Reliability and reproducibility of six-minute walk test in healthy children

Confiabilidade e reprodutibilidade do teste de caminhada de seis minutos em crianças saudáveis

Confiabilidad y reproducibilidad de la prueba de caminata de seis minutos en niños sanos

Renata Martins¹, Renata Maba Gonçalves¹, Anamaria Fleig Mayer¹, Camila Isabel Santos Schivinski³

ABSTRACT | Was verified the reproducibility of the six-minute walk test (6MWT) in Brazilian healthy children. A transversal cross-sectional observational study was carried out between October 2012 and July 2013, with healthy children aged between 6 and 14 years. Initially, the participants were assessed as biometric data (weight, height, body mass index, body surface area and length of the lower limbs) and spirometry. Two 6MWT were performed with 30min interval between them. The retest was carried out after two weeks. Statistical analysis included Shapiro-Wilk normality and analysis of variance (ANOVA) tests for comparison between the 6MWT; intraclass correlation coefficient - two way mixed model, consistency (ICC) were used to assess reproducibility, as well as the provision Bland & Altman. The level of significance was 5% (p<0,05). Were enrolled 29 children in the study (16 females, mean age 10,28 ± 2,25 years). Reproducibility was found between the two longest 6MWT distances, with ICC=0,82 (p<0,001) as well as similarity in behavior of physiological parameters. The children walked similar distances in the first and second test on both days, showing a walking distance variation (Δ) both negative on day 1 and on day 2 (-5,52 m [confidence interval of 95% (95%CI) -28,475–17,417 m] and -2,26 m [95%CI -28,503–23,982 m], respectively). The 6MWT showed to be reproducible in healthy schoolchildren. The retest showed no improvement in the performance of the studied population, suggesting no learning effect.

Keywords | Child; Walking; Reproducibility of Results.

RESUMO | Verificou-se a reprodutibilidade do teste de caminhada de seis minutos (TC6min) em crianças saudáveis brasileiras. Este estudo observacional transversal foi realizado entre outubro de 2012 e julho de 2013 com crianças saudáveis com idades entre 6 e 14 anos. Inicialmente, os escolares foram avaliados quanto aos dados biométricos (peso, altura, índice de massa corporal – IMC, área de superfície corporal e comprimento dos membros inferiores) e espirométricos. Na sequência, foram realizados 2 TC6min, com intervalo de 30 minutos entre eles. O reteste foi conduzido após duas semanas. A análise estatística incluiu os testes de normalidade Shapiro-Wilk e análise de variância (ANOVA), para comparação entre os TC6min; o coeficiente de correlação intraclassé de duas vias (consistência) (ICC) foi utilizado na verificação da reprodutibilidade; bem como a disposição gráfica de Bland e Altman. O nível de significância adotado foi de 5% (p<0,05). Participaram do estudo 29 escolares, sendo 16 do sexo feminino, com média de idade de 10,28±2,25 anos. Analisando-se os TC6min com maior distância percorrida (DP), identificou-se reprodutibilidade do teste, com ICC=0,82 (p<0,001), assim como semelhança no comportamento dos parâmetros fisiológicos considerados. As crianças caminharam distâncias semelhantes no primeiro e segundo teste em ambos os dias, apresentando uma variação (Δ) negativa da DP tanto no dia 1 como no dia 2 (-5,52 m [intervalo de confiança de 95% (IC95%) -28,475–17,417 m] e -2,26 m [IC95% -28,503–23,982 m], respectivamente). O TC6min mostrou-se reprodutível em escolares saudáveis. A repetição do teste não apresentou melhora no desempenho, sugerindo não haver efeito aprendizado na população estudada.

Descritores | Criança; Caminhada; Reprodutibilidade dos Testes.

DOI: 10.590/1809-2950/92221032014
INTRODUCTION

The individual response to exercise provides information related to respiratory, cardiac, metabolic and muscular systems, being recognized as an important clinical evaluation instrument. For the characterization of the different types of exercises, daily activities are considered to be submaximal exercises. Therefore, submaximal functional tests have been proposed to assess the physical capacity of healthy and sick individuals; among them, the Six-Minute Walk test stands out (6MWT).

The 6MWT is a result of a modification in the 12-minute walk test, due to the limited physical conditions of patients with chronic bronchitis, and, ever since then, it has been widely used. It is widely applied, has low cost and is easy to administer, since it requires little equipment and less technical experience. It can translate the individual skill for activities of daily living, since it assesses the response to exercise in an integrated and global form from all of the systems involved in physical activity. Therefore, it is an alternative to replace maximal exercise tests.

Studies have shown its use, both in the adult population and among children and adolescents. In this second age group, the conduction of cardiopulmonary exercises is especially problematic, because usually a high level of cooperation and motor coordination is required. In spite of that, its indication in pediatrics is still questioned, specifically concerning its reproducibility and reliability, once the test was created for adults, thus not considering childhood. Literature has argued about factors that can influence the performance of children in the 6MWT, among which are puberty and growth spurt, due to the important impact of step size and velocity (V) over walking distance (WD) in the test. Besides, aspects of school age related to difficulties, learning, interest and motivation should also be considered.

However, some studies of validity and reproducibility of the 6MWT in pediatrics have been published, not only with the objective of reinforcing its application as an instrument of functional assessment and clinical control, but also to know the specificities in the behavior of different ages during the test. However, the reproducibility and reliability of the 6MWT still have not been verified among Brazilian healthy children in previously published studies.

In this context, the objective of this analysis was to observe the reproducibility and the reliability (interevaluator) of the 6MWT among Brazilian healthy children.

METHODOLOGY

This is a cross-sectional observational and prospective study conducted between October, 2012, and July, 2013, with healthy children from private and public schools in Florianópolis (SC), Brazil. Data collection was carried out in schools and in the Center of Health Sciences and Sport (CEFOD) at Universidade do Estado de Santa Catarina (UDESC), after the approval of the Ethics Committee of UDESC (CAAE: 07635412.3.0000.0118). Children aged from 6 to 14 years old participated in the study, without history of cardiopulmonary, musculoskeletal, rheumatic, neurological disease or hearing and visual impairment; they were physically active or not (this aspect was verified by the health survey elaborated by the researchers) and were not enrolled...
in high performance sports federations (non-athletes). The ones who did not do some of the tests properly, for any reason, be it for inability or lack of understanding, or the ones who could not conclude any of the evaluation procedures and activities proposed on the day of data collection, were excluded from the sample. Besides, when the answer to the International Study of Asthma and Allergies in Childhood (ISAAC) \(^{15,16}\) characterized respiratory impairment, and/or spirometric parameter of Forced Expiratory Volume in the first second (FEV\(_1\)) and/or forced vital capacity (FVC) presented values lower than 80% of the predictions, according to references by Knudson \(^{17}\) et al. and Polgar \(^{18}\) et al., the student was also excluded from the study.

After the parents or people in charge signed the informed consent form, authorizing the children’s participation in the study, an identification form was filled out, and the ISAAC questionnaire was applied. Afterwards, a biometric evaluation was conducted (weight, height, body mass index – BMI) \(^{19}\). Spirometry was conducted, always by the same evaluator, by using the equipment EasyOne – Medical Technologies (Ndd Medizintechnik AG, Switzerland), according to acceptability and reproducibility criteria from the American Thoracic Society (ATS) \(^{20}\), considering the percentages of predicted values in the variables FEV\(_1\), FVC and FEV\(_1\)/FVC \(^{17,18}\). Afterwards, students were submitted to two 6MWTs, with a 30 minute interval, according to the rules by ATS \(^{3}\). The test was conducted in a flat 30 m long corridor, where participants were instructed to walk as fast as they could, without running, while listening to standardized encouragement statements \(^3\). After the 6 minutes, the WD and V were recorded. Parameters of control, such as heart rate (HR), peripheral oxygen saturation (PO\(_2\)S), dyspnea index (by using the Borg modified scale \(^{21}\)), respiratory frequency (RF) and blood pressure (BP) were monitored in the beginning and in the end of the test, and the three first parameters were also observed during the test (2\(^{nd}\) and 4\(^{th}\) minute). In order to verify the PO\(_2\)S and HR, the New Tech PM100c oximeter was used. Both 6MWTs were conducted by different evaluators (reliability), and the order of evaluation was randomized by one data (even numbers corresponded to one evaluator, and odd numbers, to another). The test–retest, according to the same procedures, was performed after 14 days (day 2) (reproducibility and reliability), to make sure there would be no overlapping effect in relation to the other test, and so that, in the meantime, anthropometric measurements would not suffer considerable changes \(^{5,22}\).

The determination of sample size was conducted according to the hypothesis test for reliability studies \(^{23}\), by adopting a 5% significance level, a 90% test power and 95% reliability. Calculation considered WD in the 6MWT from the pilot study, whose standard deviation was of 54 m. Based on these data, and on the intention to detect a difference of around 35 m, the sample size was estimated in 25 individuals. With a 10% loss prediction, 28 children would be sufficient for the study.

Data were analyzed with the software SPSS for Windows, version 20.0, and treated with descriptive analysis (mean and standard—deviation) and frequencies. The Shapiro–Wilk test was used to verify data normality. The Analysis of Variance (ANOVA) was applied to verify the differences of WD at different moments and days, when the 6MWT was conducted. Reproducibility and reliability (interevaluator) of the 6MWT were determined by the two-way intraclass correlation coefficient (consistency) (ICC). The ICC was interpreted according to Munro’s classification system \(^{24}\), considering little correlation (≤0.25), low correlation (0.26–0.49), moderate (0.50–0.69), high (0.7–0.89) and very high (0.9–1.0) correlation. Bland & Altman’s provision \(^{25}\) was also used for the reliability analysis, by using the WD data on both days of evaluation, both inter and intraevaluator, for enabling the better visualization of agreement between the individual measurements. For that, the software GraphPad Prism 5 was used. The adopted significance level for the statistical treatment was 5% (p<0.05).

RESULTS

Twenty-nine children participated in the study, being 16 female participants. The sample characterization is presented in Table 1.

The means of the distance walked by the children in the first and second tests were measured on both days, as well as the variation of WD on day 1 and on day 2 (WD on the first day: 569.59±86.96 m versus 564.06±80.85 m; WD on the second day: 564.06±80.85 m versus 554.19±76.19 m; F=697; p=0.554), which is described in Table 2. The WD variation was calculated by the difference (Δ) between the WD in the first and in the second 6MWT on both days.

In the verification of data in ANOVA, no significant difference was identified in the walked distance in the 6MWT, on days 1 and 2 (F=697; p=0.554) (Table 2).
The ICCs of the interevaluator WD variable and between the 2 days of the 6MWT are shown in Table 3. It is possible to observe moderate to high correlation (ICC = 0.742; 0.581; 0.727; 0.590) of the WD between both evaluators and both days, and these data are demonstrated by Bland and Altman’s graphs (Figure 1A to 1D). When the 6MWT performed by the same evaluator, on different days, were analyzed, graphs (Figure 1C and 1D) characterize that the 6MWT was reproducible, with ICC = 0.727 (p < 0.001) and ICC = 0.590 (p < 0.001), and the limit of agreement between both 6MWTs conducted by evaluator 1 ranged from -104.23 to 130.5 m, and by evaluator 2 it ranged from -129.54 to 149.28 m.

Table 1. Sample characterization according to age, biometric and spirometric data

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean±SD</th>
<th>95%CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>10.28±2.25</td>
<td>9.42-11.13</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>40.73±2.375</td>
<td>35.99-45.41</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>143.35±12.914</td>
<td>140.20-151.41</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>19.01±3.40</td>
<td>17.72-20.31</td>
</tr>
<tr>
<td>FEV₁ (%)</td>
<td>93±4.57</td>
<td>89.40-96.60</td>
</tr>
<tr>
<td>FVC (%)</td>
<td>98.48±10.322</td>
<td>94.96-102.41</td>
</tr>
<tr>
<td>FEV₁/FVC%</td>
<td>86.41±5.308</td>
<td>84.39-88.43</td>
</tr>
</tbody>
</table>

SD: standard deviation; BMI: body mass index; 95%CI: 95% confidence interval; FEV₁: forced expiratory volume in the first second; FVC: forced vital capacity, %: percentage of the predicted

Table 2. Walking distance in the four performed six-minute walking (two on the same day and two others at a second evaluation)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Walking distance (Mean±SD)</th>
<th>95%CI</th>
<th>Standard-error</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD6MWTmin1 – day 1</td>
<td>569.59±86.96 m</td>
<td>537.94-601.24 m*</td>
<td>16.14</td>
</tr>
<tr>
<td>WD6MWTmin2 – day 1</td>
<td>564.06±80.85 m</td>
<td>521.92-606.2 m*</td>
<td>15.01</td>
</tr>
<tr>
<td>WD6MWTmin1 – day 2</td>
<td>556.45±74.63 m</td>
<td>529.3-583.6 m*</td>
<td>13.85</td>
</tr>
<tr>
<td>WD6MWTmin2 – day 2</td>
<td>554.19±76.19 m</td>
<td>526.4-581.9 m*</td>
<td>14.41</td>
</tr>
<tr>
<td>∆WD6MWTmin1 – day 1</td>
<td>-5.52 m</td>
<td>-28.475-17.417 m</td>
<td>11.20</td>
</tr>
<tr>
<td>∆WD6MWTmin2 – day 2</td>
<td>-2.26 m</td>
<td>-28.503-23.982 m</td>
<td>12.81</td>
</tr>
</tbody>
</table>

SD: standard deviation; 95%CI: 95% confidence interval; WD6MWTmin1: walking distance in the six-minute walk test in the first 6MWT; DPT6min2: walking distance in the six-minute walk test in the second 6MWT; ∆DPT6min: walking distance in the six-minute walk test on each one of the two days and 95% confidence interval of the ∆ of the WD on each of the two days (learning effect); m: meters; *there was no statistically significant difference between the four walked distances in the 6MWT in the analysis of variance (F=697; p=0.554)

Table 3. Intraclass correlation coefficient between the walking distances in the six-minute walking tests

<table>
<thead>
<tr>
<th>ICC (95%CI)</th>
<th>Standard-error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1 – WDE1 versus WDE2</td>
<td>0.742 (0.520-0.870)</td>
<td>0.0177</td>
</tr>
<tr>
<td>D2 – WDE1 versus WDE2</td>
<td>0.581 (0.279-0.779)</td>
<td>0.0178</td>
</tr>
<tr>
<td>WDE1 – D1 versus D2</td>
<td>0.727 (0.496-0.862)</td>
<td>0.0150</td>
</tr>
<tr>
<td>WDE2 – D1 versus D2</td>
<td>0.590 (0.291-0.784)</td>
<td>0.0195</td>
</tr>
</tbody>
</table>

ICC: intraclass correlation coefficient; D1: day one; D2: day two; WDE1: walking distance in the six-minute walk test performed by evaluator 1; WDE2: walking distance in the six-minute walk test performed by evaluator 2

DISCUSSION

This study analyzed the reproducibility and the reliability (between 2 evaluators) of the 6MWT performed on the same day, with a 30 minute interval, and after 14 days, one sample of Brazilian healthy children. The 6MWT proved to be reproducible for these children, and there was moderate to high correlation between the tests performed by both evaluators (1 and 2). The study by Li et al.² also assessed the interevaluator reliability; therefore, it analyzed the performance of 52 Chinese children, also healthy, in the 6MWT, which also respected the rules established by ATS, being repeated after a two-week interval. The mean age of the participants was superior to that in this study (14.2±1.2 versus 10.28±2.25 years old), which can justify longer WD (659.8±58.1 versus 561.2±38.6 m), besides ethnic differences. Two other studies⁹,¹⁰ also observed the reproducibility of the 6MWT in the pediatric population, involving ill participants. Cunha et al.⁹ analyzed 49 children and adolescents with cystic fibrosis (11.2±1.9), with moderate bronchial obstruction, and concluded that the 6MWT is reproducible, and that the WD can be related with clinical variables in the studied population. Morinder et al.¹⁰ assessed 16 obese children and adolescents (13.2 years old) and observed that the 6MWT is valid and reproducible (ICC = 0.84), and the values found by the researchers were considered to be highly reproducible, which corroborates the findings in this study, despite the differences regarding the included population.

With regard to the comparison of WD data on the first and on the second test on both days, this study showed there was no learning effect in the analyzed population of students, unlike the studies conducted with adults²,⁹,¹⁰,²²,₂₆, whose results show that the test repetition generates a learning effect, thus improving the performance of the individual. In this context, Rodrigues et al.⁷ observed that people with chronic obstructive pulmonary disease (COPD) presented longer WD in the second 6MWT, from the two that were conducted on alternate days. In that direction, Hernandes et al.⁸ also assessed the learning effect of the 6MWT in patients with COPD, by using 2 tests conducted on subsequent days, and observed that patients had better performances in the second test, with average increase of 27 m, thus characterizing a 7% learning effect. This pattern had been identified in the classic study by Knox et al.², which analyzed the learning effect of patients with chronic bronchitis in tests performed after consecutive days and weeks. The authors observed more learning effect when tests were repeated in short
intervals of time (33% increase in WD), in comparison to the course of consecutive weeks (8–5% increase in WD)\(^2\). In this investigation, in general, the performance of children did not improve with the repetition of the 6MWT. On the contrary: there was a tendency, even if not significant, to reduce the WD after each test. This child behavior reinforces the influence of the motivational aspect in this age group, once the matter of “new” represented here by the conduction of the first test, seems to have been essential in the quality of performance, thus reflecting the longest WD among the four analyzed tests. This happens because school-age children are particularly difficult to handle, and cannot be persuaded to a very prolonged active cooperation\(^13\). Associated with that, Berleza et al.\(^27\) they state that the motivation related to the satisfaction that is inherent to the activity of learning is affected when the child is obliged to perform any action, which is implicit in the command of any physical test, like the 6MWT.

Based on that, the result of the 6MWT presented here seems to have been influenced by the motivation generated by the news and the challenge, and not by the repetition (learning effect) of the activity. This discussion is extremely relevant, once the consensus of the 6MWT\(^3\) requires that two tests should be conducted to assess functional capacity, and this study shows the reproducibility of this number of executions. Therefore, the 6MWT proved to a reproducible test among children aged from 6 to 14 years old, and the population assessed here presented the tendency to perform better in the first test, which requires new lines of investigation. However, the possibility of influence of the second examiner in the results of this investigation should be mentioned, once, despite standardization, one different evaluator may change the children’s response, since they can feel more or less motivated.

One possible setback in this study is related to the level of understanding inherent to the age group, since the younger children had difficulties to perform the spirometry test in the process of sample selection. Even though the study included only students aged more than 6 years old, which is the minimum age to obtain the acceptability and reproducibility criteria of the exam, this event was observed. Besides, this factor, even though the practice of

---

**Figure 1.** Distribution of individual differences between the walking distance (WD) in the six-minute walking test (6MWT) and individual averages between distances on day 1 between evaluators 1 and 2 (A), on day 2 (B), and both days for evaluator 1 (C) and evaluator 2 (D). (A) Mean of differences=5.52 m (95%CI: 127.7–123.74 m). (B) Mean of differences=2.26 m (95%CI: 132.96–137.48 m). (C) Mean of differences=1313 m (95%CI: 104.23–130.5 m). (D) Mean of differences=9.86 m (95%CI: 129.54–149.28 m)
physical activities has been controlled by the application of the health survey, active children, with good physical performance, may have participated in this study, and that may also have influence the results.

CONCLUSION

The 6MWT demonstrated to be reproducible and reliable among healthy Brazilian children aged between 6 and 14 years old.

ACKNOWLEDGEMENTS

The authors would like to thank the research group from Núcleo de Fisioterapia em Pneumologia Pediatrica (NuFiPP), for the collaboration and support to collect data in this study, and to the members of Núcleo de Assistência, Ensino e Pesquisa em Reabilitação Pulmonar (NuReab), for the discussions and contributions in the statistical analysis of the study.

REFERENCES