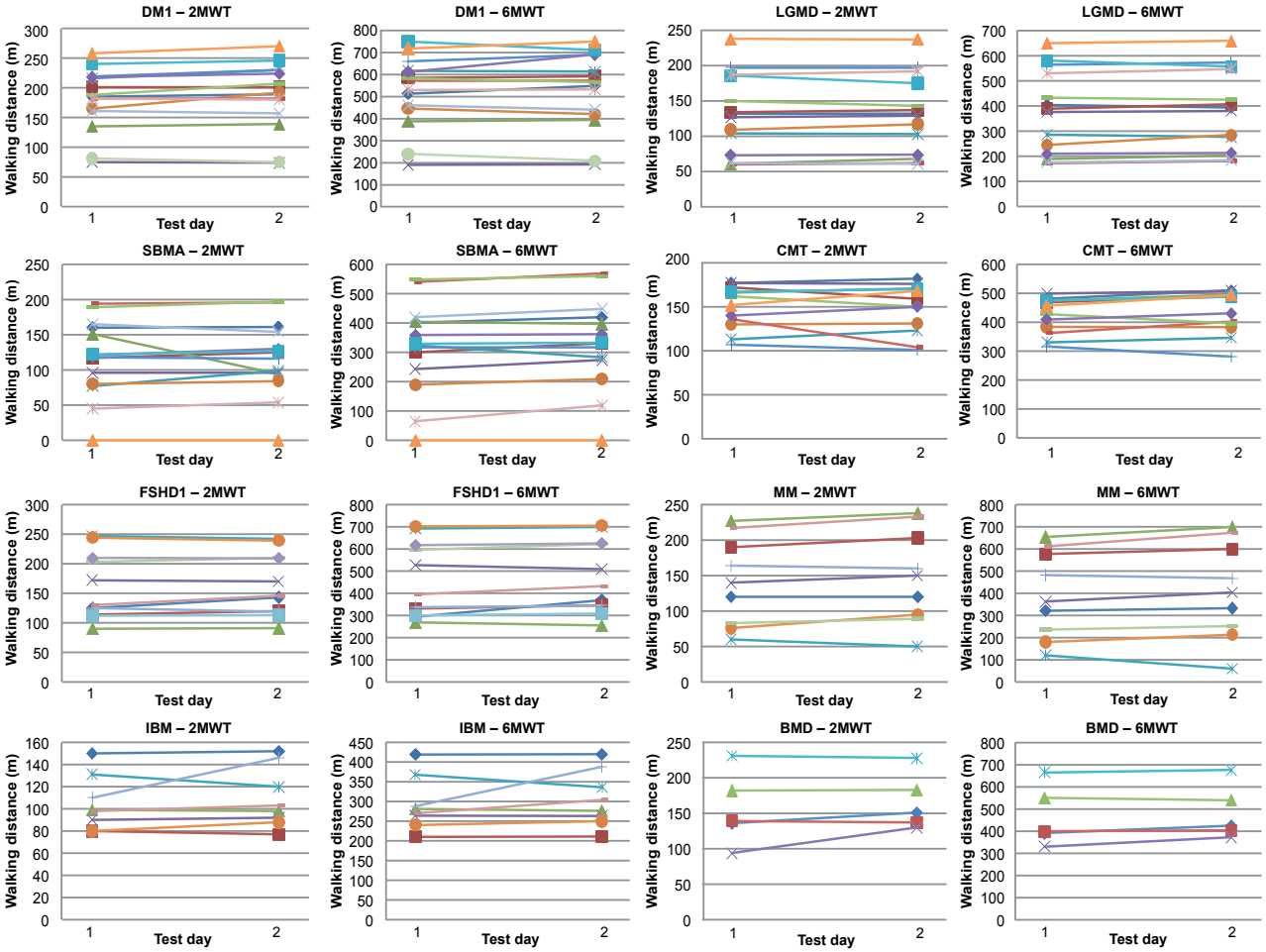

Fig. 1. Walking distance in the 2MWT and 6MWT in individual patients with neuromuscular diseases. 2MWT: 2-minute walk test; 6MWT: 6-minute walk test; DM1: myotonic dystrophy type 1; LGMD: limb-girdle muscular dystrophy; SBMA: spinobulbar muscular atrophy/Kennedy disease; CMT: Charcot-Marie-Tooth disease/hereditary motor and sensory neuropathy; FSHD1: facioscapulohumeral muscular dystrophy type 1; MM: mitochondrial myopathy; IBM: sporadic inclusion body myositis; BMD: Becker muscular dystrophy. The walking distance (m) is on the y-axis and test days on the x-axis.

Table I. Walking distance in the 2-minute walk test and 6-minute walk test in each neuromuscular diseases

		2MWT				6MWT			
	n	Day 1 Mean (SD)	Day 2 Mean (SD)	Difference Mean (SD)	*p*-value	Day 1 Mean (SD)	Day 2 Mean (SD)	Difference Mean (SD)	*p*-value
DM1	15	180.4 (52.0)	194.4 (56.5)	14.0 (9.2)	0.02*	524.5 (158.7)	528.6 (169.3)	4.1 (30.6)	0.61
LGMD	14	130.0 (56.6)	130.5 (55.0)	0.5 (5.0)	0.71	371.7 (164.1)	377.6 (160.4)	5.9 (15.5)	0.18
SBMA	13	125.8 (44.8)	125.6 (42.5)	−0.2 (18.7)	0.98	341.5 (131.6)	355.5 (127.1)	14.0 (23.3)	0.05*
CMT	12	149.9 (24.4)	148.8 (27.9)	−1.2 (13.1)	0.76	423.8 (62.8)	435.3 (74.3)	11.5 (23.9)	0.12
FSHD1	11	161.1 (56.4)	163.8 (53.4)	2.7 (8.2)	0.30	460.0 (169.0)	474.1 (164.7)	14.1 (25.5)	0.10
MM	9	141.9 (61.9)	148.7 (66.3)	6.8 (9.7)	0.07	393.8 (196.2)	411.3 (219.2)	17.6 (36.4)	0.19
IBM	8	104.8 (24.7)	109.5 (27.4)	4.8 (13.9)	0.36	292.5 (68.4)	306.0 (71.3)	13.5 (39.4)	0.36
BMD	5	156.4 (52.0)	165.8 (40.3)	9.4 (16.5)	0.27	467.6 (137.1)	483.8 (125.1)	16.2 (21.8)	0.17

*$p \leq 0.05$.
2MWT: 2-minute walk test; 6MWT: 6-minute walk test; DM1: myotonic dystrophy type 1; LGMD: limb-girdle muscular dystrophy; SBMA: spinobulbar muscular atrophy/Kennedy disease; CMT: Charcot-Marie-Tooth disease/hereditary motor and sensory neuropathy; FSHD1: facioscapulohumeral muscular dystrophy type 1; MM: mitochondrial myopathy; IBM: sporadic inclusion body myositis; BMD: Becker muscular dystrophy.

Relative and absolute reliability

For all patients as a group the mean walking distance in repeated 2MWT increased by 4.3 m (SD 9.1) (2.9 %) from 144.9 (SD 51.8) to 149.2 m (SD 52.2) ($p<0.001$) and in the 6MWT by 11.2 m (SD 26.0) (2.7 %) from 412.0 (SD 155.6) to 423.2 m (SD 158.1) ($p<0.001$). Results for subgroups are shown in Fig. 1.

Relative reliability was high for all patients as a group in 2MWT and 6MWT (ICC=0.99, $p<0.001$, 95% CI 0.98–1.00). Similar results were seen in all subgroups (ICC>0.90, $p<0.004$). For all diseases as a group SEM and MDD_{95} were 4.9 m (3.4%) and 13.7 m (9.3%), respectively, in the 2MWT, and 14.0 m (3.4%) and 38.8 m (9.3%) in the 6MWT. In subgroups, SEM varied from 0.5–9.1 m in the 2MWT and 7.1–37.0 m in the 6MWT, and MDD_{95} varied between 1.5–25.2 m in the 2MWT and between 19.7–102.5 m in the 6MWT. Results for subgroups are shown in Table II.

Agreement between repeated 2MWT and 6MWT is shown in Bland-Altman plots (Fig. 2, A+B) with a 95% CI showing a variation of –13.9 to +22.5 m for the 2MWT and –40.8 to+63.3 m for the 6MWT. The difference in walking distance from test to re-test was not dependent on walking distance on the first test day, meaning there was no heteroscedasticity.

DISCUSSION

The main finding of the present study is that the 2MWT and 6MWT are as reliable in NMD as in other diseases. Relative reliability was excellent, with an ICC of 0.99 in both the 2MWT and 6MWT. Similar results were seen in all subgroups (ICC>0.90). These findings are consistent with previous studies of the 2MWT and 6MWT in different neuromuscular (3, 4, 9, 10) and non-neuromuscular (11) conditions (ICC 0.61–1.00).

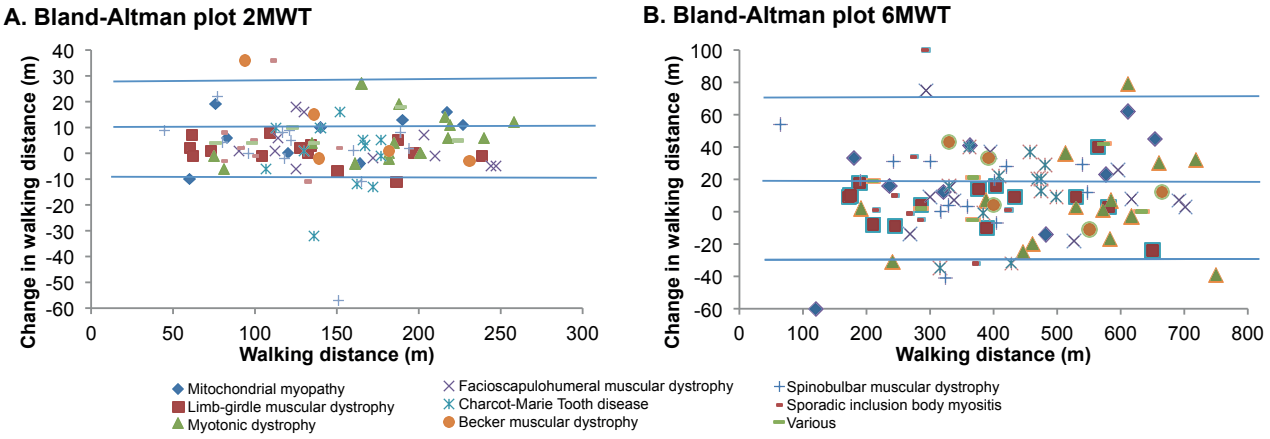

Fig. 2. Absolute reliability of the 2MWT and 6MWT in patients with neuromuscular diseases. 2MWT: 2-minute walk test, 6MWT: 6-minute walk test. The different NMD are symbolized by different colours. Each dot represents a single patient. Various (6): spinal muscular atrophy (3), myotonia congenital Thomsen disease (1), congenital myopathy (1) and polymyositis (1). A+B: The Bland-Altman plots show the change in walking distance between repeated tests for each patient compared with mean ± 1.96 standard deviation (SD) for all NMD as a group. The walking distance on test day 1 indicates whether the variability among tests depends on the walking distance on the first test day. Therefore, data on the x-axis is the walking distance from the first test day and not the mean walking distance from test day 1 and test day 2. The x-axis shows the walking distance (m) on the first test day. The y-axis shows the change in walking distance (m) from the first test day to the second test day. The closer the dots are to zero the more reliable the test is. The 3 horizontal lines in each graph illustrate the mean difference from test to re-test and the 95% confidence intervals, with the upper line showing 1.96*SD plus the mean and the lower line showing 1.96*SD minus mean.

Table II. Reliability of the 2MWT and 6MWT in each neuromuscular diseases

		2MWT			6MWT		
	n	ICC Mean (95% CI)	SEM Frequency (%)	MDD_{95} Frequency (%)	ICC Mean (95% CI)	SEM Frequency (%)	MDD_{95} Frequency (%)
DM1	15	0.99* (0.96–1.00)	0.5 (0.3)	1.5 (0.8)	0.99* (0.98–1.00)	14.4 (2.7)	40.0 (7.6)
LGMD	14	1.00* (0.99–1.00)	2.5 (1.9)	6.8 (5.2)	1.00* (0.99–1.00)	7.1 (1.9)	19.7 (5.3)
SBMA	13	0.96* (0.85–0.99)	9.1 (7.2)	25.2 (20.0)	0.99* (0.96–1.00)	13.3 (3.8)	36.9 (10.6)
CMT	12	0.94* (0.79–0.98)	6.4 (4.3)	17.7 (11.8)	0.97* (0.87–1.00)	12.6 (2.9)	35.0 (8.2)
FSHD1	11	0.99* (0.98–1.00)	4.2 (2.6)	11.5 (7.1)	0.99* (0.97–1.00)	13.6 (2.9)	37.8 (8.1)
MM	9	0.99* (0.95–1.00)	5.6 (0.7)	15.5 (10.6)	0.99* (0.96–1.00)	19.2 (4.8)	53.1 (13.2)
IBM	8	0.93* (0.66–0.99)	6.9 (6.5)	19.2 (18.0)	0.91* (0.61–0.98)	19.9 (6.6)	55.1 (18.4)
BMD	5	0.96* (0.73–1.00)	8.4 (5.2)	23.2 (14.4)	0.91* (0.91–1.00)	37.0 (7.8)	102.5 (21.6)

*$p<0.05$. 2MWT: 2-minute walk test.
6MWT: 6-minute walk test; DM1: myotonic dystrophy type 1; LGMD: limb-girdle muscular dystrophy; SBMA: spinobulbar muscular atrophy/Kennedy disease; CMT: Charcot-Marie-Tooth disease/hereditary motor and sensory neuropathy; FSHD1: facioscapulohumeral muscular dystrophy type 1; MM: mitochondrial myopathy; IBM: sporadic inclusion body myositis; BMD: Becker muscular dystrophy; ICC: intraclass correlation coefficient. SEM: standard error of measurement; MDD_{95}: minimal detectable difference for 95 % confidence interval.

A general guideline suggests that ICC >0.75=good reliability, and ICC >0.90 is recommended for clinical decisions (7). According to these guidelines, the 2MWT and 6MWT are reliable measures for clinical decisions. However, the acceptable level of ICC depends on the specific circumstances. A lower level of ICC can be accepted if 2MWT or 6MWT are secondary endpoints as opposed to a primary outcome.

Despite the good relative reliability of the 2MWT and 6MWT, there was a statistically significant increase in walking distance from test to re-test of 3% in both tests for NMD as a group (4.3 and 11.2 m in the 2MWT and 6MWT, respectively), which could be due to a learning effect. Variation was also present in subgroups, but the variation was not statistically significant in most of them. The results might be explained by small samples in subgroups and the large variability in walking capacity within subgroups. Statistically significant variability in walking distance in repeated 6MWT has been reported in previous studies of NMD with an increase of 3–5% (17.0–24.0 m) (3, 12). However, variability in post-polio syndrome (PPS) showed a non-statistically significant increase in walked distance of less than 2% in 6MWT (9) and 2MWT (10).

Absolute reliability provides information about absolute difference between repeated tests. In line with previous findings (3, 13, 14), SEM in our study was 4.9 m (3.4%) in the 2MWT and 14.0 m (3.4%) in the 6MWT for all patients as a group, but varied in subgroups. MDD_{95} was 13.7 m (9.3%) in the 2MWT and 38.8 m (9.3%) in the 6MWT for all patients as a group, but large subgroup variability was found. The MDD_{95} for the 2MWT and 6MWT is consistent with previous studies (3, 9, 13, 14), except from Stolwijk-Swüste et al (10), which reported a smallest detectable change of 22.9 m (16.8%) for the 2MWT in PPS. The inconsistency might be due to different NMD and patient characteristics (e.g. age and functional level) across studies. The larger SEM and MDD_{95} in the subgroups might be explained by: (*i*) different NMD; (*ii*) different circumstances in repeated tests in some patients (e.g. poor sleep); (*iii*) difference in sample size as a larger sample reduces SEM; and (*iv*) difference in walking capacity across subgroups as there was a tendency that subgroups with the longest walking distance (myotonic dystrophy type 1 and facioscapulohumeral muscular dystrophy) showed the smallest measurement error, while, in contrast, subgroups with the shortest walking distance (sporadic inclusion body myositis and spinobulbar muscular atrophy) showed the largest measurement error. Thus, measurement error seems to be dependent on the functional level; measurement error increases with decreasing walking capacity. Clinicians need to consider the walking capacity when interpreting a change between assessments. SEM% and MDD_{95}% vary between 2MWT and 6MWT in subgroups. Not surprisingly, those showing the highest SEM% in 2MWT also have the highest MDD_{95}% in the 2MWT. The NMD affected in the peripheral nervous system (spinobulbar muscular atrophy and Charcot-Marie-Tooth disease) show the highest SEM% and MDD_{95}% in the 2MWT. In contrast, most of the muscular dystrophies (myotonic dystrophy type 1, facioscapulohumeral muscular dystrophy and Becker muscular dystrophy) show the highest SEM% and MDD_{95}% in the 6MWT. These findings have important implications for the interpretation of a change in walking distance in repeated 2MWT and 6MWT at follow-up or in clinical trials. The variability among test results, especially in outliers, could potentially be corrected by measuring mean heart rate during walking, as suggested recently (12).

The present study was limited by only a very few protocol violations, such as testing performed on different times of day in a few patients. However, the patients with protocol violations showed similar results as the other patients and were therefore not excluded from the study.

In conclusion, our findings show good relative reliability of the 2MWT and 6MWT in patients with NMD as a group and in subgroups. However, absolute reliability demonstrated variability in NMD as a group and a considerable amount of variability across subgroups. Decreased level of walking capacity showed increased level of absolute measurement error. Thus, clinicians must be aware of the issues associated with absolute reliability when interpreting results of 2MWT and 6MWT in individual patients in the clinic.

ACKNOWLEDGEMENTS

KLK and LKA received travel grants from Lundbeck Foundation, Rigshospitalets' Anniversary Foundation and PTC Therapeutics. JV received research and travel grants from Genzyme Corporation, and serves on the global scientific advisory board for Pompe disease for Genzyme Corporation. NW reports no disclosures.

REFERENCES

1. Vill K, Ille L, Schroeder SA, Blaschek A, Müller-Felber W. Six-minute walk test versus two-minute walk test in children with Duchenne muscular dystrophy: Is more time more information? Eur J Paediatr Neurol 2015; 19: 640–646.
2. McDonald CM, McDonald CM1, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, et al. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. Muscle Nerve 2010; 41: 500–510.
3. Kierkegaard M, Tollbäck A. Reliability and feasibility of the six minute walk test in subjects with myotonic dystrophy. Neuromuscul Disord 2007; 17: 943–949.
4. Takeuchi Y, et al. Walking capacity evaluated by the 6-minute walk test in spinal and bulbar muscular atrophy. Muscle Nerve 2008; 38: 964–971.
5. Andersen LK, Knak KL, Witting N, Vissing J. Two- and 6-minute walk tests well assess walking capability equally in neuromuscular diseases. Neurology 2016; 86: 442–445.
6. 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: 111–117.
7. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 3rd edn. New Jersey: Pearson Prentice Hall; 2009: p. 594–595, 850.
8. Schwenk M, Gogulla S, Englert S, Czempik A, Hauer K. Test–retest reliability and minimal detectable change of repeated sit-to-stand analysis using one body fixed sensor in geriatric patients. Physiol Meas 2012; 33: 1931–1946.
9. Skough K, Broman L, Borg K. Test-retest reliability of the 6-min walk test in patients with postpolio syndrome. Int J Rehab Res 2013; 36: 140–145.
10. Stolwijk-Swüste J, Beelen A, Lankhorst G, Nollet F, CARPA study group. SF36 physical functioning scale and 2-minute walk test advocated as core qualifiers to evaluate physical functioning in patients with late-onset sequelae of poliomyelitis. J Rehabil Med 2008; 40: 387–394.
11. Hamilton DM, Haennel RG. Validity and reliability of the 6-minute walk test in a cardiac rehabilitation population. J Cardpulm Rehabil. 2000; 20: 156–164.
12. 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.
13. Connelly DM, Thomas BK, Cliffe SJ, Perry WM, Smith RE. Clinical utility of the 2-minute walk test for older adults living in long-term care. Physiother Can 2009; 61: 78–87.
14. Hiengkaew V, Jitaree K, Chaiyawat P. Minimal detectable changes of the Berg Balance Scale, Fugl-Meyer Assessment Scale, Timed "Up & Go" Test, gait speeds, and 2-minute walk test in individuals with chronic stroke with different degrees of ankle plantarflexor tone. Arch Phys Med Rehabil 2012; 93: 1201–1208.