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

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 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).
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).

\[
SEM = SD \times \sqrt{1 - ICC} \quad (7)
\]

\[
MDD_{95} = 1.96 \times SEM \times \sqrt{2} \quad (7)
\]

SEM% and MDD% were calculated using the formulas:

\[
SEM\% = SEM / mean \times 100 \quad (8)
\]

\[
MDD\% = MDD / mean \times 100 \quad (8)
\]

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 \( p < 0.05 \). 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 1), 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.
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 364 (SD 169) to 375.4 m (SD 171) (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 MDD95 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 (9.8%) 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 MDD95 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.

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

<table>
<thead>
<tr>
<th>Disease</th>
<th>2MWT Mean (SD)</th>
<th>6MWT Mean (SD)</th>
<th>Difference Mean (SD)</th>
<th>p-value</th>
<th>2MWT Mean (SD)</th>
<th>6MWT Mean (SD)</th>
<th>Difference Mean (SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM1</td>
<td>157.9 (54.8)</td>
<td>165.8 (55.0)</td>
<td>7.9 (9.2)</td>
<td>0.02</td>
<td>157.4 (54.8)</td>
<td>162.1 (55.0)</td>
<td>4.7 (9.2)</td>
<td>0.02</td>
</tr>
<tr>
<td>LGMD</td>
<td>190.3 (46.8)</td>
<td>197.7 (47.0)</td>
<td>7.4 (9.2)</td>
<td>0.01</td>
<td>190.3 (46.8)</td>
<td>197.7 (47.0)</td>
<td>7.4 (9.2)</td>
<td>0.01</td>
</tr>
<tr>
<td>SBMA</td>
<td>155.2 (44.8)</td>
<td>163.1 (45.0)</td>
<td>7.9 (9.2)</td>
<td>0.02</td>
<td>155.2 (44.8)</td>
<td>163.1 (45.0)</td>
<td>7.9 (9.2)</td>
<td>0.02</td>
</tr>
<tr>
<td>CMET</td>
<td>157.2 (46.8)</td>
<td>164.3 (47.0)</td>
<td>7.1 (9.2)</td>
<td>0.01</td>
<td>157.2 (46.8)</td>
<td>164.3 (47.0)</td>
<td>7.1 (9.2)</td>
<td>0.01</td>
</tr>
<tr>
<td>FSHD1</td>
<td>159.2 (54.8)</td>
<td>166.1 (55.0)</td>
<td>6.9 (9.2)</td>
<td>0.01</td>
<td>159.2 (54.8)</td>
<td>166.1 (55.0)</td>
<td>6.9 (9.2)</td>
<td>0.01</td>
</tr>
<tr>
<td>MM</td>
<td>156.3 (44.8)</td>
<td>163.9 (45.0)</td>
<td>7.6 (9.2)</td>
<td>0.01</td>
<td>156.3 (44.8)</td>
<td>163.9 (45.0)</td>
<td>7.6 (9.2)</td>
<td>0.01</td>
</tr>
<tr>
<td>IBM</td>
<td>156.4 (52.0)</td>
<td>164.1 (52.2)</td>
<td>7.7 (9.2)</td>
<td>0.02</td>
<td>156.4 (52.0)</td>
<td>164.1 (52.2)</td>
<td>7.7 (9.2)</td>
<td>0.02</td>
</tr>
<tr>
<td>BMD</td>
<td>156.5 (52.0)</td>
<td>164.2 (52.2)</td>
<td>7.7 (9.2)</td>
<td>0.02</td>
<td>156.5 (52.0)</td>
<td>164.2 (52.2)</td>
<td>7.7 (9.2)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

* 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; CMET: 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.

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 colors. 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 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.](www.medicaljournals.se/jrm)
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 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 MDD95% vary between 2MWT and 6MWT in subgroups. Not surprisingly, those showing the highest SEM% in 2MWT also have the highest MDD95% in the 2MWT. The NMD affected in the peripheral nervous system (spinobulbar muscular atrophy and Charcot-Marie-Tooth disease) show the highest SEM% and MDD95% in the 2MWT. In contrast, most of the muscular dystonephacies (myotonic dystrophy type 1, facioscapulohumeral muscular dystrophy and Becker muscular dystrophy) show the highest SEM% and MDD95% 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.

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

<table>
<thead>
<tr>
<th>Disease</th>
<th>2MWT</th>
<th>6MWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>SEM%</td>
<td>MDD95%</td>
</tr>
<tr>
<td>DM1</td>
<td>15</td>
<td>0.99 (0.96–1.00)</td>
</tr>
<tr>
<td>LGMD</td>
<td>14</td>
<td>1.00 (0.99–1.00)</td>
</tr>
<tr>
<td>SBMA</td>
<td>13</td>
<td>0.96 (0.85–0.99)</td>
</tr>
<tr>
<td>CMT</td>
<td>12</td>
<td>0.94 (0.79–0.98)</td>
</tr>
<tr>
<td>FSHD1</td>
<td>11</td>
<td>0.99 (0.98–1.00)</td>
</tr>
<tr>
<td>MM</td>
<td>9</td>
<td>0.99 (0.95–1.00)</td>
</tr>
<tr>
<td>IBM</td>
<td>8</td>
<td>0.93 (0.66–0.99)</td>
</tr>
<tr>
<td>BMD</td>
<td>5</td>
<td>0.96 (0.73–1.00)</td>
</tr>
</tbody>
</table>

*SEM: standard error of measurement; MDD95: minimal detectable difference for 95% confidence interval.*
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