



## ARTIGO ORIGINAL

# Testes funcionais validados em indivíduos hospitalizados e não hospitalizados: revisão sistemática

*Validated functional tests in hospitalized and non-hospitalized individuals: systematic review*

Natasha Cordeiro dos Santos<sup>1,2,\*</sup> , Neila Silva Soares<sup>1</sup> , Jorge Luis Motta dos Anjos<sup>1</sup> , Bruno Souza de Matos<sup>1</sup> , Daiane Barros Carvalho<sup>3</sup> 

<sup>1</sup>Hospital Geral Roberto Santos. Salvador, Bahia, Brasil.

<sup>2</sup>Departamento de Ciências da Vida, Universidade do Estado da Bahia. Salvador, Bahia, Brasil.

<sup>3</sup>Hospital da Mulher - Maria Luzia Costa dos Santos. Salvador, Bahia, Brasil.

Submetido em 22 de abril de 2020, aceito em 17 de outubro de 2020, publicado em 4 de dezembro de 2020

### PALAVRAS-CHAVE

Desempenho físico funcional  
Estudos de validação  
Reprodutibilidade dos testes

### RESUMO

**Objetivo:** Realizar uma revisão sistemática sobre os testes funcionais validados em diferentes perfis de indivíduos hospitalizados e não hospitalizados e avaliar as evidências psicométricas para confiabilidade e validade.

**Métodos:** Trata-se de uma revisão sistemática. Foram utilizadas as bases de dados EMBASE, MEDLINE, Lilacs e SciELO com as palavras-chave *Functional Tests (Walk Test, Gait Speed Test, Chair Stand Test, Timed Up And Go, Step Test), Validation Studies as Topic* e sinônimos. Foram incluídos estudos de validação de testes funcionais em indivíduos hospitalizados ou não hospitalizados que utilizaram os critérios de validade e/ou confiabilidade e relacionaram os testes a diferentes variáveis. Esses artigos poderiam ser observacionais longitudinais ou de corte transversal ou estudos de validação que utilizaram dados de ensaios clínicos. Foram excluídos os artigos que não apresentaram características importantes da amostra e a descrição do teste.

**Resultados:** A pesquisa resultou em 36.150 artigos, e 89 foram incluídos. Os estudos foram organizados em tabelas com informações como autor, ano; critérios de validação; amostra; teste; variáveis associadas; resultados. A qualidade dos artigos foi avaliada por meio da *Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies*.

**Conclusão:** Os testes funcionais são válidos e confiáveis para a avaliação de indivíduos hospitalizados e não hospitalizados, estando associados à força muscular, capacidade de caminhar, controle postural, atividades de vida diária, risco de quedas, hospitalização e mortalidade.

\*Autor de correspondência:

Rua Potiraguá, Edifício Vento Leste, 106. Pernambués. Salvador, BA, Brasil | CEP: 41100-160.

Fone: (71) 9 9311-9461

E-mail: [natasha-cordeiro@hotmail.com](mailto:natasha-cordeiro@hotmail.com) (Santos NC)

Este estudo foi realizado no Hospital Geral Roberto Santos

<https://doi.org/10.21876/rcshci.v10i4.960>

Como citar este artigo: Santos NC, Soares NS, Anjos JLM, Matos BS, Carvalho DB. Testes funcionais validados em indivíduos hospitalizados e não hospitalizados: revisão sistemática. Rev Cienc Saude. 2020;10(4):XX-XX. <https://doi.org/10.21876/rcshci.v10i4.960>

2236-3785/© 2020 Revista Ciências em Saúde. Este é um artigo de acesso aberto distribuído sob uma licença CC BY-NC-SA ([https://creativecommons.org/licenses/by-nc-sa/4.0/deed.pt\\_BR](https://creativecommons.org/licenses/by-nc-sa/4.0/deed.pt_BR))



**KEYWORDS**

Physical functional performance  
Reproducibility of results  
Validation studies

**ABSTRACT**

**Objective:** To conduct a systematic review of the functional tests validated in different profiles of hospitalized and non-hospitalized individuals and to evaluate the psychometric evidence for reliability and validity.

**Methods:** This is a systematic review. The databases EMBASE, MEDLINE, Lilacs and SciELO were used with the descriptors Functional Tests (Walk Test, Gait Speed Test, Chair Stand Test, Timed Up and Go, Step Test), Validation Studies as Topic and synonyms. Validation studies of functional tests were included in hospitalized or non-hospitalized individuals who used the criteria of validity and/or reliability and related the tests to different variables. These articles could be longitudinal or cross-sectional observational or validation studies that used data from clinical trials. Articles that did not present significant characteristics of the sample and the description of the test were excluded.

**Results:** The search resulted in 36,150 articles, and 89 were included. The studies were organized in tables with information such as author, year, validation criteria, sample, test, associated variables and results. The quality of the articles was assessed using the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies.

**Conclusion:** Functional tests are valid and reliable for assessing hospitalized and non-hospitalized individuals and are associated with muscle strength, ability to walk, postural control, activities of daily living, risk of falls, hospitalization, and mortality.

**INTRODUÇÃO**

A expectativa de vida humana vem aumentando rapidamente. No entanto, esse aumento não ocorreu de forma proporcional com a qualidade de vida e com a funcionalidade dos indivíduos. Atualmente, as doenças crônicas representam o maior desafio à saúde e são um foco predominante das políticas públicas, das pesquisas e da prática clínica. Sujeitos acometidos por essas morbidades são os maiores usuários dos serviços de saúde, utilizam um número maior de medicamentos, apresentam aumento do risco de desenvolver problemas de saúde mental e declínio funcional, muitas vezes associado a recorrentes hospitalizações<sup>1</sup>.

Há evidências de que indivíduos que foram hospitalizados por diversas condições clínicas apresentam comprometimento da função física, com perda de fibras musculares e aumento do metabolismo catabólico, além de disfunções dos sistemas fisiológicos, redução das taxas de retorno às atividades laborais e da capacidade funcional. Em virtude disso, é necessário avaliar de forma precoce e longitudinal as incapacidades do indivíduo por meio de testes funcionais, definidos como ferramentas dinâmicas utilizadas para avaliar a funcionalidade geral, abrangendo componentes como força muscular, controle neuromuscular, equilíbrio, capacidade de exercício e estabilidade articular. Para essa avaliação, deve-se selecionar instrumentos apropriados e validados, considerando suas propriedades psicométricas, pois existem diferentes fatores de interação, incluindo força, amplitude de movimento, propriocepção, equilíbrio, cognição; condições psicológicas; fatores ambientais e fatores clínicos que podem interferir na execução dos testes<sup>2,3</sup>.

Com o intuito de garantir a qualidade das avaliações, é necessário conhecer as formas de análise e as propriedades de medida. O desempenho dos resultados está relacionado principalmente aos conceitos de confiabilidade e validade. A confiabilidade, ou reprodutibilidade, é a capacidade de reproduzir um resultado de forma consistente no tempo e no espaço,

ou a partir de observadores diferentes. Os principais critérios da confiabilidade são estabilidade, consistência interna e equivalência. A validade de um teste, por sua vez, refere-se ao fato de medir exatamente o que se propõe a medir, podendo ser de conteúdo, construto e critério<sup>3</sup>.

Neste contexto, diante de uma população crescente de indivíduos com multimorbidades residentes na comunidade e de sobreviventes de recorrentes hospitalizações, que apresentam declínio funcional e piores desfechos clínicos, torna-se fundamental revisar as evidências psicométricas na literatura a respeito dos testes funcionais validados em diferentes perfis de indivíduos hospitalizados e não hospitalizados, assim como estudar suas vantagens e limitações. Através da adequada avaliação da funcionalidade é possível identificar os indivíduos com maiores riscos de perda funcional, assim como intervir com condutas mais apropriadas, e monitorar os resultados das intervenções<sup>2</sup>. Portanto, o objetivo do presente estudo foi realizar uma revisão sistemática sobre os testes funcionais validados em diferentes perfis de indivíduos hospitalizados e não hospitalizados e avaliar as evidências psicométricas para confiabilidade e validade.

**MÉTODOS**

Trata-se de uma revisão sistemática realizada conforme a Recomendação *MOOSE Statement (Meta-analysis of Observational Studies in Epidemiology)*. O protocolo das etapas de construção foi publicado na *International Prospective Register of Systematic Reviews (PROSPERO)* com registro CRD42020176918. Foram utilizadas as seguintes bases de dados: EMBASE, MEDLINE (*Medical Literature Analysis and Retrieval System Online*), Lilacs (*Literatura latina y del Caribe de Ciencias de la Salud*) e SciELO (*Scientific Electronic Library Online*). A coleta foi realizada de novembro de 2019 a fevereiro de 2020.

A elaboração desta pesquisa foi estruturada a partir

da estratégia PICO, um acrônimo para População (indivíduos hospitalizados e não hospitalizados), Interesse (testes funcionais: Teste de Caminhada de Seis Minutos, Teste de Velocidade da Marcha, Teste do Degrau, Deste de Sentar e Levantar e Teste *Timed Up and Go*) e Contexto (estudos de validação de testes funcionais que usaram critérios de validade e confiabilidade). Foram utilizadas as palavras-chave *Functional Tests (Walk Test, Gait Speed Test, Chair Stand Test, Timed Up And Go, Step Test)*, *Validation Studies as Topic* e os sinônimos de acordo com as bases de dados, identificados nos Descritores em Ciências da Saúde (DeCS), no *Medical Subject Headings (MeSH)* e *Embase Subject headings (Emtree)*, utilizando os operadores booleanos “AND” e “OR” como mostrado na Tabela 1. A busca foi realizada por meio das palavras encontradas nos títulos, assuntos e resumos dos artigos, sem restringir período de publicação.

Foram incluídos estudos de validação de testes funcionais em indivíduos hospitalizados ou não hospitalizados que utilizaram os critérios de validade e/ou confiabilidade e relacionaram os testes a diferentes variáveis. Esses artigos poderiam ser observacionais longitudinais ou de corte transversal ou estudos de validação que utilizaram dados de ensaios clínicos. Foram excluídos os estudos que não apresentaram características importantes da amostra estudada e a descrição padronizada da execução do teste.

Os artigos relevantes, coletados por meio de pesquisas em banco de dados, foram selecionados por rastreio dos títulos (primeira etapa), resumos (segunda etapa) e leitura integral (terceira etapa). Foi feita uma leitura exploratória de todo o material selecionado e, posteriormente, uma leitura seletiva e analítica mais aprofundada das partes que realmente interessavam. O processo de seleção, extração de dados dos artigos e de identificação dos aspectos metodológicos foi realizado por dois revisores independentes. Quando havia discordância entre eles, os revisores liam novamente o artigo para reavaliação. Se a divergência persistisse, um terceiro revisor independente avaliava e tomava a decisão final.

#### Análise dos critérios de confiabilidade e validade

- Confiabilidade
  - Estabilidade: consistência das repetições das medidas. Foi escolhido o método de teste-reteste<sup>4</sup>, e

para a análise o Coeficiente de Correlação Intraclasse (CCI) e os coeficientes de Pearson e de Spearman<sup>3</sup>.

- Consistência interna: indica se todas as subpartes de um instrumento medem a mesma característica. Foi escolhido para análise o coeficiente alfa de Cronbach<sup>3</sup>.
- Equivalência: grau de concordância entre dois ou mais examinadores quanto aos escores de um instrumento. O coeficiente Kappa foi a medida de análise escolhida<sup>4</sup>.
- Validade
  - Validade de conteúdo: grau em que o conteúdo de um instrumento reflete adequadamente o construto que está sendo medido. Uma abordagem qualitativa foi escolhida como método, seguida do uso do índice de validade de conteúdo (IVC)<sup>3</sup>.
  - Validade de critério: relação entre pontuações de um determinado instrumento e algum instrumento ou critério considerado padrão-ouro. Para critérios futuros, a validade preditiva é utilizada, e para contemporâneos, a concorrente. Foi escolhido para a análise o coeficiente de correlação interclasse<sup>3</sup>.
  - Validade de construto: se o conjunto de variáveis representa o construto a ser medido. Subdividida em teste de hipóteses, validade estrutural e validade transcultural. No teste de hipóteses, grupos diferentes preenchem o instrumento e os resultados são comparados. Na validade discriminante testa com construtos diferentes, com variáveis das quais deveria divergir<sup>4</sup>. Foram escolhidos os coeficientes de correlação de Pearson e Spearman.

#### Análise Metodológica

A qualidade dos estudos incluídos foi avaliada pelo uso da *Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies* da *National Heart, Lung, and Blood Institute (NHLBI)*<sup>5</sup>.

## RESULTADOS

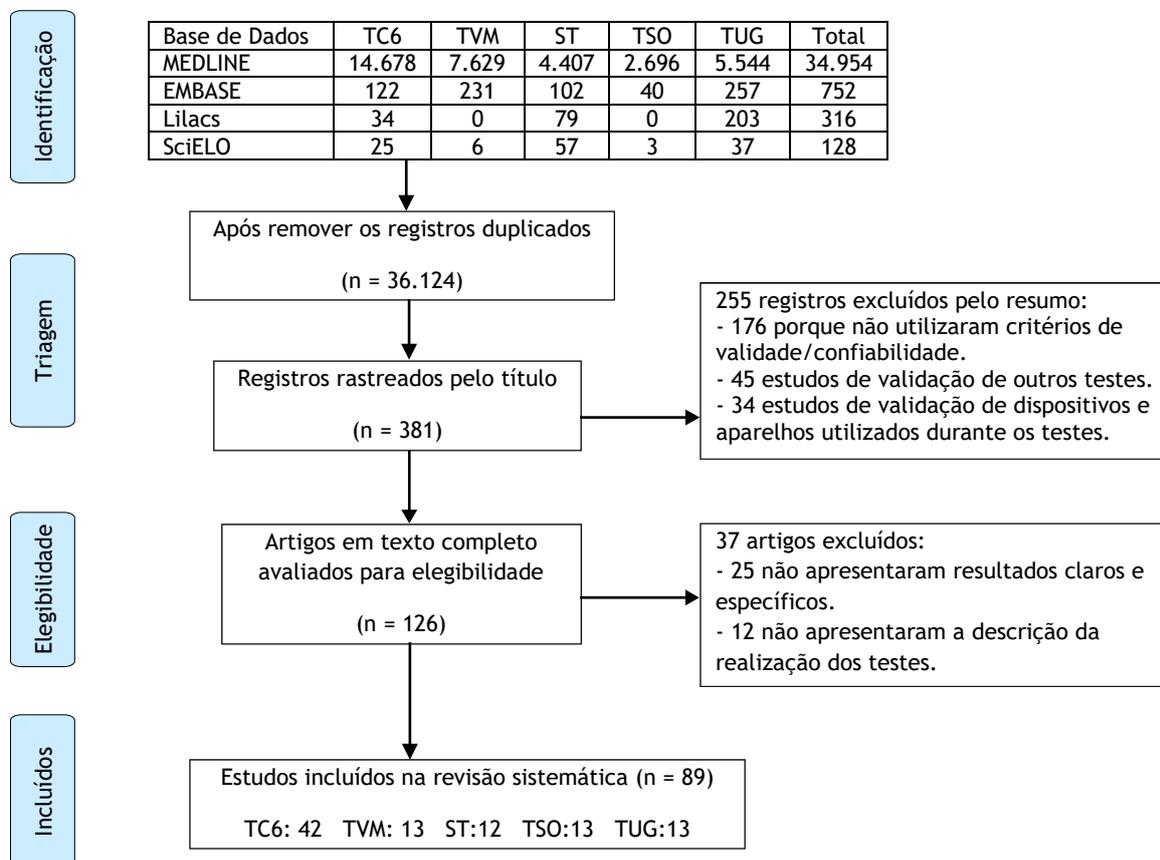
A pesquisa nas bases de dados resultou em 36.150 artigos (MEDLINE: 34.954; EMBASE: 752; Lilacs: 316; SciELO: 128), ficando 36.124 registros após remover os duplicados. Destes, 381 artigos foram selecionados por rastreio de títulos; 255 foram excluídos devido ao

**Tabela 1** - Palavras-chave utilizadas na busca eletrônica acrescidas dos operadores booleanos “AND” e “OR”.

Palavras-chave	Sinônimos
<b>Functional Tests</b>	
<i>Six-minute walk test</i>	<i>6-Minute Walk Test; 6 Minute Walk Test; 6-Minute Walk Tests; Test, 6-Minute Walk; Tests, 6-Minute Walk; Walk Test, 6-Minute; Walk Tests, 6-Minute; six-minute walk test; Walk Test; Walk Tests</i>
<i>Gait Speed Test</i>	<i>Gait Speed Tests; walking speed test; walking speed tests</i>
<i>Step Test</i>	<i>Step Tests</i>
<i>Chair Stand Test</i>	<i>Chair Stand Tests; Sit to Stand Test; Sit to Stand Tests</i>
<i>Timed Up and Go Test</i>	<i>TUG</i>
<b>Validation Studies as Topic</b>	<i>Validation Studies; Validation Study; Validation; validity</i>

resumo não atender aos critérios de seleção, restando inclusão, resultando na seleção final de 89 artigos (três estudos abordaram mais de um teste), conforme Figura 1. Os 89 artigos foram lidos de forma analítica e seletiva

e organizados em cinco tabelas com informações relevantes das pesquisas a respeito de cada teste: autor, ano; critérios de validação; amostra; teste; variáveis associadas; resultados (Tabelas 2, 3, 4, 5 e 6).



**Figura 1** - Diagrama de fluxo da pesquisa. TC6: Teste de caminhada de 6 min; TVM: Teste de velocidade de marcha; ST: teste do degrau; TSO: Teste de Sentar e Levantar; TUG: Teste *Timed Up and Go*.

### Teste de Caminhada de Seis Minutos (TC6)

Dos 42 artigos, 27 eram estudos observacionais longitudinais<sup>6,7,9-11,13,15-20,22,24,27,29-31,35-38,40,41,43,45,46</sup>, nove observacionais de corte transversal<sup>14,21,23,25,26,39,42,44,47</sup> e seis utilizaram dados de ensaios clínicos<sup>8,12,28,32-34</sup>. Apenas três foram realizados com indivíduos hospitalizados, sendo um em sujeitos com lesão traumática cerebral<sup>30</sup>, outro em indivíduos durante ou fora do tratamento do câncer<sup>40</sup> e outro em sujeitos com doença renal terminal<sup>44</sup>. Em relação à execução dos testes, foram realizados em corredores de 70 m<sup>30,34</sup>, 50 m<sup>6</sup>, 40 m<sup>14,22,47</sup>, 30 m<sup>11,17,24,26,31,38,40,42,47</sup>, 25 m<sup>13,19,39</sup>, 20 m<sup>15,21,28,33,43,44</sup>, 15 m<sup>26,35</sup>, 10 m<sup>41</sup>, 8 m<sup>20</sup>, um campo 40 m x 20 m<sup>46</sup>, em pista retangular 10 m x 50 m<sup>41</sup>, de 12 m x 6 m<sup>30</sup>, em corredor circular de 125 m<sup>25</sup> e em percurso oval de 150 e 250 pés<sup>17</sup>. Dezenove estudos seguiram as diretrizes da *American Thoracic Society* (ATS)<sup>8,9,11,13,14,17,18,22,23,27,29,31,35,36,39,40,42,46,47</sup>. Outros artigos não utilizaram o mesmo protocolo, mas forneceram incentivos com frases

padronizadas<sup>10,12,15,16,20,24,30,31,33,35,37,38,44,45</sup> e informações sobre o tempo restante para o teste terminar<sup>12,29,30,33,37,43,45</sup>. Apenas Liu et al.<sup>25</sup> informaram o tempo de descanso permitido entre dois testes consecutivos de pelo menos 60 min.

O TC6 mostrou-se como um instrumento fácil<sup>17</sup>, barato<sup>17</sup>, válido<sup>6,9-11,13,15,16,24,31,35,39</sup> e confiável<sup>7,8,10,12-15,17-19,24,25,27,28,31,33,36-38,40,41,45,46</sup>. Apresentou correlações com composição corporal, fração de ejeção, ansiedade e depressão<sup>45</sup>, função pulmonar<sup>32,39,45</sup>, carga de trabalho máxima (Wmax)<sup>20</sup>, incapacidade e dor<sup>36</sup>, atividades da vida diária (AVDs)<sup>36</sup>, capacidade de difusão de monóxido de carbono<sup>45</sup> e custo energético<sup>21</sup>.

O teste também demonstrou associação com outros instrumentos como o *Short Form 36*<sup>11</sup>, Escalas *Radiographic Global Impression of Change* e de gravidade de Rickets<sup>36</sup>, *25-foot walk test* (25FWT)<sup>39</sup>, *Timed Up and Go* (TUG)<sup>14,37,39</sup>, contração dos membros inferiores (MVIC)<sup>39</sup>, escala funcional de avaliação da Esclerose Lateral Amiotrófica<sup>39</sup>, teste muscular dos membros inferiores<sup>13</sup>, miometria de flexão do joelho<sup>13,27</sup>,

FSHD *Clinical Score*<sup>14</sup>, Medida de Independência Funcional (MIF) de locomoção (caminhada) e motora<sup>17</sup> e *New York Heart Association* (NYHA)<sup>11,12,45</sup>. O TC6 mostrou-se preditor de pico de volume de oxigênio máximo (VO<sub>2</sub>pico)<sup>6,13,14,17,18,31,33,40,45</sup>, consumo máximo de oxigênio (VO<sub>2</sub>máx)<sup>20,23,26,46</sup>, hospitalização<sup>9,32</sup>, tempo de internamento<sup>28</sup>, morte<sup>6,9,32</sup> e qualidade de vida<sup>9,32</sup>.

Dentre os estudos, sete demonstraram baixa correlação entre o TC6 e algumas das variáveis analisadas: capacidade vital forçada, índice de dispneia de Borg, capacidade de difusão em indivíduos com doença pulmonar intersticial secundária à esclerose sistêmica<sup>8</sup>, consumo máximo de oxigênio (VO<sub>2</sub>máx) em crianças obesas e com peso normal<sup>29</sup> e sujeitos com fibromialgia<sup>34</sup>, pico de volume de oxigênio máximo (VO<sub>2</sub>pico) em crianças com doença renal terminal<sup>44</sup>; função fisiológica e qualidade de vida na Fibrose Pulmonar Idiopática<sup>7</sup>, *Early Activity Scale for Endurance* (EASE) de crianças com paralisia cerebral<sup>16</sup>, Estágios Hoehn e Yahr, escala *Berg Balance* (BBS) e cadência em indivíduos com doença de Parkinson moderada<sup>21</sup>. Destes, apenas um foi realizado com indivíduos hospitalizados<sup>34</sup> (Tabela 2).

#### Teste de Velocidade da Marcha (TVM)

Dos 13 artigos, 10 eram estudos observacionais longitudinais<sup>9,37,48-52,54,55,57</sup> e três observacionais de corte transversal<sup>53,56,58</sup>. Apenas três foram realizados com indivíduos hospitalizados, sendo um em sujeitos com 65 anos ou mais<sup>48</sup>, outro em adultos após insuficiência respiratória aguda<sup>9</sup> e outro em indivíduos com lesão cerebral traumática<sup>58</sup>. Quanto à execução do teste, foram realizados em corredor de 18 m<sup>52</sup>, 14 m<sup>58</sup>, 10 m<sup>48,50,51,53</sup>, 6 m<sup>53</sup>, 4 m<sup>9,49,53-57</sup> e em um tapete de 4,47 m com *software*<sup>37</sup>. Em alguns artigos os indivíduos foram instruídos a caminhar com a velocidade habitual<sup>37,48,51,53,56,57</sup> e não o mais rápido possível como nos demais estudos.

Reychler et al.<sup>57</sup> forneceram incentivos com frases padronizadas. Em algumas pesquisas poderia descansar se necessário<sup>53,56</sup>, e em outras era permitido utilizar auxiliar de marcha<sup>48,54</sup> e suporte suplementar de oxigênio<sup>54</sup>. O teste também foi realizado três vezes<sup>49,52,53,56</sup> ou repetido uma vez<sup>9,50,55,57</sup>. No estudo de Bloem et al.<sup>49</sup> foi dado um tempo de descanso de dois minutos, no de Reychler et al.<sup>57</sup> não puderam descansar e em outros o intervalo era o mais breve possível<sup>51,55,56</sup>.

O TVM apresentou-se como um teste simples, responsivo<sup>55</sup>, válido<sup>9,49,51,53,55,58</sup> e confiável<sup>9,37,48,50,51,54-56,58</sup>. Apresentou correlações com capacidade de exercício<sup>9,54</sup>, força de preensão manual<sup>49</sup>, dispneia<sup>49,54</sup>, massa muscular, ritmo de caminhada, autoavaliação da saúde e depressão<sup>53</sup>. O teste também previu hospitalizações<sup>9</sup>, qualidade de vida<sup>9,51,54</sup>, tempo de internação, alta exigindo maior nível de cuidados e *delirium*<sup>48</sup>. O instrumento demonstrou associação com o Teste de Sentar e Levantar 5 vezes<sup>49</sup>.

Dos 13 estudos, apenas dois mostraram baixas correlações com algumas variáveis estudadas: o de Peters et al.<sup>56</sup>, no qual o TVM de 4 m (TVM4m) não apresentou um grau alto de validade concorrente com o de 10 m em idosos saudáveis; e o de Reychler et al.<sup>57</sup> que não mostrou correlação entre o Teste de Degrau de 6

Minutos (ST6) e TVM4m em crianças, ambos estudos foram realizados com indivíduos não hospitalizados (Tabela 3).

#### Teste do Degrau (ST)

Dos 12 artigos, quatro eram estudos observacionais longitudinais<sup>60,62,63,67</sup> e oito observacionais de corte transversal<sup>57,59,61,64-66,68,69</sup>. Nenhum estudo foi realizado em indivíduos hospitalizados. Quanto à execução do teste, foram utilizados degraus de 25,4 cm<sup>62</sup>, 20 cm<sup>57,59,67</sup>, 18 cm<sup>61</sup>, 15 cm<sup>68</sup>, 10 cm<sup>66</sup>, 7,5 cm<sup>64</sup>, dois degraus de 20 cm<sup>65</sup>, três degraus de 20, 75 e 40 cm<sup>67</sup> e outro de 0,1, 0,2 e 0,3 m<sup>63</sup>. A maioria foi realizada em 6 min<sup>57,59,67-69</sup>, outros em 2 min em cada altura<sup>63</sup>, 3 min<sup>62</sup>, 15 s<sup>64</sup> e realizados 8 vezes<sup>61</sup>. Em relação à repetição, quatro foram executados duas vezes<sup>61,63,66,69</sup> e dois três vezes<sup>62,64</sup>, com intervalos de 30 min de descanso<sup>68</sup>, de 15 min<sup>63</sup>, de 1 min<sup>62,63</sup>, ou sem descanso<sup>66</sup>. Alguns utilizaram incentivos com frases padronizadas baseadas na ATS para o TC6<sup>57,67-69</sup>.

O ST é uma ferramenta de avaliação da aptidão cardiorrespiratória<sup>65</sup>, válida<sup>62,66,67,69</sup> e confiável<sup>59,61-64,66,69</sup>. Apresentou correlações com valores metabólicos e cardiovasculares<sup>68</sup>, VO<sub>2</sub>max<sup>62,63,65</sup> e VO<sub>2</sub>pico<sup>62,68</sup>. O teste também demonstrou associação com outros instrumentos como *Fugl-Meyer Assessment-Lower Extremity* (FMA-LE)<sup>62</sup>, TSO de 5 vezes, *Berg Balance Scale* (BBS), TUG<sup>61</sup>, tempo de reação motora (MRT), a força isométrica do extensor do joelho e o tempo de permanência em uma perna (SLST)<sup>66</sup>.

O teste não apresentou algum resultado satisfatório em três estudos, no de Bonnevie et al.<sup>60</sup> o uso do ST6 e da carga máxima de trabalho obtida durante o teste (Wpeak) na prescrição do treinamento de força resultou em um erro clinicamente não aceitável em indivíduos com Doença Pulmonar Obstrutiva Crônica (DPOC) não hospitalizados. Enquanto que no artigo de Reychler et al.<sup>57</sup>, o ST6 e o TVM4m não se apresentaram como substitutos válidos para o TC6 em crianças saudáveis (Tabela 4).

#### Teste de Sentar e Levantar (TSO)

Dos 13 artigos, 10 eram estudos observacionais longitudinais<sup>70-75,77,78,80,82</sup>, dois observacionais de corte transversal<sup>76,79</sup> e um utilizou dados de ensaios clínicos<sup>81</sup>. Apenas dois foram realizados em indivíduos hospitalizados, sendo um no departamento de emergência<sup>70</sup> e outro em adultos mais velhos<sup>77</sup>. Quanto à execução do teste, foram utilizadas cadeiras de 48 cm<sup>74,81</sup>, 46 - 48 cm<sup>71</sup>, 46 cm<sup>75</sup>, 45,7 cm<sup>77</sup>, 45 cm<sup>78</sup>, 43,2 cm<sup>73</sup>, 43 cm<sup>79</sup>, cadeira ajustável para manter os quadris a 90 graus<sup>76,82</sup> e sem apoio para os braços<sup>70-75,79</sup>. Em apenas três artigos os indivíduos puderam usar os braços<sup>76,77,80</sup>. A duração dos testes variou entre um minuto<sup>71,75</sup>, 30 s<sup>70,74,77</sup>; repetição em 5 vezes<sup>72,74,75,78,77,82</sup> e 10 vezes<sup>80</sup>. No artigo de Staartjes et al.<sup>81</sup> foi realizada uma familiarização com o teste, sendo executado sem cronometrar. O teste foi realizado três vezes<sup>76,79,82</sup> e duas vezes<sup>78</sup>. Kumban et al.<sup>76</sup> e Mong et al.<sup>79</sup> utilizaram a média dos resultados. O tempo de intervalo para descanso variou em 3 min<sup>78</sup>, 2 min<sup>82</sup> e 1 min<sup>76,79</sup>.

O TSO é um teste válido<sup>71-74,75-77,80,82</sup> e confiável<sup>71-</sup>

<sup>74,76,79,81,82</sup>. Apresentou correlações com VO<sub>2</sub>máx, produção de dióxido de carbono<sup>71</sup> e capacidade funcional e de equilíbrio<sup>76</sup>, força dos flexores e extensores de joelho<sup>75,79</sup>. O teste também demonstrou associação com outros instrumentos como TC6<sup>71,75</sup>, *Berg Balance Scale* (BBS), Índice de Barthel modificado<sup>77</sup>, *Roland-Morris Disability Questionnaire* (RMDQ), *Oswestry Disability Index* (ODI), escala analógica visual (EVA) e Escala Europeia de Qualidade de Vida (EQ-5D)<sup>81</sup>.

Em duas pesquisas os indivíduos foram informados sobre o tempo 15 s antes de terminar o teste<sup>71,75</sup>. Apenas Nilsagård et al.<sup>80</sup> apresentaram correlações não significativas quanto ao TSO, especificamente em relação ao número de quedas ou incidentes próximos em indivíduos com esclerose múltipla não hospitalizados (Tabela 5).

### Teste Timed Up and Go (TUG)

Dos 13 artigos, quatro eram estudos observacionais longitudinais<sup>37,86,89,93</sup>, oito observacionais de corte transversal<sup>83-85,88,90-92,94</sup> e um utilizou dados de ensaios clínicos<sup>87</sup>. Apenas dois foram realizados com indivíduos hospitalizados, sendo um em sujeitos com insuficiência cardíaca crônica<sup>87</sup>, e outro com falência crônica avançada de órgãos (doença pulmonar obstrutiva crônica, insuficiência cardíaca e insuficiência renal)<sup>90</sup>. O Protocolo desenvolvido por Podsiadlo e Richardson foi utilizado por três artigos<sup>83,85,90</sup>. Foram usadas cadeiras de 46 cm<sup>85,93</sup>, 45 cm<sup>87,89</sup>, e 43 cm<sup>84</sup>. Em alguns estudos os indivíduos puderam utilizar as mãos<sup>92</sup>, auxílio de marcha<sup>86-88,90,92-94</sup> e suporte de oxigênio<sup>90</sup>.

Nicolini-panisson et al.<sup>91</sup> realizaram treinamento dos participantes antes da execução do teste. Em alguns artigos foram executados duas vezes<sup>37,83,87,89,94</sup> e outros três vezes<sup>84,90</sup>, com descanso de 30 a 60 min<sup>37</sup> ou até a resolução dos sintomas<sup>87,90</sup>. Sebastião et al.<sup>94</sup> usaram a média das duas repetições.

O TUG mostrou-se como um teste confiável<sup>37,83,84,86,87,89-91,93</sup> e válido<sup>83-89,92-94</sup>. Apresentou correlações com status de incapacidade, equilíbrio, controle postural<sup>94</sup> e força de abdutor do quadril, flexor do joelho, dorsiflexor do tornozelo e flexor plantar<sup>84</sup>. O teste também demonstrou associação com outros instrumentos como Avaliação de Fugl-Meyer para membros inferiores e *Berg Balance Scale* (BBS)<sup>84</sup>, Teste Ativo de Elevação da Perna (ASLR), Questionário de Cintura Pélvica (PGQ)<sup>85</sup>, Escala Visual Analógica (EVA), Índice de Incapacidade Roland-Morris, Índice de Incapacidade de Oswestry, Escala Europeia de Qualidade de Vida (EQ-5D)<sup>86</sup>, Índice de Incapacidade do Questionário de Avaliação de Saúde (HAQ-DI), Escala Europeia de Qualidade de Vida (EQ-5D)<sup>89</sup> e medida da função motora grossa (GMFM)<sup>91</sup>.

Apenas Rolenz et al.<sup>92</sup> mostrou uma baixa sensibilidade do TUG de apenas 20% comparado ao *8-Foot Up and Go* (8UG) de 64%. O TUG identificou quedas em taxas diferentes das do 8UG ( $p < 0,05$ ) em idosos não hospitalizados com e sem comprometimento cognitivo (Tabela 6).

### Qualidade metodológica

Todos os artigos selecionados apresentaram a

questão da pesquisa ou objetivo de forma clara, a população foi especificada e recrutada de forma semelhante de acordo com os critérios de seleção e a taxa de participação das pessoas elegíveis foi de pelo menos 50%. Quarenta e dois artigos apresentaram uma justificativa do tamanho da amostra<sup>7,9,10,12,14,15,16,21,24,26,28,30,31,33,35-37,40,42,46-48,51,57,59,62,64,65,67-70,72,77,79,81,83-85,88,89,92,94</sup>.

Apenas cinco estudos<sup>7,11,12,28,32</sup> analisaram as exposições de interesse antes do resultado a ser medido e o prazo foi suficiente para esperar uma associação. Nos 89 estudos foram avaliados diferentes níveis de exposição quando relacionados ao resultado, e as medidas de exposição e de resultados foram claramente definidas, válidas, confiáveis e implementadas de forma consistente. As principais variáveis foram medidas e ajustadas estatisticamente por seu impacto na relação entre exposição e resultado em todos os estudos. Apenas nove artigos<sup>10,37,41,54,57,59,84,85,89</sup> utilizaram cegamento. Por fim, 46 estudos foram classificados com boa qualidade metodológica e nenhum com uma qualidade ruim (Tabela 7).

## DISCUSSÃO

A presente revisão sistemática sugere que o Teste de Caminhada de Seis Minutos, o Teste de Velocidade da Marcha, o Teste do Degrau, o Teste de Sentar e Levantar e o Timed Up and Go Test são instrumentos válidos e confiáveis para avaliar incapacidades, para diagnosticar as disfunções do movimento, e para gerenciar e aprimorar as habilidades físicas e funcionais de indivíduos hospitalizados e não hospitalizados.

Em cada perfil de sujeitos é fundamental considerar questões específicas para a adequada escolha do teste a ser utilizado, como idade, nível de consciência, capacidade de seguir instruções, motivação, fraqueza, fadigabilidade, e a própria condição clínica do sujeito<sup>2</sup>, pois selecionar uma medida inadequada de resultado pode levar a conclusões incorretas sobre a condição do indivíduo e sobre os benefícios do tratamento<sup>2,95,96</sup>. Entretanto, os testes funcionais possuem algumas limitações. De modo geral, existe uma ampla variabilidade no esforço dispendido durante a execução dos testes, podendo variar devido aos incentivos fornecidos por diferentes examinadores e às diferentes interpretações das instruções, por exemplo, há uma conotação inerentemente subjetiva da frase “velocidade normal de caminhada”. Alguns participantes interpretam como uma caminhada rápida, enquanto outros interpretam “normal” como um ritmo mais lento<sup>97-99</sup>.

Em associação, há controvérsias na padronização dos testes, como pode-se observar nos artigos incluídos nesta revisão, existe uma diversidade de protocolos utilizados. Há também uma falta de valores de referência para determinar o que seria uma baixa funcionalidade; e uma relativa escassez de informações sobre mecanismos de limitação ao esforço. Os resultados podem ser afetados por uma variedade de fatores não relacionados ao status cardiopulmonar, incluindo idade, sexo, altura e peso, além de causas ortopédicas e outras comorbidades<sup>97</sup>. Essas condições justificam os resultados não satisfatórios em alguns estudos.

**Tabela 2** - Características dos estudos sobre Teste de Caminhada de Seis Minutos (TC6) (n = 42).

Autor, ano	Crítérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Boer PH, 2016 <sup>6</sup>	Validade de critério.	43 indivíduos com Síndrome de Down (SD) (não hospitalizados).	Os indivíduos foram instruídos a andar tantas voltas quanto possível em um corredor de 50 m dentro dos 6 min. Puderam descansar a qualquer momento, mas foram encorajados a continuar o mais rápido possível.	PACER de 16 m, VO <sub>2</sub> pico, IMC, idade e sexo.	- A regressão linear revelou que o PACER (R <sup>2</sup> = 0,86) e o TC6 (R <sup>2</sup> = 0,75) estavam significativamente relacionados ao VO <sub>2</sub> pico (p < 0,05). - A correlação entre o VO <sub>2</sub> pico e os dois testes, controlando idade, sexo e IMC, permaneceu significativa (r > 0,7; p < 0,05). - PACER, TC6 e IMC são preditores significativos do VO <sub>2</sub> pico (p < 0,05).
Bois RM, 2011 <sup>7</sup>	Confiabilidade (estabilidade), validade de critério; de construto.	822 indivíduos com Fibrose Pulmonar Idiopática (FPI) (não hospitalizados).	Instruídos a caminhar o mais longe possível sem correr; desacelerar ou parar se necessário, e retomar assim que fosse capaz. Interrompido se apresentasse dor no peito, dispneia intolerável, câibras, diaforese ou dessaturação abaixo de 83%. Realizada titulação de oxigênio.	CVF, DLCO; <i>University of California San Diego Shortness of Breath Questionnaire</i> ; <i>St. George's Respiratory questionnaire</i> e morte.	- Boa confiabilidade (CCI = 0,83; p < 0,001). A MCID foi de 24 a 45 m. - Pouco correlacionado com medidas da função fisiológica e qualidade de vida (menores para os indivíduos com pior estado funcional). - Altamente preditiva de mortalidade: um declínio de 24 semanas maior que 50 m foi associado a um aumento de quatro vezes no risco de morte em um ano ( <i>Hazard Ratio</i> : 4,27; IC 95%: 2,57 - 7,10; p < 0,001).
Buch MH, 2007 <sup>8</sup>	Confiabilidade (estabilidade).	163 indivíduos com doença pulmonar intersticial secundária à esclerose sistêmica (não hospitalizados).	De acordo com a ATS. Instrução: "O objetivo é caminhar o mais longe possível por 6 min. Você vai e volta neste corredor. Seis min é muito tempo para caminhar, você estará se exercitando. Provavelmente ficará sem fôlego ou exausto. Você tem permissão para desacelerar, parar e descansar. Pode encostar na parede enquanto descansar, mas volte assim que puder".	Índice de dispneia de Borg; teste de função pulmonar (CVF, capacidade de difusão de monóxido de carbono).	- A distância média (DP) do TC6 1 foi de 396,6 (84,55) m e do TC6 2 de 399,5 (86,28) m. - A correlação intra-sujeito, determinada pelo teste do coeficiente de correlação de Pearson, foi de 0,95 (p < 0,001). - Correlações fracas com o percentual de capacidade vital forçada e o Índice de Dispneia de Borg; nenhuma com capacidade de difusão. - A falta de correlação com parâmetros fisiológicos sugere uma base multifatorial para a capacidade limitada e põe em questão a utilidade.
Chan KS, 2015 <sup>9</sup>	Validade de construto (convergente e discriminante); critério (preditiva).	641 indivíduos sobreviventes à Insuficiência Respiratória Aguda (não hospitalizados).	A medida primária do estudo TC6 foi baseada nas diretrizes da ATS com variações modestas, incluindo o uso de um único TC6 em cada acompanhamento e a maior distância disponível durante as visitas domiciliares.	Saúde física e mental, risco de morte, hospitalização e qualidade de vida relacionada à saúde.	- Boa validade convergente e discriminante, correlações moderadas a fortes com medidas de saúde física ( r  = 0,36 - 0,76) e mais fracas com medidas de saúde mental ( r  = 0,03 - 0,45). - A validade dos grupos foi demonstrada pelas diferenças entre os grupos com força muscular e função pulmonar diferentes (p < 0,01). - Previu mortalidade futura, hospitalização e qualidade de vida.
Chan AW, 2019 <sup>10</sup>	Confiabilidade (estabilidade); validade de construto; de conteúdo (teste de hipóteses).	39 idosos com demência ou doença de Alzheimer (não hospitalizados).	Seguindo a diretriz de Crapo <i>et al.</i> (2002). Foram convidados a caminhar "o mais longe possível" nos 6 min, correr não foi permitido. Utilizou-se incentivo padronizado: "você está indo bem, continue assim", a cada minuto.	Teste de Caminhada de 2 min (TC2), de velocidade da marcha de 10m (TVM10m), Escala de Equilíbrio de Berg (BBS).	- Confiabilidade teste-reteste (CCI = 0,91 - 0,98) e interexaminadores (CCI = 0,86 - 0,96). Os testes foram correlacionados entre si (p = 0,85 - 0,94) e foram capazes de distinguir quem podia caminhar ao ar livre e no interior (p ≤ 0,036). O IC de 95% foi 9,1 m no TC2, 28,1 m no TC6 e 16 m/s no TVM10m. Correlações com medidas funcionais (p = 0,34 - 0,55).
Chen YC, 2018 <sup>11</sup>	Validade de construto (convergente e discriminante).	125 indivíduos agendados para cirurgia cardíaca (não hospitalizados).	De acordo com a ATS. Os indivíduos foram convidados a caminhar em um corredor de 30 m, e sua distância foi calculada. Para compensar um possível efeito de aprendizado, completaram o TC6 duas vezes e a maior distância foi selecionada.	<i>36-Item Short Form Health Survey</i> (SF-36), Classificação Funcional da <i>New York Heart Association</i> (NYHA).	- O TC6 médio foi de 308,9 ± 77,3 m no pré-operatório, 277,3 ± 85,7 m na alta e 378,1 ± 95,2 m em três meses. As distâncias na linha de base e em três meses tiveram moderada a alta correlação com o funcionamento físico do SF-36 (rs = 0,44 e 0,54). Maiores níveis da NYHA: menor TC6. - Sensível à alteração no perioperatório (tamanho de efeito -0,51 a 1,72).

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico= pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 2** - Características dos estudos sobre Teste de Caminhada de Seis Minutos (TC6) (n = 42) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Demers C, 2001 <sup>12</sup>	Confiabilidade (estabilidade); validade de construto.	768 indivíduos com Insuficiência Cardíaca (IC) (não hospitalizados)	Instruídos a caminhar no seu ritmo enquanto tentavam percorrer a maior distância. O tempo foi informado a cada 2 min. O incentivo a cada 30 s foi “Você está indo bem” ou “Continue com o bom trabalho”. Com permissão para descansar, parar e continuar assim que pudessem.	Classificação Funcional da <i>New York Heart Association</i> (NYHA-FC), qualidade de vida.	- A confiabilidade teste-reteste na linha de base (CCI = 0,90), 18 semanas (CCI = 0,88) e 43 semanas (CCI = 0,91) foi muito boa. - A distância na linha de base foi inversamente correlacionada com qualidade de vida (r = -0,26, p = 0,0001) e moderadamente inversamente com a NYHA-FC (r = -0,43, p = 0,001). - O TC6 e o NYHA-FC foram comparáveis, mas menos responsivos.
Dunaway YS, 2016 <sup>13</sup>	Validade de construto (convergente); confiabilidade (estabilidade).	30 indivíduos com atrofia muscular da coluna (não hospitalizados).	Após descanso de 10 min, percorreram um corredor de 25 m. Instruídos a andar o mais longe possível sem correr. Incentivo padronizado administrado com mesmos tons para fornecer orientação antes e durante de acordo com a ATS. Permitido descansar, sem sentar-se, se necessário.	<i>Hammersmith Functional Motor Scale Expanded</i> (HFMSE), teste muscular dos membros inferiores, miometria de flexão do joelho, VO <sub>2</sub> pico.	- A confiabilidade teste-reteste foi excelente para todos os participantes (CCI: 0,992; IC 95%, 0,979 - 0,997) com idades entre 10 e 49 anos. - Associação moderada entre TC6 e o VO <sub>2</sub> pico (r = 0,558; P = 0,038). - Associado positivamente ao VO <sub>2</sub> pico, HFMSE, teste muscular dos membros inferiores, miometria de flexão do joelho e inversamente com velocidade de marcha.
Eichinger K, 2017 <sup>14</sup>	Confiabilidade (estabilidade).	86 indivíduos com Distrofia Muscular Facio-escapulo-umeral (FSHD) (não hospitalizados).	Realizado de acordo com as diretrizes da ATS. Ambos os locais usaram um percurso reto, mas a duração do percurso variou entre os dois locais; um percurso de 40 m foi utilizado no Centro Médico da Universidade de Rochester (URMC) e um percurso de 50 m no Instituto Kennedy Krieger (KKI).	Teste Muscular Manual (MMT), <i>Timed Up and Go</i> (TUG), <i>30 foot Go</i> , <i>10 meter/walk run</i> , FSHD <i>Clinical Score</i> , gravidade da doença.	- Confiável (CCI = 0,99) e com correlações moderadas a fortes com força na extremidade inferior, resultados funcionais e FSHD <i>Clinical Score</i> . - A confiabilidade teste-reteste (n = 25; CCI = 0,99). - Relações moderadas a fortes com gravidade; forte com MMT da extremidade inferior (p = 0,79; P < 0,0001). - Correlacionado inversamente com o TUG (p = -0,81; p < 0,0001), <i>30 foot Go</i> (p = -0,88; p < 0,0001) e <i>10 meter/walk run</i> (p = -0,78; P < 0,0001).
El-mahgoub SS, 2012 <sup>15</sup>	Confiabilidade (estabilidade); validade de critério (preditiva).	61 adolescentes com sobrepeso ou obesidade e deficiência intelectual (não hospitalizados)	Realizado em um corredor de 20 m. Realizaram o TC6 duas vezes com intervalo de uma semana, no mesmo local e hora do dia. As instruções foram dadas de acordo com as recomendações aceitas. Durante o teste, os participantes foram incentivados verbalmente com frases padrão.	VO <sub>2</sub> pico, IMC, força de quadríceps, Quociente de Inteligência (QI).	- Não houve diferença significativa entre as médias dos dois TC6 (571,3 vs. 576,5 m; p = 0,452). O CCI foi de 0,82, indicando boa confiabilidade. A menor diferença real foi de 82,6 m. - O TC6 apresentou correlação significativa com o VO <sub>2</sub> pico relativo (B = 0,69), indicando validade. O VO <sub>2</sub> pico relativo, a estatura, o QI, o IMC e a força do quadríceps são preditores da distância percorrida no TC6.
Fiss AL, 2019 <sup>16</sup>	Validade de construto (convergente).	708 crianças com paralisia cerebral (não hospitalizados)	Instruções padronizadas para caminhar o máximo possível em 6 min. Os avaliadores forneceram incentivos verbais pré-estabelecidos a cada minuto. As conversas eram limitadas para não impactar ou diminuir o ritmo da caminhada.	<i>Early Activity Scale for Endurance</i> (EASE) e Sistema de classificação da função motora bruta (GMFCS) níveis I a V.	- O EASE e o TC6 apresentaram correlação estatisticamente significativa, mas baixa (r = 0,30; P < 0,001) entre os níveis GMFCS I a III, indicando que EASE e TC6 parecem medir diferentes construtos. - Baixas correlações não significativas entre o EASE e o TC6 de r = 0,14 para GMFCS nível I, r = 0,07 para nível II e r = 0,02 para nível III.
Fulk GD, 2008 <sup>17</sup>	Confiabilidade (estabilidade); validade de critério (concorrente) e de construto.	37 indivíduos após Acidente Vascular Cerebral (AVC) (hospitalizados).	Os indivíduos foram orientados a caminhar o mais longe possível no seu ritmo confortável por 6 min em um percurso oval. A distância tinha 150 pés em um hospital de reabilitação e 250 pés nos outros dois.	Velocidade da marcha (Caminhada de 5 m) e a Medida de Independência Funcional (MIF).	- O TC6 apresentou alta confiabilidade teste-reteste; CCI (2,1) 0,973 (IC 95% = 0,925 - 0,988) e uma MDC (90) de 54,1 m. - Correlação forte a moderada com a velocidade da marcha (r = 0,89), MIF de locomoção (caminhada) (r = 0,69) e MIF motora (r = 0,52). - Medida clinicamente útil e confiável da capacidade de caminhar.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico= pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 2** - Características dos estudos sobre Teste de Caminhada de Seis Minutos (TC6) (n = 42) (cont.).

Autor, ano	CrITÉrios de ValidaÇão	Amostra	Teste	Variáveis associadas	Resultados
Guerra BM, 2015 <sup>18</sup>	Confiabilidade (estabilidade).	46 indivíduos com diferentes níveis de dificuldade intelectual (não hospitalizados).	Realizados em ritmo próprio em 30 m, de acordo com a ATS (2002). Eles foram instruídos a não correr e andar o mais rápido que pudessem, e podiam parar em qualquer momento, mas foram incentivados a voltar o mais rápido possível. Foram incentivados com frases padronizadas.	VO <sub>2</sub> pico, sexo, altura, massa gorda e massa livre de gordura, nível de dificuldade intelectual e força isométrica das pernas.	- As distâncias a pé para T1, T2 e T3 foram de 460,3 ± 76,9 m; 489,4 ± 81,2 m e 491,4 ± 77,9 m, respectivamente. - Os TC6 entre T1-T2 e T1-T3 foram diferentes (p < 0,001), mas entre T2 e T3 não. O CCI entre T2 e T3 foi 0,96, alta confiabilidade. - VO <sub>2</sub> pico e a força isométrica da perna contribuíram significativamente para o TC6 (R <sup>2</sup> = 0,55). É fácil, barato, confiável e válido.
Gomes E, 2016 <sup>19</sup>	Confiabilidade (estabilidade); validade de critério.	51 indivíduos com esquizofrenia (não hospitalizados).	Em um corredor com dois cones separados a 25 m indicavam o caminho. Instruídos a ir e voltar durante 6 min, sem correr. Interrompido se sintomas ameaçadores aparecessem. Fornecidos incentivos recomendados pela ATS.	Percentual de gordura, tabagismo, IMC, atividade física, VO <sub>2</sub> pico.	- Não foram encontradas diferenças significativas entre os dois TC6s. - O CCI foi de 0,94, indicando boa confiabilidade. - Apresentou correlação significativa com VO <sub>2</sub> pico (r = 0,67). - Altura, massa gorda corporal, comportamento de fumar e minutos de atividade física por semana foram significativamente associados ao TC6.
Gulmans VAM, 1996 <sup>20</sup>	Confiabilidade (estabilidade); validade de critério.	23 crianças com sintomas leves de Fibrose Cística e 15 moderados (não hospitalizados).	Os indivíduos foram instruídos a percorrerem a maior distância possível em 6 min no percurso de 8 m, percorrendo repetidamente as distâncias. Foram utilizadas frases padronizadas de incentivo a cada 16 m (“Continue!”, “Você está indo bem!” “Tudo está indo bem!”).	Consumo máximo de oxigênio (VO <sub>2</sub> máx), Volume Residual (VR), Capacidade pulmonar total, Carga de Trabalho Máxima (Wmax).	- Correlação entre a distância (DP = 697 ± 104 m) e a Wmáx (118 ± 44 W; r = 0,76, p < 0,001) e VO <sub>2</sub> máx (1.688 ± 495 mL; r = 0,76, p < 0,001). - VR/CPT%: correlação negativa (r = -0,72; P < 0,01) com a distância. - A análise de regressão múltipla mostrou um coeficiente de regressão múltipla de R <sup>2</sup> = 0,84 (P < 0,001) para Wmax e VR/CPT% vs. distância. - Correlacionaram-se negativamente com a hiperinflação pulmonar.
Kobayashi E, 2016 <sup>21</sup>	Validade de construto (convergente).	24 indivíduos com doença de Parkinson moderada (não hospitalizados).	O TC6 foi realizado em um corredor de 20m. Os indivíduos foram instruídos a caminhar ao longo dos 20 m tanto quanto possível dentro de 6 min.	Estágios Hoehn e Yahr (HY), <i>Unified Parkinson Disease Rating Scales</i> (UPDRS), escala <i>Berg Balance</i> (BBS), <i>Timed Up and Go</i> (TUG), custo energético (EC).	- TUG e EC: boas correlações com TC6; idade e UPDRS: moderadas. - O estágio HY, BBS e cadência mostraram fracas correlações com TC6. - O modelo de regressão representou uma quantidade significativa da variação nos resultados do TC6 (R <sup>2</sup> = 0,77; p < 0,001). - Os resultados TUG e EC foram correlacionados independentemente com os resultados do TC6, considerando que idade, estágio HY e cadência não apresentaram correlações significativas.
Larsson UE, 2008 <sup>22</sup>	Confiabilidade (estabilidade); validade de conteúdo (teste de hipóteses).	43 indivíduos obesos (não hospitalizados).	De acordo com a ATS em corredor de 40 m. Dois cones indicavam o percurso. Instruções: “Agora você deve andar o mais longe possível por 6 min, mas não corra”. Incentivo dado a cada minuto, dizendo quantos minutos faltavam.	Índice de Massa Corporal (IMC).	- O grupo obeso andou 534 m (IC 508 - 560 m) o primeiro e 552 m (IC 523-580) o segundo passeio (p < 0,001). O SEM foi de 25 m, coeficiente de variação 4,7%, CCI 0,96. O limite de concordância foi de -46 m + 80 m. - O grupo obeso andou 162 m menos (p < 0,001) e executou um trabalho muito mais pesado (p < 0,001) do que o grupo magro.
Lee MC, 2018 <sup>23</sup>	Confiabilidade (estabilidade); e validade de critério.	54 mulheres com diabetes mellitus tipo 2 (não hospitalizados).	Realizado de acordo com as diretrizes da ATS. Os sujeitos foram instruídos a caminhar por 6 min, ao longo de uma linha de 30 m em um corredor externo. Foram encorajados a andar o mais rápido possível.	Consumo máximo de oxigênio (VO <sub>2</sub> máx), Teste de Degrau (ST), nível de glicemia e nível de hemoglobina glicada.	- Foi encontrada correlação significativa entre as distâncias percorridas durante o teste de aptidão cardiorrespiratória (VO <sub>2</sub> máx), no TC6 (r = 0,542; p < 0,01) e no teste de degrau (r = -0,490; p < 0,01). - O TC6 e o teste de degrau são confiáveis para a avaliação da aptidão cardiorrespiratória em indivíduos com diabetes mellitus tipo 2.
Li AM, 2005 <sup>24</sup>	Validade de critério (concorrente); confiabilidade (estabilidade).	52 jovens saudáveis do sexo feminino (não hospitalizados).	Instruídos a percorrer um corredor o máximo possível em 6 min. Foi permitido descansar. Incentivos: “continue”, “você está indo bem”, “tudo está indo bem”, e dadas pela mesma pessoa em horários determinados.	Altura, frequência cardíaca, consumo máximo de oxigênio (VO <sub>2</sub> máx).	- A validade concorrente foi demonstrada pela boa correlação entre TC6 e VO <sub>2</sub> máx na esteira ergométrica. A confiabilidade teste-reteste foi realizada em 52 indivíduos e o CCI (IC de 95%) foi calculado como 0,94 (0,89 -0,96). - Análises de Bland e Altman demonstraram alto grau de repetibilidade. - TC6 é confiável e válido: CCI (IC de 95%) de 0,94 (0,89 - 0,96).

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico= pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 2** - Características dos estudos sobre Teste de Caminhada de Seis Minutos (TC6) (n = 42) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Liu WY, 2016 <sup>25</sup>	Validade de construto e confiabilidade (estabilidade).	61 indivíduos com DPOC (não hospitalizados).	GRAIL (Laboratório Interativo de Análise em Tempo Real da <i>Gait</i> ) na esteira. Familiarização (15 - 20 min). No solo, um corredor circular de 125 m. Realizaram duas vezes, descanso de 60 min.	TC6 no GRAIL.	- Caminharam mais no segundo teste no solo (24,8 m; IC 95% 15,2 - 34,4 m, $p < 0,001$ ) e no segundo teste GRAIL (26,8 m; IC 95% 13,9 - 39,6 m). - O GRAIL TC6 foi reprodutível (CCI = 0,65-0,80). A melhor distância no GRAIL nos pacientes foi menor do que no solo (-27,3 ± 49,1 m, $p < 0,001$ ).
Mänttari A, 2018 <sup>26</sup>	Validade de critério.	75 adultos (não hospitalizados).	Em corredores de 15 e 30 m em ordem aleatória. Cones de plástico nas duas extremidades para marcar os pontos da pista. Instrução: "Ande de um lado para outro pelos cones o mais rápido possível por 6 min. Você não pode correr, e você deve girar rapidamente em torno dos cones, como eu mostrei. Você está pronto? Comece!"	Consumo máximo de oxigênio (VO <sub>2</sub> máx), IMC, frequência cardíaca (FC), Teste de Exercício Graduado (GXT).	- A distância média foi 652 m (DP ± 74). O VO <sub>2</sub> máx médio no GXT e O <sub>2</sub> no final do TC6 foi 34,4 ml/kg.min (DP ± 7,6) e 27,2 ml/kg.min (DP ± 6,5). - Para homens, os melhores preditores para o VO <sub>2</sub> máx foram distância, idade, IMC, FC e altura, e para as mulheres, distância, idade e peso. - Os preditores explicaram 82% e 79% do VO <sub>2</sub> máx medido por homens e mulheres com SEM de 3,6 ml/kg.min e 3,5 ml/kg.min, respectivamente. - É válido para prever o VO <sub>2</sub> máx com precisão de cerca de 1 MET.
McDonald CM, 2013 <sup>27</sup>	Confiabilidade (estabilidade); e validade de critério (concorrente).	174 rapazes com distrofia muscular de Duchenne (não hospitalizados).	Realizado seguindo procedimentos padronizados, desenvolvidos na Universidade da Califórnia em Davis, medindo o TC6 em metros. O TC6 para este estudo incluiu modificações no método recomendado pelo ATS para uso em adultos.	Testes de função cronometrada (TFTs), força por miometria, qualidade de vida, índice de gasto energético.	- A confiabilidade teste-reteste para a maioria das medidas foi alta. - O TC6 teve a maior confiabilidade teste-reteste de qualquer desfecho clínico (CCI = 0,92). - Os CCI para TC6, TFTs e miometria portátil eram todos fortes (0,72 -0,92).
Moriello C, 2008 <sup>28</sup>	Validade de conteúdo (teste de hipóteses); de construto (convergente).	63 indivíduos após ressecção eletiva do cólon (hospitalizado)	Instruídos a caminhar por 6 min em um trecho de 20 m de corredor, de acordo com as diretrizes. Eles puderam descansar durante o teste, se necessário, e lhes foi fornecido um período de descanso antes dos testes.	Tempo de internamento (LOS), idade, <i>American Society of Anesthesiologists</i> (ASA)	- Coeficientes de correlação de Pearson ou Spearman entre o TC6 no pré-operatório, 3 e 6 semanas após a alta. A idade e o ASA correlacionaram-se negativamente com TC6. Albumina sérica no pré-operatório correlacionou-se com TC6 no início ( $r = 0,36$ ) e 6 semanas ( $r = 0,35$ ). - Correlação negativa após 3 semanas ( $r = -0,39$ ; $p < 0,05$ ) com LOS.
Morinder G, 2009 <sup>29</sup>	Confiabilidade (estabilidade); e validade de conteúdo (teste de hipóteses).	250 crianças obesas e 97 crianças com peso normal (não hospitalizados).	Realizado em um corredor de 70 m. Foram instruídos a usar sapatos confortáveis e a andar o máximo possível, sem correr (ATS, 2002), mas o mais rápido possível. As informações foram fornecidas durante o teste, dizendo quantos minutos eles caminharam ou minutos restantes.	Consumo máximo de oxigênio (VO <sub>2</sub> máx).	- Crianças obesas: 571 m no primeiro teste, 57 m no segundo ( $p = 0,578$ ). - O SEM foi de 24 m, coeficiente de variação (CV): 4,3% e CCI 0,84. - Repetibilidade de 68 m, e limites de concordância de +71 e -65 m. - A distância percorrida nas crianças obesas foi de 571 m (65,5) e nas crianças com peso normal, 663 m (61,1) ( $p < 0,001$ ). - A correlação entre TC6 e VO <sub>2</sub> máx estimado ( $r = 0,34$ ) foi baixa.
Mossberg KA, 2012 <sup>30</sup>	Validade de critério (concorrente).	21 indivíduos em reabilitação após lesão cerebral traumática (hospitalizado).	Foram instruídos a andar tantas vezes quanto possível de forma rápida, por 6 min, em uma pista retangular de 12 m x 6 m. O encorajamento verbal foi padronizado e incluiu "você está indo bem" após cada minuto. Nos minutos 4 e 5, foi dito que tinham 2 e 1 minuto.	Frequência cardíaca, índice de custo fisiológico (batimentos por metro - PCI), VO <sub>2</sub> pico.	- Entre a admissão e a alta, a distância média aumentou de 342,6 m (DP = 127,0) para 408,9 m (DP = 124,2), e o trabalho aumentou de 27,185 kg·m (DP = 10,528) para 34,114 kg·m (DP = 12,057). - Índices de tamanho de efeito: 1,10 e 1,12 para distância e trabalho. - As correlações ( $r$ ) entre o pico de VO <sub>2</sub> e a distância do TC6, PCI e trabalho foram de 0,58, -0,61 e 0,47, respectivamente.
Nasuti G, 2013 <sup>31</sup>	Confiabilidade (estabilidade); e validade de critério (concorrente).	13 adultos com deficiência intelectual (não hospitalizados).	De acordo com a ATS em corredor de 30m. Com frases de incentivo a cada 15s. Uma prática foi fornecida antes; estavam familiarizados com o protocolo.	VO <sub>2</sub> pico.	- A confiabilidade teste-reteste do TC6 modificado foi CCI = 0,98, e a regressão linear revelou que a distância do TC6 modificado estava significativamente relacionada ao pico de VO <sub>2</sub> ( $R = 0,84$ , $p < 0,001$ ). - Validade aceitável e excelente confiabilidade.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico= pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 2** - Características dos estudos sobre Teste de Caminhada de Seis Minutos (TC6) (n = 42) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Nathan SD, 2015 <sup>32</sup>	Validade de critério.	338 indivíduos com fibrose pulmonar idiopática (não hospitalizados).	Caminhar o mais longe, sem correr. Autorizados a desacelerar ou parar e incentivados a retomar assim que pudessem. Interrompido se apresentasse dor no peito, dispnéia, câibras, diaforese ou dessaturação <83%.	<i>San Diego Shortness of Breath Questionnaire</i> (SOBQ), função pulmonar e qualidade de vida.	- Correlações mais fortes com função pulmonar e qualidade de vida. - A dispnéia, mostrou as correlações mais fortes com o TC6 (linha de base: coeficiente -0,35; em 48 semanas: coeficiente -0,37; p < 0,001). - MCID com referência de critério (hospitalização ou morte) foi 21,7 m.
Nsenga LU, 2012 <sup>33</sup>	Confiabilidade (estabilidade); validade de critério.	24 crianças/adolescentes com paralisia cerebral (não hospitalizados).	Realizado em 20 m. Instruídos a caminhar o mais longe em 6 min; muletas no antebraço não foram usadas para evitar aumentar o custo energético. Com incentivos, a cada 30 s eram avisados da distância e do tempo restantes.	VO <sub>2</sub> pico, pico de ventilação, pico de frequência cardíaca.	- As correlações teste-reteste para o TC6 foram boas (VO <sub>2</sub> pico: r = 0,90, p < 0,001, CCI = 0,55; pico de ventilação: r = 0,88, p < 0,001, CCI = 0,83; pico de frequência cardíaca: r = 0,86, p < 0,001, CCI = 0,82; distância percorrida: r = 0,87, p = 0,007, CCI = 0,80). - As pontuações corresponderam aos dados do ciclo-ergômetro.
Pankoff B, 2000 <sup>34</sup>	Validade de critério.	28 indivíduos com fibromialgia (não hospitalizados).	Realizado em um corredor de 70 m. Instruídos a caminhar o mais rápido, mas que pudesse ser mantido por 6 min. Nenhum incentivo foi dado, mas o <i>feedback</i> do tempo no último minuto. Usavam cronômetro e pararam quando o tocou.	Questionário de Impacto da Fibromialgia (FIQ) e VO <sub>2</sub> pico.	- As correlações entre a distância de TC6 e o VO <sub>2</sub> pico antes (r = 0,328) e depois (r = 0,420) não foram significativas. Correlações entre TC6 e FIQ (r = -0,494, p < 0,01) e comprometimento físico (r = -0,403, p < 0,05). - A mudança na distância foi correlacionada (p < 0,05) com pontuação de FIQ, mas nenhuma alteração no VO <sub>2</sub> pico. O TC6 não foi preditor válido.
Pecorelli N, 2016 <sup>35</sup>	Validade de construto.	151 indivíduos após ressecção colorretal (não hospitalizados).	Realizado em um trecho plano de 15 m e foram instruídos a caminhar por 6 min no ritmo. Foram utilizados incentivos dados minuto a minuto, de acordo com as diretrizes da ATS.	Estado físico, cirurgia aberta e ocorrência de complicações pós-operatórias.	- Todas as hipóteses para validade de construto foram suportadas. - Idade avançada, pior estado físico, cirurgia aberta e complicações foram associadas à redução clinicamente relevante (> 19 m). - Correlação entre distância percorrida e atividade física (r = 0,46).
Phillips D, 2019 <sup>36</sup>	Confiabilidade (estabilidade).	Crianças (n = 11), adolescentes (n = 4) e adultos (n = 9) com hipofosfatase (não hospitalizados).	O TC6 foi administrado por um fisioterapeuta de acordo com as diretrizes da ATS.	Escala <i>Radiographic Global Impression of Change</i> e de gravidade de Rickets, incapacidade e dor, atividades da vida diária (AVDs), função dos membros inferiores (MMII)	- Confiabilidade para crianças (r = 0,95; p < 0,0001), adolescentes (r = 0,81; p = 0,125) e adultos (r = 0,94; p = 0,0001). As MCID foram 31 m (crianças e adultos) e 43 m (adolescentes). Correlação com escala <i>Radiographic Global Impression of Change</i> (r = 0,50; p < 0,0001), gravidade de Rickets (r = -0,78; p < 0,0001), distância (p < 0,0001), incapacidade (r = -0,67), AVDs (r = 0,71 a 0,77) e dor (r = -0,39). Em adolescentes e adultos, a distância correlacionou-se (p < 0,05) com função de MMII (r = 0,83 e 0,60), gravidade da dor (r = -0,41 e -0,36) e interferência da dor (r = -0,41 e -0,49).
Ries JD, 2009 <sup>37</sup>	Confiabilidade (estabilidade).	51 pessoas com doença de Alzheimer (DA) (não hospitalizados).	Realizado em longos corredores. Caminharam em um ritmo acelerado, foram instruídos a não conversar e foram informados a cada minuto sobre o tempo. Se estivessem distraídos ou parados, eram solicitados "Continue andando".	<i>Mini-Mental Status Examination, Timed Up and Go</i> (TUG), Teste de Velocidade da Marcha (TVM).	- A confiabilidade teste-reteste para o TUG, o TC6 e a TVM (CCI ≥ 0,973). A variabilidade individual do desempenho também foi alta. - Os escores MDC calculados no IC 90% foram: TUG = 4,09 s, TC6 = 33,5 m (110 pés) e velocidade da marcha = 9,4 cm/s. - Os grupos foram significativamente diferentes na realização dos testes.
Sandberg A, 2020 <sup>38</sup>	Confiabilidade (estabilidade).	102 indivíduos com claudicação intermitente (não hospitalizados).	Realizado em corredor de 30 m, marcados com cones. O indivíduo sentou-se em uma cadeira, por 10 min antes de iniciar. Frases padronizadas de encorajamento foram utilizadas todo minuto.	Índice tornozelo-braquial (ITB), Classificação de Esforço Percebido pela Escala de Borg.	- A distância média no teste-reteste foi de 397,8 m (DP 81,2 m; N = 100), e a distância individual de caminhada variou de 175 a 600 m. - Excelente confiabilidade teste-reteste para o TC6 (CCI 0,95; IC 95% 0,9 - 0,97). O SEM foi de 16,6 m (IC 95% 14,6 - 19,3), a MDC foi de 46 m.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico= pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 2** - Características dos estudos sobre Teste de Caminhada de Seis Minutos (TC6) (n = 42) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Sanjak M, 2017 <sup>39</sup>	Validade de construto.	186 indivíduos com esclerose lateral amiotrófica (ELA) (não hospitalizados).	De acordo com a ATS. Caminharam sobre uma superfície plana, usando sapatos e seus dispositivos em torno de 25 m de pista. Foram instruídos a caminhar o mais longe possível por 6 min; e a parar, descansar e encostar-se na parede, se necessário, sem sentar, e retomar a caminhada o mais rápido.	25-foot walk test (25FWT), <i>Timed Up and Go</i> (TUG), contração dos membros inferiores (MVIC), escala funcional de avaliação da ELA (ALSFRS-R) e CVF.	- O TC6 correlacionou-se com todas as outras medidas de desfecho em pacientes ambulatoriais com ELA (25FWT: $r = -0,74$ , $p < 0,0001$ ; teste TUG: $r = -0,80$ , $p < 0,0001$ ; MVIC: $r = 0,64$ , $p < 0,0001$ ; CVF percentual prevista: $r = 0,25$ , $p = 0,0007$ ; ALSFRS-R: $r = 0,52$ , $p < 0,0001$ ; sub-pontuação bruta do motor ALSFRS-R: $r = 0,71$ , $p < 0,0001$ ). - O TC6 é uma medida válida da capacidade de deambulação de pacientes ambulatoriais com ELA.
Schmidt K, 2013 <sup>40</sup>	Confiabilidade (estabilidade).	50 indivíduos durante ou fora do tratamento do câncer (hospitalizado).	Realizado de acordo com a ATS. O curso de teste foi 30 m de comprimento em um corredor interno. Os sujeitos foram instruídos a andar no seu ritmo mais rápido, cobrindo a maior distância possível. Incentivo padronizado foi dado e o tempo restante foi informado a cada minuto.	VO <sub>2</sub> pico, <i>European Organization for Research and Treatment of Cancer Quality of Life Questionnaire</i> (EORTC QLQ-C30), esforço percebido (EPR).	- A distância percorrida se correlacionou ( $p < 0,001$ ) com o VO <sub>2</sub> pico ( $r = 0,67$ ) e a função física percebida (EORTC QLQ-C30) ( $r = 0,55$ ). - Confiabilidade: CCI $r = 0,93$ (IC 95% +0,86; +0,97; $p < 0,001$ ) e o coeficiente de variação de 3%. - No reteste, os indivíduos caminharam 3,1% (IC95% +1,1; +5,2) mais longe e alcançaram um EPR mais alto (+1,0; IC95%: +0,3; +1,8). - Os limites de concordância estavam entre - 43,1 e 76,4 m.
Scivoletto L, 2011 <sup>41</sup>	Validade de critério e a confiabilidade (equivalência).	37 indivíduos com lesão medular (não hospitalizados).	Duas condições: 1) pista retangular 10 x 50 m; 2) pista curta de 10 m. Instruídos a caminhar a uma velocidade auto-selecionada sem descansar. Com uso de dispositivos (aparelhos ortodônticos/aparelhos para caminhar) e assistência física.	Teste de Velocidade de 10 m (TVM10m)	- Diferenças pelo comprimento da via: mediana de 226,7 m (25°- 75° intervalo interquartil 123,2 - 319 m) na mais longa e 187,6 m (25°- 75° intervalo interquartil 69,7 - 240,6) na curta ( $P < 0,001$ ). Correlação de 0,91 a 0,93. - A confiabilidade inter e intraexaminadores foi de 0,98 e 0,99. - As diferentes condições de teste resultam em diferenças significativas.
Sinclair RC, 2012 <sup>42</sup>	Validade de critério.	110 indivíduos antes de uma grande cirurgia não cardíaca (não hospitalizados).	Conforme descrito nas orientações publicadas pela ATS. Indivíduos caminharam no seu próprio ritmo máximo ao longo de um corredor plano, marcado com uma pista de 30 m, com o objetivo de percorrer a maior distância possível nos 6 min cronometrados.	Teste de Esforço Cardiopulmonar (TECP).	- A análise da curva ROC ( <i>Receiver Operating Characteristic</i> ) revelou uma área abaixo da curva de 0,85 (IC 95% 0,77 - 0,91). - Os pontos de corte ótimos foram < 440 m (método convencional) e < 502 m (abordagem ponderada pela sensibilidade). Os pontos de corte da distância inferior e superior do TC6 com base na regressão foram < 427 e > 563 m, respectivamente.
Stellmann JP, 2015 <sup>43</sup>	Validade de critério.	28 indivíduos com esclerose múltipla (não hospitalizados).	Foram instruídos a caminhar na velocidade mais rápida de acordo com as diretrizes. Realizaram o teste em um corredor de 20 m com giro em cada extremidade. Após cada minuto, eram informados sobre o tempo restante, mas não incentivado.	Velocidade de marcha.	- A maioria dos parâmetros da velocidade de marcha foram associados ao TC6, mas distância média e o número de passos não. - A correlação com diferentes velocidades foi consistentemente moderada ( $R^2: 0,47 - 0,52$ ). As melhores correlações foram encontradas para sequências com 50 ou 100 passos ( $R^2 = 0,68$ e $0,71$ , respectivamente).
Takken T, 2009 <sup>44</sup>	Validade de construto (discriminante).	30 crianças com doença renal terminal (DRT) (hospitalizado).	Realizado em uma pista de 20 m no corredor do hospital. Foram instruídos a percorrer a maior distância possível em 6 min em uma caminhada escolhida. Foram feitas curvas nas duas extremidades da pista de 20 m. A cada intervalo de 1 min, receberam incentivo.	Altura, peso, IMC, hematócrito, hemoglobina, duração da doença, força muscular, Volume corrente e VO <sub>2</sub> pico.	- Os preditores mais fortes do desempenho: hematócrito e altura. - Os modelos de regressão explicaram 59% (hematócrito e altura) a 60% (hematócrito) da variância no desempenho no TC6. - O desempenho no TC6 não foi associado ao VO <sub>2</sub> pico, força ou outras variáveis antropométricas, mas ao hematócrito e estatura. - Crianças com doença renal terminal tiveram pontuação mais baixa.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico= pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 2** - Características dos estudos sobre Teste de Caminhada de Seis Minutos (TC6) (n = 42) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Uszko-Lencer NHMK, 2017 <sup>45</sup>	Validade de construto (convergente e discriminante) e de conteúdo (teste de hipóteses).	337 indivíduos com Insuficiência Cardíaca Crônica (ICC) (não hospitalizados).	De acordo com diretrizes internacionais. Foram instruídos caminhar no seu próprio ritmo enquanto tentavam cobrir a maior distância possível. O incentivo padronizado era fornecido a cada minuto, dizendo: "Vocês estão indo bem" ou "Continue com o bom trabalho", informando-lhes o tempo restante.	Dados clínicos, fração de ejeção, capacidade máxima de exercício, composição corporal, função pulmonar e sintomas de ansiedade e depressão, VO <sub>2</sub> pico.	- Confiável (CCI = 0,90; P < 0,0001). Efeito aprendizagem de 31 m (IC 95% 27 - 35m). Idade (≥ 65), menor DLCO e maior NYHA foram associadas a menor aumento no segundo teste (OR 0,45 - 0,56; P < 0,05). - Melhores TC6: correlações moderadas a boas com VO <sub>2</sub> pico (rs = 0,54 -0,69) e composição corporal, função pulmonar, fração de ejeção, ansiedade e depressão (rs = 0,04 - 0,49). Associado à potência durante o exercício, à taxa de filtração glomerular e à idade (51,7% da variabilidade).
Vanhelst J, 2013 <sup>46</sup>	Confiabilidade (estabilidade); validade de critério.	52 meninas e 45 meninos, jovens obesos (não hospitalizados).	Instruídos a caminhar o mais longe em 6 min em um campo (40 × 20m). Fatores ambientais: sem vento, nível de solo, superfície e temperatura 18°C. Roupas e sapatos adequados e foram incentivados de acordo com a ATS.	IMC, consumo máximo de oxigênio (VO <sub>2</sub> máx).	- Andaram 689,4 ± 98,6 m no TC6 inicial e 690,4 ± 98,2 m no segundo. - O CCI foi r = 0,99. O método de Bland e Altman também mostrou boa concordância entre os dois TC6. - A regressão múltipla determinou que a distância percorrida e IMC são os fatores mais significativos na previsão do VO <sub>2</sub> máx.
Vis JC, 2009 <sup>47</sup>	Confiabilidade (estabilidade).	54 indivíduos com Síndrome de Down com grave doença cardíaca (não hospitalizados)	Realizado em terreno de 30 ou 40m em quatro locais diferente, de acordo com a ATS. Foram instruídos a caminhar o mais longe possível por 6 min, sem correr. O descanso era permitido, se necessário, mas a caminhada era retomada assim que pudessem fazê-lo.	Idade avançada, sexo e o nível grave de deficiência intelectual.	- O TC6 médio no grupo com doença cardíaca grave foi de 289 ± 104m e no grupo sem de 280 ± 104m (P = 0,70). - Idade avançada, sexo feminino e nível grave de deficiência intelectual estavam correlacionados com TC6 inferior (r = 0,67; P < 0,001). - O TC6 não foi significativamente diferente (310 ± 88m vs. 317 ± 85 m; P = 0,40) quando realizado duas vezes. Coeficiente de variação de 11%.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico= pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 3** - Características dos estudos sobre Teste de Velocidade da Marcha (TVM) (n = 13).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Arjunan A, 2018 <sup>48</sup>	Validade de critério (preditiva).	258 adultos com 65 anos ou mais (hospitalizado).	A velocidade lenta da marcha (GS) foi avaliada pelos fisioterapeutas da enfermaria na admissão usando o TVM de 10 m em ritmo confortável, permitindo o uso de auxiliares de marcha. Zonas de aceleração e desaceleração estavam presentes. Aqueles incapazes de completar o teste GS na admissão receberam um GS de 0.	Índice de fragilidade (FI), velocidade lenta da marcha (GS), queda, delirium, tempo de internamento (LOS).	- FI e GS tiveram uma correlação negativa fraca a moderada, com Coeficiente de Pearson de -0,396 e P < 0,001. À medida que o FI aumentou, o GS diminuiu. FI e GS mostraram associações com tempo de internação, alta exigindo maior nível de cuidados e delirium. - Cada aumento de 0,1 na FI aumentou a probabilidade de maior duração de permanência, pior resultado da alta e delirium. <i>Odds Ratios</i> (OR) de 2,1 (IC95% 1,6 - 2,8), 1,4 (IC95% 1,2 - 1,8) e 1,4 (IC95% 1,1 - 1,8).
Bloem AEM, 2018 <sup>49</sup>	Validade de critério.	51 indivíduos com fibrose pulmonar idio-pática (não hospitalizado).	Instruídos a "andar o mais rápido possível" por 4 m; espaço para acelerar e desacelerar de 2 m, antes/depois da distância definida. As distâncias (0-2-6-8 m) foram marcadas com fita no chão. Foi realizado três vezes (intervalo de 2 min).	Teste de Sentar e Levantar 5 vezes (TSO5x) e Teste de Caminhada de 6 min (TC6).	- TC6: 453 ± 111m. TVM4m: 2,0 ± 0,5 s. TSO5x: 12,0 ± 3,8 s. - O TVM4m (r = 0,77; p < 0,01) e o TSO5x (r = -0,41; p < 0,01) correlacionaram-se significativamente com a distância percorrida no TC6. - O TVM4m combinado com força de prensão manual e dispneia do pode explicar 75% da variância no TC6.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min= minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 3** - Características dos estudos sobre Teste de Velocidade da Marcha (TVM) (n = 13) (cont.).

Autor, ano	CrITÉRIOS de ValidaÇÃO	Amostra	Teste	Variáveis associadas	Resultados
Boonstra AM, 1993 <sup>50</sup>	Validade de critério e confiabilidade (estabilidade).	15 indivíduos normais, 16 amputados transfemorais e 8 desarticulações do joelho (não hospitalizado).	O teste foi feito em um corredor de 10 m. Dois raios infravermelhos no início e o final de uma trajetória de 7 m começaram e pararam as medições automaticamente. Os indivíduos andaram pela primeira vez sem nenhum eletrodo. A velocidade de andar confortável e rápido a caminhada (“o mais rápido possível”) foi medida. Foi repetido uma vez.	Registro simultâneo de eletromiografia (EMG) e goniometria em velocidades de caminhada confortáveis e rápidas.	<ul style="list-style-type: none"> <li>- A variação de velocidade nas sessões durante a caminhada em alta velocidade foi maior ao caminhar sem equipamento do que ao caminhar com equipamento.</li> <li>- As variações de velocidade nas sessões dos sujeitos normais foram superiores às dos dois grupos de amputados.</li> <li>- A confiabilidade teste-reteste, expressa pelo CCI, foi boa: entre 0,83 e 0,98.</li> </ul>
Chan KS, 2015 <sup>9</sup>	Confiabilidade (estabilidade e equivalência); validade de construto (convergente, discriminante) e critério (preditiva)	306 indivíduos adultos após Insuficiência Respiratória Aguda (hospitalizado).	A velocidade da marcha de 4 m, em m/s é a principal medida para este estudo psicométrico. O teste foi realizado duas vezes em cada avaliação. Conforme recomendado, a velocidade mais rápida foi utilizada.	Força muscular e função pulmonar, medidas da função física, hospitalizações futuras e qualidade de vida relacionada à saúde.	<ul style="list-style-type: none"> <li>- Excelente confiabilidade teste-reteste (CCI 0,89 - 0,99 nos estudos e acompanhamento) e confiabilidade entre avaliadores (CCI 0,97).</li> <li>- A validade convergente foi apoiada por correlações moderadas a fortes (69% de 32 &gt; 0,40) com medidas da função física, a validade discriminante por correlações fracas (86% de 28 &lt; 0,30) com saúde mental.</li> <li>- Sobreviventes com medidas prejudicadas vs. não prejudicadas de força muscular e função pulmonar tiveram velocidade mais lenta (p &lt; 0,05).</li> <li>- Previu hospitalizações e qualidade de vida. MCID de 0,03 - 0,06 m/s.</li> </ul>
Combs SA, 2014 <sup>51</sup>	Confiabilidade (estabilidade) e validade de construto (convergente e discriminante).	88 indivíduos com doença de Parkinson (DP) leve a moderada (não hospitalizados).	Os testes de velocidade da marcha de 10 m confortáveis (CWT) e rápidos (FWT) foram realizados em um corredor aberto de 14 m, foram cronometrados nos 10 m do meio. Instruídos a caminhar em um ritmo auto-selecionado, o mais rápido e com segurança possível. Concluíram três tentativas, com breves intervalos.	Mobilidade, qualidade de vida, Teste de Caminhada de 6 min (TC6).	<ul style="list-style-type: none"> <li>- A confiabilidade teste-reteste foi alta para as medidas de velocidade de caminhada (CWT, ICC (2,1) = 0,98; FWT, ICC (2,1) = 0,99).</li> <li>- A MDC para o CWT e o FWT foi de 0,09 m/s e 0,13 m/s.</li> <li>- Os níveis 3/4 de Hoehn &amp; Yahr correlacionaram-se com velocidade de caminhada mais lenta com CWT e FWT do que os níveis 1 e 2 (P &lt; 0,01).</li> <li>- O CWT e o FWT foram significativamente correlacionados (P &lt; 0,002) com todas as atividades e medidas baseadas na participação.</li> </ul>
Dolmage TE, 2017 <sup>52</sup>	Validade de construto.	99 indivíduos com DPOC (não hospitalizados).	Caminharam usando auxílio habitual por 18 m na velocidade usual, viraram em um cone e retornaram a uma velocidade rápida. O temporizador foi acionado por sensores colocados a 4 e 14 m. A velocidade foi calculada nos primeiros 4 m, e nos 10 m subsequentes.	IMC, Volume Expiratório Forçado no primeiro segundo (VEF <sub>1</sub> ), CVF, VEF <sub>1</sub> /CVF.	<ul style="list-style-type: none"> <li>- A velocidade da marcha subestimou a velocidade usual da marcha (diferença = 6,1 m/min [5,3 - 6,9 m/min]) com concordância ruim (8 m/min [6,6 - 9,4 m/min]) entre os dois métodos de relatório de velocidade.</li> <li>- A velocidade inicial em pé identificou incorretamente alguns indivíduos como de maior risco para problemas de saúde. A velocidade da marcha produziu 11 avaliações falso-negativas da capacidade de concluir.</li> </ul>
Kim HJ, 2016 <sup>53</sup>	Confiabilidade (estabilidade); validade de critério (concorrente).	54 idosas (não hospitalizadas).	Velocidade da marcha em 4, 6 e 10 m. Foi repetido duas vezes, usado a média. Instruídos a caminhar em um ritmo normal e confortável ou o mais rápido. Um avaliador treinado caminhou atrás do participante e parou de cronometrar no final. Podiam descansar se necessário.	Composição corporal, autoavaliação da saúde, depressão e função física ( <i>Short physical performance battery</i> -SPPB).	<ul style="list-style-type: none"> <li>- Correlações positivas entre a massa muscular e ritmo de caminhada (r = 0,301 - 0,308; p &lt; 0,05), mas não com a gordura corporal.</li> <li>- Todos os testes de marcha correlacionaram-se positivamente com a autoavaliação da saúde (ritmo normal r = 0,328 - 0,346, p &lt; 0,05; ritmo máximo r = 0,427 - 0,472, p &lt; 0,001) e depressão (ritmo normal r = 0,279 - 0,430, p &lt; 0,05; ritmo máximo r = 0,413 - 0,456, p &lt; 0,001).</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= American Thoracic Society, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 3** - Características dos estudos sobre Teste de Velocidade da Marcha (TVM) (n = 13) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Kon SS, 2013 <sup>54</sup>	Confiabilidade (estabilidade e equivalência).	138 indivíduos com DPOC (não hospitalizados).	Realizado em 4 m (TVM4m), com demonstração. Posicionados com os dedos tocando a linha de partida. Instruções: “Este é o percurso. Quero que caminhe até o outro lado na velocidade normal, como se estivesse andando na rua para ir às lojas. Caminhe sem parar”. O cronômetro iniciou quando começou a se mover e foi parado quando o primeiro pé cruzou a linha. Permitido dispositivo e suporte de oxigênio.	Volume Expiratório Forçado no primeiro segundo (VEF <sub>1</sub> ), <i>Incremental Shuttle Walk Test</i> (ISW), <i>Medical Research Council</i> (MRC) para dispneia e o <i>St George's Respiratory Questionnaire</i> (SGRQ).	<ul style="list-style-type: none"> <li>- Os CCI para a confiabilidade teste-reteste e interobservadores foram de 0,97 e 0,99, respectivamente. Houve correlação positiva significativa entre TVM4m com ISW (<math>p = 0,78</math>; <math>p &lt; 0,001</math>) e correlação negativa com a escala de dispneia MRC e SGRQ (<math>p = -0,55</math> e <math>-0,44</math>; <math>p &lt; 0,001</math> para ambos).</li> <li>- Indivíduos com DPOC com TVM4m lentos apresentaram redução do ISW e maior escala de dispneia por MRC e SGRQ do que aqueles com velocidade de caminhada preservada, apesar do VEF1% previsto.</li> <li>- O TVM4m é confiável na DPOC, correlaciona-se com a capacidade de exercício, dispneia e qualidade de vida relacionada à saúde.</li> </ul>
Nolan CM, 2018 <sup>55</sup>	Confiabilidade (estabilidade e equivalência), Validade de critério (concorrente).	279 indivíduos com fibrose pulmonar idiopática (FPI) (não hospitalizados).	O teste de velocidade da marcha de 4m (TVM4m) foi medido em um percurso plano de 4 m com os participantes caminhando na velocidade usual. A avaliação foi realizada por equipe de pesquisa treinada e assistentes de fisioterapia que seguiram um procedimento operacional padrão. Dois testes foram realizados com descanso mínimo e o mais rápido dos dois, medido pelo cronômetro, foi usado para calcular a velocidade da marcha.	Tomografia computadorizada de alta resolução (TCAR), Teste de Caminhada de 6 min (TC6), dispneia pelo <i>Medical Research Council</i> (MRC), King's Brief Interstitial Lung Disease (KBILD), índice prognóstico de sexo, idade e fisiologia pulmonar (GAP).	<ul style="list-style-type: none"> <li>- Os CCI para confiabilidade inter-observador e teste-reteste foram 0,996 e 0,983, respectivamente.</li> <li>- Houve forte associação entre TVM4m e TC6 (<math>r = 0,76</math>; <math>p &lt; 0,0001</math>) e correlações moderadas com MRC (<math>r = -0,56</math>), KBILD (<math>r = 0,44</math>) e índice GAP (<math>r = -0,41</math>); todos <math>p &lt; 0,05</math>.</li> <li>- Melhorou significativamente com a reabilitação pulmonar (alteração média (IC 95%): 0,16 m/s (0,12 - 0,20 m/s)), tamanho do efeito 0,65.</li> <li>- Indivíduos com TVM4m lentos tiveram desempenho de exercício significativamente pior (TC6: -167 m (-220 a -133 m)), dispneia, estado de saúde e índice de prognóstico do que aqueles com TVM4m preservado, apesar da função pulmonar e parâmetros de TCAR semelhantes.</li> <li>- O TVM4m é simples, confiável, válida e responsivo na FPI estável.</li> </ul>
Peters DM, 2013 <sup>56</sup>	Confiabilidade (estabilidade) e validade de critério (concorrente).	43 idosos saudáveis (não hospitalizados).	Cada participante realizou o teste três vezes em um corredor de 4 m. Foram instruídos a “andar no seu ritmo confortável e habitual”. Temporizadores sem fio foram usados e colocados no início e no final do corredor. Os participantes puderam usar dispositivo de marcha e receberam descanso necessário durante a sessão de testes.	Teste de Velocidade da marcha de 10m (TVM10m).	<ul style="list-style-type: none"> <li>- As medidas para os testes mostraram excelente confiabilidade no teste-reteste (CCI 0,96 - 0,98), com resultados semelhantes para avaliações de cronômetro e cronômetro automático (CCI de 0,99 - 1,00).</li> <li>- O SEM foi pequeno (0,004 - 0,008 m/s). Enquanto o CCI para medições da velocidade entre os dois testes foi de 0,93, a análise de Bland-Altman revelou uma discrepância de <math>\pm 0,15</math> a <math>\pm 0,17</math> m/s entre os métodos.</li> <li>- não apresenta um grau alto de validade concorrente com o de 10 m.</li> </ul>
Reychler G, 2018 <sup>57</sup>	Validade de critério.	61 crianças (não hospitalizados).	Realizado em um corredor de 4m marcado com fita após demonstração. Foram posicionadas com os dedos tocando a primeira marca. Instruções a caminhar na velocidade habitual e o cronômetro foi parado quando o primeiro pé atravessou completamente. Incentivos padronizados foram dados. O teste foi repetido sem descanso.	Frequência cardíaca (FC), Teste do Degrau de 6 min (ST6) e Teste de Caminhada de 6 min (TC6).	<ul style="list-style-type: none"> <li>- A distância percorrida durante o TC6 foi correlacionada com o número de etapas durante o ST6 (<math>r = 0,320</math>; <math>p = 0,013</math>) e não correlacionada ao TVM4m (<math>r = -0,074</math>; <math>p = 0,575</math>).</li> <li>- não foi encontrada correlação entre ST6 e TVM4m (<math>r = -0,129</math>; <math>p = 0,332</math>).</li> <li>- O aumento da FC foi maior no ST6 e significativamente menor no TVM4m do que no TC6. A dispneia e o esforço percebido foram menores após o TC6 do que após o ST6 e maiores que no TVM4m. Maiores durante o ST6.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 3** - Características dos estudos sobre Teste de Velocidade da Marcha (TVM) (n = 13) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Ries JD, 2009 <sup>37</sup>	Confiabilidade (estabilidade).	51 indivíduos com doença de Alzheimer (DA) (não hospitalizados).	Foi avaliada usando a passarela GAITRite, tapete portátil com sensores. Instruídos a caminhar em um ritmo “confortável” pelo comprimento do tapete (4,57 m [15 pés]), e o caminho a pé foi estabelecido de modo que a aceleração e desaceleração não ocorressem no tapete.	<i>Mini-Mental Status Examination, Functional Assessment Staging, Timed Up and Go</i> (TUG), Teste de Caminhada de 6 min (TC6).	- A confiabilidade teste-reteste para o TUG, o TC6 e a velocidade da marcha foram altos para todos os participantes juntos (CCI > ou = 0,973); no entanto, a variabilidade individual do desempenho foi alta. - Os escores MDC no IC de 90% foram: TUG = 4,09 s, TC6 = 33,5 m (110 pés) e velocidade da marcha = 9,4 cm/s. - O TUG, o TC6 e a velocidade da marcha são medidas confiáveis.
van Loo MA, 2003 <sup>58</sup>	Confiabilidade (equivalência) e validade de critério (concorrente).	12 indivíduos com lesão cerebral traumática (hospitalizado)	Foram instruídos a caminhar cinco vezes em um o mais rápido sem correr. Um ensaio de familiarização foi realizado. Percorreram 14 m, apenas os 10 m internos foram cronometrados. Os 2 m antes e após foram usados para avaliar aceleração e desaceleração.	Velocidade da marcha.	- A confiabilidade interexaminadores foi muito alta, com um coeficiente de correlação intraclasse de pelo menos 0,998 para testes confortáveis e em ritmo acelerado. - A validade concorrente foi excelente para testes rápidos e confortáveis, com correlações perfeitas entre os procedimentos de medição do cronômetro e do cronômetro infravermelho.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, IMC= Índice de Massa Corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= American Thoracic Society, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 4** - Características dos estudos sobre o Teste do Degrau (ST) (n = 12)

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Arcuri JF, 2016 <sup>59</sup>	Confiabilidade (estabilidade), validade de critério.	91 indivíduos saudáveis e com capacidade normal de exercício (não hospitalizados).	Um degrau de madeira de 20 cm. Instruídos a subir e descer a uma velocidade que eles pensavam permitir que executassem a quantidade máxima durante 6 min. Poderiam começar com o membro inferior que se sentiam mais à vontade e alterá-lo a qualquer momento.	Circunferência abdominal, porcentagem de gordura corporal e massa livre de gordura, Teste de Caminhada de 6 min (TC6).	- O desempenho (passos médios ± DP) no primeiro e segundo teste foi de 149 ± 34 e 149 ± 36 passos, respectivamente, o que foi correlacionado com a distância (em m) no TC6 (r = 0,72; P < 0,05). - ST6 era confiável (ICC = 0,9; IC 95%: 0,85 - 0,93). A equação para prever valores de referência foi significativa (P < 0,001 e R = 0,48): Desempenho (etapas) = 174 a 1,05 × Idade (anos) para mulheres e Desempenho (etapas) = 209 a 1,05 × Idade (anos) para homens.
Bonnevie T, 2019 <sup>60</sup>	Validade de critério (preditiva) e validade de conteúdo (teste de hipóteses).	35 indivíduos com DPOC (não hospitalizados).	Realizaram o ST6 duas vezes. O segundo teste começou quando os valores da frequência cardíaca (FC) e da saturação de oxigênio (SpO <sub>2</sub> ) retornaram aos valores basais após pelo menos 20 min. O segundo teste foi registrado. As instruções foram baseadas na ATS para o TC6.	Teste de Caminhada de 6 min (TC6), carga máxima de trabalho obtida durante o teste (Wpeak) e uma repetição máxima (1RM).	- Houve uma relação moderada entre o ST6, o Wpeak e o 1RM (r = 0,44 e r = 0,41, respectivamente, ambos P ≤ 0,01). - A prescrição usando o 1RM estimado em comparação com 1RM medido resultou em uma diferença absoluta média e viés de cerca de 30 kg. - O uso do ST6 e do Wpeak na prescrição do treinamento de força resultaria em um erro clinicamente não aceitável.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, IMC= índice de massa corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= American Thoracic Society, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 4** - Características dos estudos sobre o Teste do Degrau (ST) (n = 12) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Chung MM, 2014 <sup>61</sup>	Confiabilidade (estabilidade e equivalência).	86 indivíduos: 45 com AVC crônico e 41 idosos saudáveis (não hospitalizados).	Um degrau de 18 cm, 40 cm de profundidade e 60 cm de largura. O participante estava de frente para o degrau com apoio duplo de perna e sem ajuda de dispositivo. Foi instruído a colocar alternadamente o pé esquerdo e direito no degrau o mais rápido possível, 8 vezes para cada pé em resposta para a contagem de "1. 2. 3. Vá". Cada etapa envolvia a colocação do pé inteiro no degrau e devolvendo-o ao chão.	Avaliação de Extremidade Inferior de Fugl-Meyer (FMA-LE), Teste de Sentar e Levantar 5 vezes (TSO5x), limites de estabilidade, Berg Balance Scale (BBS), Timed Up and Go (TUG).	<ul style="list-style-type: none"> <li>- Foi encontrada excelente confiabilidade intra-avaliador, inter-avaliador e teste- reteste, com uma alteração detectável mínima de 3,26 s.</li> <li>- Os tempos do Teste do Degrau foram significativamente associados à avaliação FMA-LE, tempos do TSO5x, limites de estabilidade nas direções para frente e para trás e para o lado afetado, classificações BBS e TUG.</li> <li>- O tempo do Teste do Degrau é uma ferramenta de avaliação confiável que se correlaciona com diferentes deficiências específicas do AVC em pessoas com AVC crônico.</li> </ul>
Cooney JK, 2013 <sup>62</sup>	Validade de critério e a confiabilidade (estabilidade).	30 indivíduos com artrite reumatoide (não hospitalizados).	Um degrau portátil de 25,4 cm foi utilizado em 3 min, três vezes. Inicialmente, o estágio 1 foi 17 vezes/min; aumentado para 26 e 34 para os estágios 2 e 3, respectivamente. Separado por 1 min de descanso.	Consumo máximo de oxigênio (VO <sub>2</sub> máx).	<ul style="list-style-type: none"> <li>- O VO<sub>2</sub>máx estimada correlacionou-se bem com a VO<sub>2</sub>pico medida diretamente (r = 0,79, LoA ± 5,7 mL/kg·min). A reprodutibilidade final para a VO<sub>2</sub>pico estimada foi excelente (CCI = 0,97, LoA ± 2,2 mL/kg·min).</li> <li>- O teste submáximo de etapas estudado aqui representa um método válido e reproduzível para estimar a aptidão cardiorrespiratória na AR.</li> </ul>
Hayes RM, 2019 <sup>63</sup>	Confiabilidade (estabilidade), Validade de critério.	139 crianças de 5 a 10 anos (não hospitalizados).	Realizado em 22 passos por min. As alturas consistiam em 0,1, 0,2 e 0,3 m. Realizaram por 2 min em cada altura e a frequência cardíaca foi medida após cada altura por 10 s. Foi repetido após 15 min no mesmo dia para reduzir a influência do horário das refeições e outras.	Sexo e IMC, consumo máximo de oxigênio (VO <sub>2</sub> máx).	<ul style="list-style-type: none"> <li>- O Teste do Degrau apresentou excelente confiabilidade (CCI ≥ 0,92; N = 170), independentemente do nível de esforço durante o teste.</li> <li>- Questões significativas de esforço e colinearidade entre as variáveis independentes levou a excluir as crianças de 5 a 6 anos.</li> <li>- Em crianças de 7 a 10 anos com esforço adequado (n = 111) mostrou que ST, sexo e IMC foram preditores de VO<sub>2</sub>máx (R<sup>2</sup> = 0,51).</li> </ul>
Hong SJ, 2012 <sup>64</sup>	Confiabilidade (estabilidade e equivalência).	30 indivíduos com AVC (n = 15) e adultos saudáveis com mais de 50 anos (n = 15) (não hospitalizados).	Em um degrau com 7,5 cm de altura por 15 s. Foram realizados três ensaios com cada perna, com 1 min de descanso. No início do teste, o sujeito colocava o pé inteiro no degrau e devolvia ao chão, era contado um passo. Os indivíduos não foram autorizados a usar qualquer ajuda no teste. Se precisasse de assistência, o teste era interrompido e o número de etapas concluídas registrado.	Força muscular do membro inferior, teste de coordenação motora dos membros inferiores (LEMOCOT).	<ul style="list-style-type: none"> <li>- Excelente confiabilidade intra-avaliador, com CCI de 0,981 a 0,995 e entre avaliadores de 0,996 a 0,999. Ponto de corte de 13 no lado parético foi encontrado para distinguir os saudáveis dos indivíduos com AVC com sensibilidade de 87% e especificidade de 87%. Ponto de corte de 11 no lado não-parético foi encontrado para distinguir os adultos saudáveis dos indivíduos com AVC, sensibilidade de 100% e especificidade de 67%.</li> <li>- O ST do membro parético demonstrou correlação com a força, LEMOCOT da perna parética e a velocidade. Do membro não-parético demonstraram correlação com a força e LEMOCOT da perna parética.</li> </ul>
Knight E, 2014 <sup>65</sup>	Validade de critério (preditiva).	40 indivíduos adultos (não hospitalizados).	Os participantes subiram e desceram um conjunto padronizado de 2 degraus, 20 cm cada, em um ritmo auto-selecionado. Os participantes completaram 1 ou 2 passos de treinos para se familiarizar com o padrão antes de iniciar o teste.	Consumo máximo de oxigênio (VO <sub>2</sub> máx), a frequência cardíaca pós-teste, o peso corporal, a idade e o sexo.	<ul style="list-style-type: none"> <li>- Forte relação entre o VO<sub>2</sub>máx previsto do ST e as medidas diretas do VO<sub>2</sub>máx no teste de esteira máxima (r = 0,78; p &lt; 0,001).</li> <li>- A análise de Bland-Altman demonstrou concordância estatística, embora houvesse superestimação sistemática de 6,4 mL/kg/min.</li> <li>- O ST é uma ferramenta de avaliação da aptidão cardiorrespiratória.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, IMC= índice de massa corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= American Thoracic Society, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 4** - Características dos estudos sobre o Teste do Degrau (ST) (n = 12) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Miyamoto K, 2008 <sup>66</sup>	Confiabilidade (estabilidade), validade de conteúdo (teste de hipóteses), validade de critério.	828 adultos independentes residentes na comunidade (não hospitalizados).	Um degrau de 10 cm de altura. O tempo necessário para repetir 10 rodadas foi medido. Instruídos a realizar a sequência o mais rápido possível. Três pontos: (1) antes de iniciar, a ponta de cada pé deve entrar em contato com o bloco, (2) o calcanhar deve ser colocado no bloco completamente e (3) o salto deve ser evitado. O teste foi realizado duas vezes seguidas.	Tempo de reação motora (MRT), a força isométrica do extensor do joelho, o tempo de permanência em uma perna (SLST).	- Excelente confiabilidade do ST para inter-relação (CCI = 0,96), e a confiabilidade teste-reteste foi suficiente (CCI = 0,86). - A validade de critério foi encontrada através da alta correlação com SLST (r = 0,68) e alta correlação com MRT (r = 0,59). - A validade do conteúdo foi encontrada através de baixa correlação com a força dos extensores do joelho (-0,35) e SLST (-0,36) em 112 mulheres com mais de 70 anos de idade. - A taxa de erro do ST para prever quedas (35,2%) foi menor que a taxa de erro da força muscular (44,4%) e do equilíbrio (38,7%).
Pessoa BV, 2014 <sup>67</sup>	Validade de critério (concorrente e preditiva).	32 indivíduos com Doença Pulmonar Obstrutiva Crônica (DPOC) estágios leves - muito graves (não hospitalizados).	O Teste do Degrau de 6 min (ST6) foi realizado por dois avaliadores, um para comandar o teste e o outro para a contagem. Foi utilizado um degrau de 20 cm de altura, com piso de borracha antiderrapante. O teste seguiu os princípios da ATS para o TC6, com as mesmas frases de incentivo. Os indivíduos foram orientados a subir e descer o degrau por 6 min, objetivando o maior número possível (cadência livre), podendo intercalar os membros inferiores, sem o apoio dos membros superiores.	Teste de Caminhada de 6 min (TC6),  Volume Expiratório Forçado no primeiro segundo (VEF1), IMC, função pulmonar, Body mass index, airway Obstruction, Dyspnea, and Exercise capacity (Índice BODE).	- Validade concorrente: correlação forte positiva (Pearson) entre o número de subidas do primeiro (T1), segundo (T2) e do melhor dos dois (T1 ou T2) ST6 com a distância percorrida no TC6. - Correlações negativas entre os desempenhos no ST6 com o índice BODE e VEF <sub>1</sub> , não foram suficientes para caracterizar validade. - A validade preditiva do ST6 para identificar baixa capacidade física (comparação com o TC6), utilizando o desempenho do T1 ou T2 ou somente do T1, foi considerada acurada para alguns propósitos, sendo a área abaixo da curva de 0,8 (IC 95% 0,62 - 0,98) e 0,85 (IC 95% 0,70 - 0,99). - Ponto de corte de 86 e 78 degraus, apresentando sensibilidade de 90% e 90% e especificidade de 64% e 68%, para T1 ou T2 e somente T1.
Reychler G, 2018 <sup>57</sup>	Validade de critério.	61 crianças saudáveis de 6 a 12 anos (não hospitalizados).	O Teste do Degrau de 6 min (ST6) foi realizado em um degrau de 20 cm de altura sem alças. O teste seguiu as mesmas diretrizes do ATS para o TC6 em relação aos incentivos. As crianças foram instruídas a subir e descer o dispositivo o mais rápido possível em um período de 6 min. A cadência estava livre. O desempenho físico no teste foi determinado pelo número total de subidas e descidas.	Saturação de oxigênio pulsado (SpO <sub>2</sub> ), frequência cardíaca (FC), dispneia e fadiga (PCERT), Teste de Velocidade da Marcha de 4 m (TVM4m), Teste de Caminhada de 6 min (TC6).	- A distância percorrida durante o ST6 foi correlacionada com o número de etapas durante o TC6 (r = 0,320; p = 0,013) e não correlacionada com o TVM4m (r = -0,074; p = 0,575). Não foi encontrada correlação entre o número de etapas durante ST6 e TVM4m (r = -0,129; p = 0,332). - A FC foi menor que a FC máxima após os testes. O aumento da FC foi significativamente maior para o ST6 e significativamente menor para o TVM4m do que para o TC6 e foi maior para o ST6 do que para o TVM4m. - A dispneia e o esforço percebido foram menores após o TC6 do que após o ST6 e maiores que após o TVM4m. Eles também foram maiores durante o ST6 do que durante os TVM4m. O ST6 e o TVM4m não são substitutos válidos para o TC6 em crianças saudáveis.
Thommazo-Luporini L, 2015 <sup>68</sup>	Validade de critério (concorrente).	56 mulheres obesas (não hospitalizadas).	O Teste do Degrau de 6 min (ST6) foi realizado em degrau de 15 cm sem corrimãos. Foram instruídos a subir e descer, iniciando com seu lado dominante, quantas vezes eles pudessem por 6 min e autorizados a desacelerar e/ou parar com base em sintomas de fadiga ou dispneia. Receberam incentivo padronizado a cada minuto, de acordo com a ATS para o TC6.	Teste de Exercício Cardiopulmonar (CPX), VO <sub>2</sub> pico e frequência cardíaca (FC).	- Validade concorrente satisfatória entre o CPX e o ST6, com diferença média de 5,1 ± 3,6 mL/kg·min e 23 ± 13 bpm em relação ao VO <sub>2</sub> pico e FC. - O ST6 demonstrou precisão moderada na previsão de respostas de CPX (área sob a curva: 0,72, IC 95% 0,59 - 0,83). O número de ciclos de passos e o pico de FC durante o ST6 explicaram 31% e 39% da variação total no pico de VO <sub>2</sub> e FC obtidos durante o CPX, respectivamente. - O ST6 é capaz de prever valores metabólicos e cardiovasculares.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, IMC= índice de massa corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= American Thoracic Society, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 4** - Características dos estudos sobre o Teste do Degrau (ST) (n = 12) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Travensolo CF, 2020 <sup>69</sup>	Confiabilidade (estabilidade), validade de critério.	35 indivíduos com doença arterial coronariana (DAC) (não hospitalizados).	O Teste do Degrau de 6 min (ST6) foi realizado em um degrau de 20 cm de altura, 75 cm de largura, 40 cm de largura, sem apoio da mão. Instruídos a subir e descer o máximo de vezes. Utilizadas frases de incentivo da ATS para o TC6. Podiam diminuir a velocidade, parar ou se sentar sem interromper o cronômetro. Realizado duas vezes, 30 min após o primeiro teste.	Teste de Caminhada de 6 min (TC6).	<ul style="list-style-type: none"> <li>- A confiabilidade teste-reteste foi excelente (CCI = 0,967; IC 95% 0,766 - 0,989; p &lt; 0,05). Houve um aumento de 7,9 ± 8,2 etapas do primeiro para o segundo ST6, com erro técnico de medida = 8,0; p &lt; 0,001.</li> <li>- As correlações entre a distância percorrida no TC6 e o número de etapas no TC6 no primeiro e segundo testes foram de 0,6 e 0,7 (p &lt; 0,001), respectivamente. O ST6 é um teste confiável para medir a capacidade funcional em indivíduos com DAC.</li> <li>- O teste é válido ao usar o TC6 como padrão-ouro.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, IMC= índice de massa corporal, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= American Thoracic Society, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 5** - Características dos estudos sobre o Teste de Sentar e Levantar (TSO) (n = 13).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Bruun IH, 2019 <sup>70</sup>	Validade de construto, de critério (concorrente).	207 idosos em departamento de emergência (hospitalizado).	O teste de Teste de Sentar e Levantar de 30 segundos (TSO30s) foi realizado em uma cadeira contando o número de ciclos concluídos em 30 s com as mãos cruzadas contra o peito.	Índice de Mobilidade de Morton (DEMMI): "DEMMI walking" e "DEMMI balance".	<ul style="list-style-type: none"> <li>- Diminuição no TSO30s foi seguida por aumento na ajuda nas atividades diárias. Houve associação entre TSO30s e DEMMI (r = 0,72); para cada repetição extra no TSO30s, a pontuação do DEMMI aumentou 4,9.</li> <li>- Associação com "DEMMI walking" e "balance"; mas com um efeito piso.</li> <li>- O DEMMI tem uma melhor responsividade do que os TSO30s.</li> </ul>
Crook S, 2017 <sup>71</sup>	Confiabilidade (estabilidade), validade do construto.	255 indivíduos com DPOC (não hospitalizados).	Cadeira com 46 - 48 cm, sem braços, contra a parede. As pernas na largura dos quadris e flexionados a 90°, mãos nos quadris sem usar as mãos ou os braços. Instruídos a ficar retos e tocar a cadeira com a parte inferior quando sentados, não precisava se sentar totalmente. Realizado em 1 min. Informados "você tem 15 s até terminar".	Teste de Caminhada de 6 min (TC6), qualidade de vida relacionada à saúde, dispneia e dados cardiorrespiratórios do exercício na admissão e alta.	<ul style="list-style-type: none"> <li>- Correlacionado com o TC6 na admissão (r = 0,59 e 0,64, respectivamente) e alta (r = 0,67 e 0,68). Os CCI (IC95%) entre os TSO foram 0,93 (0,83 - 0,97) para efeito de aprendizado e 0,99 (0,97 - 1,00) para confiabilidade.</li> <li>- As médias de resposta padronizada (IC95%) foram de 0,87 (0,58 - 1,16) e 0,91 (0,78-1,07). A MCID foi de três repetições.</li> <li>- O consumo de O<sub>2</sub>, a produção de CO<sub>2</sub>, a ventilação, a frequência respiratória e cardíaca foram semelhantes nos testes.</li> </ul>
Goldberg UM, 2012 <sup>72</sup>	Confiabilidade (estabilidade).	29 mulheres mais velhas (não hospitalizados).	Os participantes se levantaram de uma cadeira e retornaram à posição sentada o mais rápido possível por cinco repetições (TSO5x). Os braços do participante foram cruzados sobre o peito durante o teste.	Timed up and go (TUG) e testes de alcance funcional.	<ul style="list-style-type: none"> <li>- O coeficiente de correlação de Pearson entre TSO5x e TUG (r = 0,64, p &lt; 0,001) indica que é válido. CCI 2,1 de 0,95: excelente confiabilidade.</li> <li>- O SEM foi de 0,9 S e o MDC 95 foi de 2,5 S para o TSO5x. Os valores de SEM (6,3% da média do TSO5x) e MDC (17,5% da média do TSO5x) baixos, excelente confiabilidade.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= American Thoracic Society, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição

**Tabela 5** - Características dos estudos sobre o Teste de Sentar e Levantar (TSO) (n = 13) (cont.).

Autor, ano	Crítérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Jones CJ, 1999 <sup>73</sup>	Confiabilidade (estabilidade), validade de critério e construto.	76 idosos residentes na comunidade (não hospitalizados).	Cadeira dobrável sem braços, com altura de 43,2 cm, pés de borracha, contra a parede. Sentado, costas retas, pés aproximadamente na largura dos ombros e em um ângulo atrás dos joelhos, com um pé ligeiramente na frente do outro. Braços cruzados contra o peito. No sinal “vá”, levantou-se e depois se sentou, maior número em 30 s.	<i>Keiser LegPress.</i>	<ul style="list-style-type: none"> <li>- O CCI teste-reteste de 0,84 para homens e 0,92 para mulheres.</li> <li>- Correlação moderada a alta entre o desempenho do suporte da cadeira e desempenho máximo do <i>leg-press</i> ajustado ao peso para homens e mulheres (<math>r = 0,78</math> e <math>0,71</math>, respectivamente) apoia a validade de critério.</li> <li>- A validade do construto do suporte da cadeira foi demonstrada.</li> <li>- O desempenho da cadeira diminuiu entre as faixas etárias em décadas - dos 60 aos 70 e 80 (<math>p &lt; 0,01</math>) e foi menor para baixa atividade (<math>p &lt; 0,0001</math>).</li> </ul>
Jones SE, 2013 <sup>74</sup>	Confiabilidade (estabilidade e equivalência), validade de construto.	50 indivíduos com DPOC (não hospitalizados).	Uma cadeira sem braços com encosto reto e assento rígido colocada contra a parede. A altura do chão até o assento era de 48 cm. Sentados com os pés apoiados no chão e membros superiores dobrados sobre o peito. Instruídos a se levantar e sentar uma vez sem usar os membros superiores. Se fosse bem-sucedido, deveriam realizar o mais rápido possível, cinco vezes (TSO5x) sem usar os braços. A cronometragem foi iniciada com o comando “vai”.	<i>Incremental Shuttle Walk Test (ISW),</i> Contração voluntária máxima do quadriceps (QMVC), <i>St. George's Respiratory Questionnaire (SGRQ),</i> <i>Age, Dyspnea, Obstruction Index (ADO)</i> e <i>Body mass index, airway Obstruction, Dyspnea, and Exercise capacity (BODE).</i>	<ul style="list-style-type: none"> <li>- Os CCI teste-reteste e interobservadores foram de 0,97 e 0,99, respectivamente.</li> <li>- O tempo TSO5x correlacionou-se significativamente com ISW, QMVC, SGRQ, ADO e BODE (<math>r = -0,59, -0,38, 0,35, 0,42</math> e <math>0,46</math>, respectivamente; todos <math>p &lt; 0,001</math>). A mediana (25°, 75° centil) do tempo TSO5x diminuiu com RP (pré: 14,1 s (11,5 - 21,3 s) vs. pós: 12,4 s (10,2 - 16,3 s); <math>p &lt; 0,001</math>).</li> <li>- Uma estimativa conservadora para o MCID foi de 1,7 s.</li> <li>- O TSO5x é confiável, válido e responsivo em pacientes com DPOC com um MCID estimado de 1,7 s. É uma medida de resultado funcional prática, adequada para uso na maioria dos estabelecimentos de saúde.</li> </ul>
Kohlbrener D, 2019 <sup>75</sup>	Validade de construto.	38 candidatos a transplante pulmonar (não hospitalizados).	Cadeira sem braços de 46 cm, contra a parede. Instruídos a levantar e se sentar o mais rápido possível. Sem incentivo, mas foram informados quando faltavam 15 s. Uma repetição: pernas retas, e quadril em contato com a cadeira. Mãos nos quadris, não puderam usar mãos ou braços.	Teste de Caminhada de 6 min (TC6), Teste do Degrau de 1 min (ST).	<ul style="list-style-type: none"> <li>- Fortes correlações entre o TSO de 1 min e o TC6 (<math>r = 0,79, P &lt; 0,001</math>) e moderadas entre o ST de 1 min e a força extensora do joelho (<math>r = 0,53, P = 0,001</math>) e entre o TC6 e o joelho força extensora (<math>r = 0,44, P = 0,008</math>).</li> <li>- O TSO de 1 min provocou maior dispneia (<math>P = 0,008</math>) e menor dessaturação de oxigênio em comparação com o TC6 (<math>P &lt; 0,001</math>).</li> <li>- Considerar o uso do TSO quando o TC6 não pode ser realizado.</li> </ul>
Kumban W, 2013 <sup>76</sup>	Confiabilidade (estabilidade e equivalência), validade de critério (concorrente).	33 crianças com paralisia cerebral (PC) (não hospitalizados).	Cadeira ajustável, quadris flexionados a 90° com pés no chão e maléolos laterais 10 cm atrás do joelho. Instruções: “Levante e sente-se cinco vezes (TSO5x) o mais rápido possível, quando disser ‘vá’”. Permitido usar os braços. Realizado 3 vezes, intervalo de 1 min. A média foi utilizada.	<i>Timed up and go (TUG)</i> e a escala de equilíbrio de Berg (BBS).	<ul style="list-style-type: none"> <li>- O CCI para a confiabilidade teste-reteste e interexaminadores do TSO5x foi de 0,91 e 0,88, respectivamente.</li> <li>- TSO5x mostrou correlação moderada com TUG (<math>r = 0,552; P &lt; 0,01</math>) e com BBS (<math>r = -0,561; P &lt; 0,01</math>).</li> <li>- O TSO5x é uma ferramenta de avaliação confiável e se correlaciona com os testes de capacidade de equilíbrio funcional em crianças com PC.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição

**Tabela 5** - Características dos estudos sobre o Teste de Sentar e Levantar (TSO) (n = 13) (cont.).

Autor, ano	Crítérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
McAllister LS, 2019 <sup>77</sup>	Confiabilidade (equivalência), validade de construto (convergente).	30 adultos mais velhos (hospitalizados e não hospitalizados).	Duas cadeiras com 45,7 cm, contra a parede. Na área de internação e no ambulatório. Se não pudesse usava cadeira de rodas. Participante sentado, as costas retas e os pés na largura dos ombros. No sinal "vá", levantava-se e retornava por 30 s (TSO30s). Poderia usar os braços e soltar na postura ereta. Dispositivos não eram permitidos.	Escala de equilíbrio de Berg (BBS), Índice de Barthel modificado.	<ul style="list-style-type: none"> <li>- A confiabilidade entre avaliadores foi CCI 2,1 = 0,737 (P &lt; 0,001).</li> <li>- A confiabilidade intra-avaliador foi calculada CCI 2, k = 0,987 (P &lt; 0,001). Na fase II, validade concorrente para ST30s e BBS inicial (Spearman <math>\rho</math> = 0,737; P = 0,01) e alta (Spearman <math>\rho</math> = 0,727; P = 0,01), bem como as pontuações da escala de Barthel (pontuação inicial Spearman <math>\rho</math> = 0,711; P = 0,01; da alta Spearman <math>\rho</math> = 0,824; P = 0,01).</li> <li>- Não demonstrou correlação com força dos extensores de joelho.</li> </ul>
Møller AB, 2012 <sup>78</sup>	Confiabilidade (estabilidade e equivalência).	22 indivíduos com esclerose múltipla (EM) (não hospitalizados).	Cadeira sem apoio de braço, altura de 45 cm e profundidade de 41 cm. O mais rápido possível em cinco vezes (TSO5x) e em duas tentativas com pausa de 3 min. Sem incentivos. Foram instruídos a iniciar sentados com peso total na cadeira e braços cruzados sobre o peito, pés paralelos e posição ereta com joelhos estendidos.	Força muscular de membros inferiores (MMII), avaliação do equilíbrio.	<ul style="list-style-type: none"> <li>- Correlações significativas entre o TSO5x e a força muscular flexora e extensora isométrica e isocinética do joelho da perna mais afetada (R = -0,60 a -0,77) e entre o ST5x e o desempenho do equilíbrio (R = 0,69).</li> <li>- A variabilidade diária intra-avaliador, a variabilidade intra-avaliador teste-reteste e a variabilidade intra-avaliador no teste foram 25,5, 22,3 e 23,1%.</li> <li>- A variabilidade interavaliadores dentro do teste e a variabilidade interavaliadores foram 23,4 e 5,9%, respectivamente.</li> </ul>
Mong Y, 2010 <sup>79</sup>	Confiabilidade (estabilidade e equivalência).	36 indivíduos: 12 indivíduos com AVC crônico, 12 idosos saudáveis e 12 jovens (não hospitalizados).	Cadeira sem apoio de braço, 43 cm de altura e 47,5 cm de profundidade. Os dois primeiros para familiarização, e a média dos próximos três foi usada, com descanso de 1 min. Instruções: "Até a contagem de 3, levante-se e sente-se o mais rápido possível por cinco vezes (TSO5x). Coloque as mãos no colo e não as use. Apoie as costas contra o encosto no final".	Força muscular de flexores do quadril e flexores e extensores do joelho; de dorsi-flexores do tornozelo e flexores plantares; <i>Berg Balance Scale</i> (BBS); e teste de limites de estabilidade.	<ul style="list-style-type: none"> <li>- Excelente confiabilidade intra-avaliador (intervalo CCI 0,970 a 0,976), interexaminador (CCI = 0,999) e teste-reteste (CCI 0,989 a 0,999).</li> <li>- Os escores do TSO5x também foram associados à força muscular dos flexores do joelho afetados e não afetados (<math>\rho</math> = -0,753 a -0,830; P = 0,00556).</li> <li>- Não foram encontradas associações significativas entre o teste STS de 5 repetições e o BBS e limites de estabilidade.</li> <li>- Escores de corte de 12 s foram discriminatórios entre idosos saudáveis e com AVC com sensibilidade 83% e especificidade 75%.</li> </ul>
Nilsagård Y, 2017 <sup>80</sup>	Validade de critério (concorrente), de construto (convergente e discriminante).	39 indivíduos com esclerose múltipla (não hospitalizados).	Modificados por razões de segurança, permitindo o apoio das mãos no início do levante. O tempo foi medido a partir da posição sentada, quando recebeu o comando "vá" até se sentar após concluir a décima transição (TSO10x). Foi realizado um tempo intermediário após a quinta (TSO5x). Instruídos a realizar rapidamente, mas a levantar-se adequadamente antes de se sentar.	<i>Berg Balance Scale</i> (BBS), <i>Timed Up and Go</i> (TUG), Teste de Velocidade da Marcha de 10 m (TVM10m), Teste de Caminhada de 2 min (TC2), quedas e auxiliares de marcha.	<ul style="list-style-type: none"> <li>- As correlações foram superiores a 0,60 entre TSO e BBS, TUG, TVM10m e TC2 e acima de 0,50 entre TSO e o número de quedas.</li> <li>- Os testes discriminaram aqueles que usaram e não usaram auxiliares de marcha, mas não entre os que caíram e não caíram. Não houve correlações entre TSO e número de quedas ou incidentes próximos.</li> <li>- Os TSO não diferenciaram entre os participantes com resultados alterados e inalterados no BBS. São válidos, mas não identificam quedas e capacidade limitada para alterações.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição

**Tabela 5** - Características dos estudos sobre o Teste de Sentar e Levantar (TSO) (n = 13) (cont.).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Staatjes VE, 2018 <sup>81</sup>	Validade de construto (convergente), confiabilidade (estabilidade).	157 indivíduos com patologias degenerativas da coluna lombar e 80 saudáveis (não hospitalizados).	De acordo com o protocolo de Jones <i>et al.</i> Uma cadeira de 48 cm, sem braços, contra a parede. Instruídos a cruzar os braços no peito e manter os pés no chão, com sapatos estáveis. Para familiarização, realizaram uma vez sem usar os braços. Se fosse necessário, era cancelado. Caso contrário, iniciava no comando “vá”, levante-se e sente-se cinco vezes (TSO5x).	Roland-Morris Disability Questionnaire (RMDQ), Oswestry Disability Index (ODI), escala analógica visual (EVA), Escala Europeia de Qualidade de Vida (EQ-5D)	<ul style="list-style-type: none"> <li>- Correlação direta com RMDQ (<math>r = 0,49</math>), ODI (<math>r = 0,44</math>) e EVA para dor nas costas (<math>r = 0,31</math>) e indireta com o EQ-5D (<math>r = -0,41</math>) (<math>p &lt; 0,001</math>). Não mostrou correlação com EVA para dor nas pernas e EQ-5D EVA (<math>p &gt; 0,05</math>).</li> <li>- Excelente confiabilidade teste-reteste com um CCI de 0,98. Distinguindo indivíduos com e sem comprometimento funcional em 10,35 s.</li> <li>- Uma estratificação de gravidade: 10,5 - 15,2 s, 15,3 - 22,0 s ou maiores que 22,0 s como tendo comprometimento leve, moderado ou grave.</li> <li>- Em 10,4 s pode ser considerado sem comprometimento relevante.</li> </ul>
Wang TH, 2012 <sup>82</sup>	Confiabilidade (estabilidade), validade de construto (convergente).	108 crianças com diplegia espástica e 62 com desenvolvimento típico (não hospitalizados).	Cadeira ajustável sem apoio para braços. Uma corda no nível da altura, descalças sobre um tapete. Posição inicial com o quadril flexionado a 90° e o joelho a 105°. Instrução: “Dobre os braços sobre o peito. Levante-se e sente-se o mais rápido possível por cinco vezes (TSO5x). Toque a corda com a cabeça quando estiver de pé. Se tentar, mas não conseguir se levantar, me avise. Vá!” Praticaram antes três vezes. Realizado três vezes com descanso de 2 min.	Força muscular isométrica, escores da Medida da Função Motora Grossa e função da marcha.	<ul style="list-style-type: none"> <li>- Os CCI de confiabilidade intra-sessão e teste-reteste foram de 0,95 e 0,99, respectivamente. A diferença mínima detectável foi de 0,06 rep/s.</li> <li>- A validade convergente foi apoiada por correlação com o teste <i>sit-to-load</i>, força muscular isométrica, escores da Medida da Função Motora Grossa e função da marcha (<math>r</math> ou <math>\rho = 0,40 - 0,78</math>).</li> <li>- Para validade de grupo conhecida, crianças com desenvolvimento típico e crianças classificadas como Nível I do Sistema de Classificação da Função Motora Grossa apresentaram taxas mais altas de cinco repetições do que as crianças classificadas como nível II, e crianças classificadas como nível II apresentaram taxas mais altas que o nível III.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição

**Tabela 6** - Características dos estudos sobre o Teste *Timed Up and Go* (TUG) (n = 13).

Autor, ano	Critérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Botolfesen P, 2008 <sup>83</sup>	Confiabilidade (estabilidade e equivalência, consistência interna), validade de critério (concorrente).	28 idosos com mobilidade reduzida (não hospitalizados).	Protocolo de Podsiadlo e Richardson (1991). Enquanto Podsiadlo usou apenas o segundo teste, foram incluídos o primeiro e o segundo nas análises. Um cronômetro regular foi usado para registrar a hora (em 1/10 s).	<i>Expanded Timed Up-and-Go</i> (ETUG)	<ul style="list-style-type: none"> <li>- A confiabilidade intra e interavaliadores CCI variou de 0,55 a 0,97, teste-reteste CCI de 0,54 a 0,85. O SEM do tempo (1,96 Sw) foi 2,8 s.</li> <li>- A consistência interna (alfa de Cronbach) foi de 0,74. A correlação (<math>r</math> de Pearson) entre o tempo total do ETUG e o TUG após a correção da atenuação causada pela confiabilidade restrita foi de 0,85.</li> <li>- O ETUG mostrou uma confiabilidade mais alta que o TUG, e a alta validade concorrente sugere que podem ter propriedades semelhantes.</li> </ul>

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição

**Tabela 6** - Características dos estudos sobre o Teste *Timed Up and Go* (TUG) (n = 13) (cont.)

Autor, ano	Crítérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Chan PP, 2017 <sup>84</sup>	Confiabilidade (estabilidade e equivalência).	65 sobreviventes de AVC crônico (n = 33) e idosos saudáveis (n = 32) (não hospitalizados).	A cadeira tinha braços e altura de 43 cm. Um cone a 3 m. O sujeito segurou um copo de água a 3 cm, com a mão não parética. Instruídos a se levantar, caminhar até o cone, virar para o lado parético, voltar e sentar. Quando um avaliador disse “vá”, os dois cronometraram de quando as costas deixaram o encosto até que as nádegas tocaram o assento. Caminharam na própria velocidade três vezes, com calçados habituais.	Avaliação de Fugl-Meyer para membros inferiores;  força do abductor do quadril, flexor do joelho, dorsiflexor do tornozelo e flexor plantar; Teste de Sentar e Levantar cinco vezes (TSO5x); <i>Berg Balance Scale</i> (BBS)	- Excelentes confiabilidades intra-avaliador, interavaliador e teste-reteste. - O TUGmotor correlacionou-se com os escores da Avaliação de Fugl-Meyer para membros inferiores e BBS, com força de abdutor do quadril, flexor do joelho, dorsiflexor do tornozelo e flexor plantar do lado parético, com TSO5x e com o TUG cronometrado convencional. - A mudança mínima detectável no tempo do TUGmotor foi de 3,53 s nos sobreviventes de AVC. Um tempo de corte do TUGmotor de 13,49 s foi encontrado para discriminar melhor o desempenho dos sobreviventes de AVC do que os adultos mais velhos sem AVC.
Evensen NM, 2016 <sup>85</sup>	Validade de construto (convergente).	18 gestantes com dor na cintura pélvica (não hospitalizados).	Protocolo de Podsiadlo e Richardson (1991). Em uma cadeira de 46 cm, as costas no encosto, os braços apoiados e os pés em uma linha no chão. Outra fita a 3 m. Instruções: “Depois de ouvir ‘vá’, se levante e ande o mais rápido possível até cruzar a linha, vire-se, volte e sente”.	Teste de Velocidade da Marcha de 10 m (TVM10m); Teste Ativo de Elevação da Perna (ASLR); Questionário de cintura pélvica (PGQ).	- O coeficiente de correlação de classificação de Spearman (rs) foi usado para determinar a validade convergente. - Fortes correlações foram encontradas entre o TUG e ASLR (rs = 0,73; p = 0,001) e os TVM10m e ASLR (rs = -0,65; p = 0,003). As relações entre o TUG e o PGQ foram moderadas (rs = 0,41 a 0,52) e entre os TVM10m e PGQ de baixa a moderada (rs = -0,25 a -0,56).
Gautschi OP, 2016 <sup>86</sup>	Validade de critério e confiabilidade (estabilidade e equivalência).	2253 indivíduos com doença degenerativa do disco lombar (não hospitalizados).	Protocolo comum. Solicitado que se sentassem e recostassem em uma poltrona, com os braços apoiados. O tempo foi registrado e os indivíduos tiveram que levantar-se e andar o mais rápido possível por uma linha a 3 m da cadeira. Quando chegaram, fizeram uma curva de 180°, voltaram e se sentaram. O tempo parou quando o participante se sentou. Autorizados a usar seus sapatos e auxílio para caminhar, se necessário.	Escala Visual Analógica (EVA), Índice de Incapacidade Roland-Morris, Índice de Incapacidade de Oswestry, pontuação do Formulário 12, Escala Europeia de Qualidade de Vida (EQ-5D).	- Excelente confiabilidade intra (CCI: 0,97) e interobservadores (CCI: 0,99), com SEM de 0,21 e 0,23 s, respectivamente. - Boa correlação para dor nas costas (coeficiente de correlação de Pearson [PCC]: 0,25), dor nas pernas (PCC: 0,29), Índice de Incapacidade Roland-Morris (PCC: 0,38), Índice de Incapacidade de Oswestry (PCC: 0,34), Formulário 12 do Componente Mental (PCC: -0,25), Formulário 12 do Resumo Físico (PCC: -0,32) e EQ-5D [PCC: -0,28]). - O limite superior do “normal” foi de 11,52 s. Mais de 12 s podem ser considerados com comprometimento funcional.
Hwang R, 2016 <sup>87</sup>	Confiabilidade (estabilidade) e validade de critério.	278 indivíduos com insuficiência cardíaca (ICC) (Hospitalizado)	Cadeira de 45 cm com os braços apoiados. Caminharam 3 m em um ritmo confortável, viraram 180°, retornaram e se sentaram. Com calçado regular e dispositivos, realizado duas vezes com descanso até resolução dos sintomas.	Distância de caminhada de 6 min (DTC6), tempo de teste de caminhada de 10 m.	- Em 278 participantes (idade média de 62 anos), o TUG demonstrou excelente confiabilidade teste-reteste (CCI 0,93). Um tempo TUG mais curto (melhor) foi associado a um TC6 mais longo (r = -0,81; P < 0,001) e menor tempo de teste de caminhada de 10 m (rs = 0,80; P < 0,001). - O TUG é uma medida funcional confiável e válida na ICC.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO2pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 6** - Características dos estudos sobre o Teste *Timed Up and Go* (TUG) (n = 13) (cont.)

Autor, ano	Crítérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Kalron A, 2017 <sup>88</sup>	Validade de construto.	285 indivíduos com Esclerose Múltipla (EM) (não hospitalizados).	Usavam calçados habituais e dispositivos. Sentados em uma cadeira de altura padrão com as costas niveladas e braços apoiados. Instruídos a se levantar, andar 3 m, virar, voltar e sentar. Cronometragem começou quando começou a levantar e terminou quando e se sentou.	Teste de Quatro Degraus (ST), Teste de Caminhada de 2 min (TC2), Escala internacional de eficácia de quedas, avaliação cognitiva.	- Diferenças significativas entre os quatro grupos, com exceção dos grupos “muito leve” a “leve”. Os caídores recorrentes demonstraram um TUG significativamente mais lento em comparação aos não-caídores. - Correlação fortemente modestas entre o TUG e testes de caminhada. - As pontuações mais fortes foram encontradas para o ST de 4 degraus e TC2; R de Pearson = 0,740 e -0,731, respectivamente.
Kojima T, 2018 <sup>89</sup>	Validade de critério e confiabilidade (estabilidade).	435 indivíduos com artrite reumatoide (não hospitalizados).	Os indivíduos estavam sentados em uma cadeira de 45 cm de altura e receberam o comando: “Levante-se e vá”, movendo-se 3 m e voltando à cadeira para se sentar novamente a uma velocidade confortável. O melhor tempo de duas medições foi selecionado como valor de TUG.	Índice de Incapacidade do Questionário de Avaliação de Saúde (HAQ-DI), Escala Europeia de Qualidade de Vida (EQ-5D).	- O HAQ-DI médio foi de 1,14 e o TUG de 11,1 s. - O TUG se correlacionou com envelhecimento, EQ-5D e HAQ-DI - Valores médios de TUG foram de 9,0 s (IC 95%, 7,7 - 10,3 s) em pacientes com remissão do HAQ e 8,7 s (7,4 - 10,4 s) nos pacientes com HAQ normal. - TUG para HAQ foi 7,2 s (5,9 - 8,5 s) em jovens ( $\leq$ 61 anos), 9,1 s (7,6 - 10,5 s) em meia idade (62 - 70 anos) e 10,0 s (5,7 - 14,2 s) em idosos ( $\geq$ 71 anos).
Mesquita R, 2013 <sup>90</sup>	Confiabilidade (estabilidade).	235 indivíduos: Doença Pulmonar Obstrutiva Crônica (n = 95) Insuficiência Cardíaca Crônica (n = 68) e insuficiência renal crônica (n = 72) (hospitalizado).	Protocolo de Podsiadlo e Richardson (1991). Solicitou-se que se levantassem de uma cadeira, andassem à velocidade normal de uma distância de 3 m, se virassem, voltassem e sentassem novamente. O tempo começou quando se levantou da cadeira e terminou quando tocou na cadeira ao sentar-se. Três ensaios foram realizados pelo mesmo avaliador e com tempo de descanso suficiente (até a recuperação dos sintomas). Todos os participantes usavam seus próprios calçados, e ajudas para caminhar e oxigênio eram permitidos.	Idade, Índice de Massa Corporal (IMC).	- Boa concordância, em geral, para amostra total e subgrupos, com CCI variando de 0,85 a 0,98 e coeficientes de Kappa de 0,49 a 1,00. - Houve melhora estatística na amostra total do primeiro ao segundo estudo, com grandes limites de concordância (diferença média: -97 s; IC 95% 3,00 a -4,94 s; P < 0,01). - O terceiro teste adicionou pouca ou nenhuma informação aos dois primeiros estudos. - Para a amostra total, um SEM de 1,6 s, um valor absoluto de MDC no nível de confiança de 95% (MDC 95%) de aproximadamente 4,5 s e um valor relativo de MDC no nível de confiança de 95% (MDC 95%) de aproximadamente 35% foram encontrados entre os dois primeiros ensaios, com valores semelhantes para os subgrupos.
Nicolini RD, 2014 <sup>91</sup>	Confiabilidade (estabilidade) e validade de critério.	40 indivíduos com síndrome de Down (SD) (não hospitalizados).	Levantar-se da cadeira sem braços, 90° (DP 10°) de flexão do joelho, pés no chão. Caminhar 3 m até a parede, virar, retornar e se sentar. Três medidas, o resultado foi o menor. Modificações: tocar um alvo na parede; instruções repetidas. Roupas e sapatos confortáveis. Foi feito um treinamento.	Sexo, idade, medida da função motora grossa (GMFM).	- Excelente reprodutibilidade (ICC = 0,82). O tempo foi significativamente maior (p < 0,001) em indivíduos com SD em comparação com sexo, idade, e crianças com desenvolvimento ponderal típico desenvolvendo uma diferença média de -3,53 (IC 95% -4,05 a -3,00). - A dimensão E do GMFM (Caminhada, Corrida e Salto) apresentou a maior correlação (r = -0,55; p < 0,001) com o teste.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO<sub>2</sub>pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição

**Tabela 6** - Características dos estudos sobre o Teste *Timed Up and Go* (TUG) (n = 13) (cont.)

Autor, ano	Crítérios de Validação	Amostra	Teste	Variáveis associadas	Resultados
Ries JD, 2009 <sup>37</sup>	Confiabilidade (estabilidade).	51 indivíduos com doença de Alzheimer (DA) leve a moderada (n = 20); moderada a grave (n = 31) (não hospitalizados).	Levantar-se da cadeira com braços, caminhar 3 m, virar, voltar e se sentar. Circularam em um pequeno cone colocado a distância de 3 m. Instrução: “vá o mais rápido que você conseguir”. O cronômetro começou quando a parte inferior deixou a cadeira e terminou quando a reião posterior fez contato com a cadeira. Duas vezes, separadas por um período de descanso de 30 a 60 min.	Teste de Caminhada de 6 min (TC6) e a velocidade da marcha, <i>Mini-Mental Status Examination</i> e pontuações na escala <i>Functional Assessment Staging</i> (FAST).	- Os valores de confiabilidade teste-reteste para o TUG, o TC6 e a velocidade da marcha foram altos para todos os participantes juntos e para a DA leve a moderada e grupos com DA moderadamente grave a grave separadamente (CCI $\geq 0,973$ ); no entanto, a variabilidade individual do desempenho também foi alta. - Os escores MDC calculados no intervalo de confiança de 90% foram: TUG = 4,09 s, TC6 = 33,5 m (110 pés) e velocidade da marcha = 9,4 cm/s. Os dois grupos foram significativamente diferentes.
Rolenz E, 2016 <sup>92</sup>	Validade de critério (concorrente).	62 idosos com (n = 41) e sem (n = 21) comprometimento cognitivo (não hospitalizados).	Levantar-se de uma cadeira com braços, andar 3 m para uma linha, voltar e sentar. Instruções: “Quando eu disser ‘vá’, quero se levante e caminhe o mais rápido e seguro possível, vire-se, volte e sente-se”. Começou ao comando e parou quando as nádegas entraram em contato com a cadeira. Podiam usar as mãos e dispositivos.	<i>8-Foot Up and Go</i> (8UG), <i>Activities-specific Balance Confidence</i> (ABC), <i>mild cognitive impairment</i> (MCI).	- A sensibilidade do TUG foi de apenas 20% e especificidade de 94,6% e a sensibilidade do 8UG foi de 64% e especificidade de 75,7%. O TUG identificou quedas em taxas diferentes das do 8UG e do ABC (P < 0,05). - O 8UG é recomendado como uma medida de resultado mais apropriada para identificar o risco de queda em idosos da comunidade. - O histórico de queda foi o único preditor do resultado o TUG, 8UG e ABC, indicando que o MCI não é um determinante do desempenho.
Schoppen T, 1999 <sup>93</sup>	Validade de critério (concorrente), confiabilidade (equivalência).	32 indivíduos com amputação unilateral de membros inferiores (não hospitalizados).	Uma poltrona (altura de 46cm, altura do braço 67cm), contra a parede, os braços apoiados e o auxílio para caminhar. Usaram calçado regular e auxílios de locomoção. Na palavra “vá”, se levantou, caminhou até uma linha no chão a 3 m, virou, voltou e se sentou. O final definido quando as nádegas tocaram o assento. Escolheram a própria velocidade confortável e segura.	<i>Sickness Impact Profile</i> , na versão de 68 itens (SIP68) e na <i>Groningen Activity Restriction Scale</i> (GARS).	- O TUG mostrou boa confiabilidade intra e interavaliadores (r = 0,93 e 0,96, respectivamente). - Existe um relacionamento moderado entre o TUG e o GARS, existe um bom relacionamento com as “subescalas físicas” do SIP68 e não há relacionamento com as “subescalas mentais” do SIP68. - O TUG é um instrumento confiável com validade concorrente adequada para medir a mobilidade física de pacientes com amputação da extremidade inferior.
Sebastião E, 2016 <sup>94</sup>	Validade de construto (convergente e discriminante).	47 indivíduos com esclerose múltipla (EM) (não hospitalizados).	Instruídos a concluir rapidamente, levantando-se (sem a ajuda das mãos), caminhando em direção e em torno de um cone/marca no chão; voltando para a cadeira e sentando. Foram autorizados a usar dispositivos auxiliares. Realizaram duas vezes, o tempo foi utilizado.	<i>Multiple Sclerosis Walking Scale</i> , <i>Late-Life Function and Disability Instrument</i> , <i>Expanded disability status scale</i> .	- Fortemente associado a outras medidas válidas de mobilidade ambulatorial (correlação de Spearman, rs = 0,71 a 0,90) e status de incapacidade (rs = 0,80), moderadamente a fortemente à confiança no equilíbrio (rs = 0,66) e fracamente ao controle postural (rs = 0,31). - Moderadamente associado à velocidade de processamento cognitivo (rs = 0,59), mas não a outras medidas não-ambulatorias.

CVF= Capacidade Vital Forçada, DLCO= Capacidade de difusão de monóxido de carbono, CCI= Coeficiente de Correlação Interclasse, m= metros; min: minuto, IC= Intervalo de Confiança, MCID= Diferença Mínima Clinicamente Importante, MDC= Mínima Alteração Detectável, DP= Desvio Padrão, ATS= *American Thoracic Society*, VO2pico = pico de volume de oxigênio máximo, SEM= Erro Padrão de Medição.

**Tabela 7 - Avaliação metodológica dos artigos selecionados (n = 89).**

Estudos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Qualidade
Arcuri JF, 2016 <sup>59</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Arjunan A, 2018 <sup>48</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Bloem AEM, 2018 <sup>49</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Boer PH, 2016 <sup>6</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Bois RM, 2011 <sup>7</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Boa
Bonnevie T, 2019 <sup>60</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Boonstra AM, 1993 <sup>50</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Botolfsen P, 2008 <sup>83</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Bruun IH, 2019 <sup>70</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Buch MH, 2015 <sup>8</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Chan KS, 2015 <sup>9</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Chan PP, 2017 <sup>84</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Chan AW, 2019 <sup>10</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Chen YC, 2018 <sup>11</sup>	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Boa
Chung MM, 2014 <sup>61</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Combs SA, 2014 <sup>51</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Cooney JK, 2013 <sup>62</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Crook S, 2017 <sup>71</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Demers C, 2001 <sup>12</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Boa
Dolmage, 2017 <sup>52</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Dunaway S, 2016 <sup>13</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Eichinger K, 2017 <sup>14</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Elmahgoub SS, 2012 <sup>15</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Evensen NM, 2016 <sup>85</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Fiss AL, 2019 <sup>16</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Fulk GD, 2008 <sup>17</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Gautschi OP, 2016 <sup>86</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Goldberg A, 2012 <sup>72</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Gomes E, 2016 <sup>19</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Guerra-Balic M, 2015 <sup>18</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Gulmans VAM, 1996 <sup>20</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Hayes RM, 2019 <sup>63</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Hong SJ, 2012 <sup>64</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Hwang R, 2016 <sup>87</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Jones CJ, 1999 <sup>73</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Jones SE, 2013 <sup>74</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Kalron A, 2017 <sup>88</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Kim HJ, 2016 <sup>53</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Knight E, 2014 <sup>65</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Kobayashi E, 2016 <sup>21</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Kohlbrenner D, 2019 <sup>75</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Kojima T, 2018 <sup>89</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Kon SS, 2013 <sup>54</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Kumban W, 2013 <sup>76</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Larsson UE, 2008 <sup>22</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Lee MC, 2018 <sup>23</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Li AM, 2005 <sup>24</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Liu WY, 2016 <sup>25</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Mänttari A, 2018 <sup>26</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
McAllister LS, 2019 <sup>77</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
McDonald CM, 2013 <sup>27</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Mesquita R, 2013 <sup>90</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Miyamoto K, 2008 <sup>66</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Møller AB, 2012 <sup>78</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Mong Y, 2010 <sup>79</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Moriello C, 2008 <sup>28</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Boa
Morinder G, 2009 <sup>29</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Mossberg KA, 2012 <sup>30</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Nasuti G, 2013 <sup>31</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Nathan SD, 2015 <sup>32</sup>	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Boa

**Tabela 7** - Avaliação metodológica dos artigos selecionados (n = 89) (cont.).

Estudos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Qualidade
Nicolini RD, 2014 <sup>91</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Nilsagård Y, 2017 <sup>80</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Nolan CM, 2018 <sup>55</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Nsenga LA, 2012 <sup>33</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Pankoff B, 2000 <sup>34</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Pecorelli N, 2016 <sup>35</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Pessoa BV, 2014 <sup>67</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Peters DM, 2013 <sup>36</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Phillips D, 2019 <sup>36</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Reychler G, 2018 <sup>57</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Ries JD, 2009 <sup>37</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Rolenz E, 2016 <sup>92</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Sandberg A, 2020 <sup>38</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Sanjak M, 2017 <sup>39</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Schmidt K, 2013 <sup>40</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Schoppen T, 1999 <sup>93</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Scivoletto L, 2011 <sup>41</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Moderada
Sebastião E, 2016 <sup>94</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Sinclair RC, 2012 <sup>42</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Boa
Staartjes VE, 2018 <sup>81</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Stellmann JP, 2015 <sup>43</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Takken T, 2009 <sup>44</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Thommazo L, 2015 <sup>68</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Travensolo CF, 2019 <sup>69</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Uszko NHMK, 2017 <sup>45</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
van Loo MA, 2003 <sup>58</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada
Vanhelst J, 2013 <sup>46</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Vis JC, 2009 <sup>47</sup>	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Boa
Wang TH, 2012 <sup>82</sup>	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Moderada

1. A questão da pesquisa ou objetivo neste artigo foi claramente indicado?

2. A população do estudo foi claramente especificada e definida?

3. A taxa de participação de pessoas elegíveis foi de pelo menos 50%?

4. Todos os assuntos foram selecionados ou recrutados nas mesmas populações ou semelhantes (incluindo o mesmo período de tempo)? Os critérios de inclusão e exclusão para estar no estudo pré-especificados e aplicados uniformemente a todos os participantes?

5. Uma justificativa de tamanho de amostra, descrição de potência ou estimativas de variação e efeito foram fornecidas?

6. Para as análises neste artigo, a(s) exposição(ões) de interesse foi(ram) medida(s) antes do(s) resultado(s) a ser(em) medido(s)?

7. O prazo era suficiente para que se pudesse razoavelmente esperar uma associação entre exposição e resultado se existisse?

8. Para as exposições que podem variar em quantidade ou nível, o estudo examinou diferentes níveis de exposição como relacionados ao resultado (por exemplo, categorias de exposição ou exposição medida como variável contínua)?

9. As medidas de exposição (variáveis independentes) foram claramente definidas, válidas, confiáveis e implementadas de forma consistente em todos os participantes do estudo?

10. A exposição (s) foi avaliada mais de uma vez ao longo do tempo?

11. As medidas de resultado (variáveis dependentes) foram claramente definidas, válidas, confiáveis e implementadas de forma consistente em todos os participantes do estudo?

12. Os avaliadores dos resultados foram cegados ao status de exposição dos participantes?

13. A perda do seguimento após a linha de base foi de 20% ou menos?

14. As principais variáveis de confusão variáveis foram medidas e ajustadas estatisticamente por seu impacto na relação entre exposição(ões) e resultado(s)?

Y, sim; N, não

Contudo, ainda que possuam algumas limitações, os testes funcionais possuem grande valor clínico para a obtenção de um indicador da capacidade funcional que se relaciona com força muscular, capacidade de caminhar e controle postural, com as atividades de vida diária, risco de quedas, hospitalização e mortalidade. Além de serem medidas gerais de integração entre os sistemas respiratório, cardiovascular e neuromioarticular<sup>97</sup> e de avaliação do impacto de medidas terapêuticas e reabilitadoras<sup>100</sup>.

Quanto às limitações da presente revisão sistemática, são a variabilidade de protocolos de execução dos testes utilizados pelos estudos, o que não permite a adequada comparação dos resultados entre os

artigos, e o fato de que algumas pesquisas não especificaram o critério de confiabilidade (estabilidade, consistência interna e equivalência) e de validade (conteúdo, construto e critério) que foram utilizados, assim como o protocolo de avaliação dessas propriedades psicométricas.

## CONCLUSÃO

Os testes funcionais são instrumentos válidos e confiáveis para a avaliação da funcionalidade de indivíduos hospitalizados e não hospitalizados, estando associados à força muscular, capacidade de caminhar,

controle postural, atividades de vida diária, risco de quedas, hospitalização e mortalidade. Além de fornecer dados para avaliação e acompanhamento de indivíduos com risco de perda funcional e para monitorar resultados de intervenções.

Entretanto, são necessários novos estudos que busquem promover uma maior padronização da

execução dos testes e que considerem os fatores de interação como idade, sexo, altura e peso, causas ortopédicas e comorbidades que podem interferir diretamente na execução e nos resultados das avaliações. Em associação, é fundamental que novas pesquisas forneçam valores de referências para perfis específicos de indivíduos.

## REFERÊNCIAS

- Hajat C, Stein E. The global burden of multiple chronic conditions: A narrative review. *Prev Med Rep.* 2018;12:284-93. <https://doi.org/10.1016/j.pmedr.2018.10.008> PMID:30406006 PMCID:PMC6214883
- Parry SM, Huang M, Needham DM. Evaluating physical functioning in critical care: considerations for clinical practice and Research. *Critical Care.* 2017;21:249. <https://doi.org/10.1186/s13054-017-1827-6> PMID:28978333 PMCID:PMC5628423
- Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol Serv Saúde.* 2017;26(3):649-59. <https://doi.org/10.5123/S1679-49742017000300022> PMID:28977189
- Polit DF, Beck CT. Fundamentos de pesquisa em enfermagem: métodos, avaliação e utilização (7th Ed). Porto Alegre: Artmed; 2011.
- National Institute of Health. Quality assessment tool for observational cohort and cross-sectional studies [Internet Site]. 2017 [cited 2020 Oct 31]. Available from: <https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools>
- Boer PH, Moss SJ. Validity of the 16-metre PACER and six-minute walk test in adults with Down Syndrome. *Disabil Rehabil.* 2016;38(26):2575-83. <https://doi.org/10.3109/09638288.2015.1137982> PMID:26860699
- Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, et al. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *Am J Respir Crit Care Med.* 2011;183(9):1231-7. <https://doi.org/10.1164/rccm.201007-1179OC> PMID:21131468
- Buch MH, Denton CP, Furst DE, Guillevin L, Rubin LJ, Wells AU, et al. Submaximal exercise testing in the assessment of interstitial lung disease secondary to systemic sclerosis: reproducibility and correlations of the 6-min walk test. *Ann Rheum Dis.* 2007;66(2):169-73. <https://doi.org/10.1136/ard.2006.054866> PMID:16868020 PMCID:PMC1798506
- Chan KS, Pfoh ER, Denehy L, Elliott D, Holland AE, Dinglas VD, et al. Construct validity and minimal important difference of 6-minute walk distance in survivors of acute respiratory failure. *Chest.* 2015;147(5):1316-26. <https://doi.org/10.1378/chest.14-1808> PMID:25742048 PMCID:PMC4420183
- Chan AW, Pinb TW. Reliability, validity and minimal detectable change of 2-minute walk test, 6-minute walk test and 10-meter walk test in frail older adults with dementia. *Exp Gerontol.* 2019;115:9-18. <https://doi.org/10.1016/j.exger.2018.11.001> PMID:30423359
- Chen YC, Chen KC, Lu LH, Wu YL, Lai TJ, Wang CY. Validating the 6-minute walk test as an indicator of recovery in patients undergoing cardiac surgery. A prospective cohort study. *Medicine.* 2018;97(42):e12925. <https://doi.org/10.1097/MD.00000000000012925> PMID:30335028 PMCID:PMC6211883
- Demers C, McKelvie RS, Negassa A, Yusuf S. Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. *Am Heart J.* 2001;142(4):698-703. <https://doi.org/10.1067/mhj.2001.118468> PMID:11579362
- Dunaway YS, Montes J, Kramer SS, Marra J, Salazar R, Cruz R, et al. Six-minute walk test is reliable and valid in spinal muscular atrophy. *Muscle Nerve.* 2016;54(5):836-42. <https://doi.org/10.1002/mus.25120> PMID:27015431
- Eichinger K, Heatwole C, Heining S, Stinson N, Matichak Stock C, Grossmann C, et al. Validity of the 6 minute walk test in facioscapulohumeral muscular dystrophy. *Muscle Nerve.* 2017;55(3):333-7. <https://doi.org/10.1002/mus.25251> PMID:27421252 PMCID:PMC5793210
- Elmahgoub SS, Van De Velde A, Peersman W, Cambier D, Calders P. Reproducibility, validity and predictors of six-minute walk test in overweight and obese adolescents with intellectual disability. *Disabil Rehabil.* 2012;34(10):846-51. <https://doi.org/10.3109/09638288.2011.623757> PMID:22149772
- Fiss AL, Jeffries L, Yocum A, Westcott MS. Validity of the Early Activity Scale for Endurance and the 6-Minute Walk Test for Children with cerebral palsy. *Pediatr Phys Ther.* 2019;31(2):156-63. <https://doi.org/10.1097/PEP.0000000000000577> PMID:30865148
- Fulk GD, Echternach JL, Nof L, O'Sullivan S. Clinometric properties of the six-minute walk test in individuals undergoing rehabilitation poststroke. *Physiother Theory Pract.* 2008;24(3):195-04. <https://doi.org/10.1080/09593980701588284> PMID:18569856
- Guerra BM, Oviedo GR, Javierre C, Fortuño J, Barnet-López S, Niño O, et al. Reliability and validity of the 6-min walk test in adults and seniors with intellectual disabilities. *Res Dev Disabil.* 2015;47:144-53. <https://doi.org/10.1016/j.ridd.2015.09.011> PMID:26426514
- Gomes E, Bastos T, Probst M, Ribeiro JC, Silva G, Corredeira R. Reliability and validity of 6MWT for outpatients with schizophrenia: A preliminary study. *Psychiatry Res.* 2016; 237:37-42. <https://doi.org/10.1016/j.psychres.2016.01.066> PMID:26921049
- Gulmans VA, Van Veldhoven NH, De Meer K, Helders PJ. The six-minute walking test in children with cystic fibrosis: reliability and validity. *Pediatr Pulmonol.* 1996; 22(2):85-9. [https://doi.org/10.1002/\(SICI\)1099-0496\(199608\)22:2<85::AID-PPUL1>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1099-0496(199608)22:2<85::AID-PPUL1>3.0.CO;2-1)
- Kobayashi E, Himuro N, Takahashi M. Clinical utility of the 6-min walk test for patients with moderate Parkinson's disease. *Int J Rehabil Res.* 2017;40(1):66-70. <https://doi.org/10.1097/MRR.0000000000000205> PMID:27977464
- Larsson EU, Reynisdottir S. The six-minute walk test in outpatients with obesity: reproducibility and known group validity. *Physiother Res Int.* 2008;13(2):84-93. <https://doi.org/10.1002/pri.398> PMID:18446882
- Lee MC. Validity of the 6-minute walk test and step test for evaluation of cardio respiratory fitness in patients with type 2 diabetes mellitus. *J Exerc Nutrition Biochem.* 2018;22(1):49-55. <https://doi.org/10.20463/jenb.2018.0008> PMID:29673246 PMCID:PMC5909078
- Li AM, Yin J, Yu CC, Tsang T, So HK, Wong E, et al. The six-minute walk test in healthy children: reliability and validity. *Eur Respir J.* 2005;25(6):1057-60. <https://doi.org/10.1183/09031936.05.00134904> PMID:15929962
- Liu WY, Meijer K, Delbressine JM, Willems PJ, Franssen FM, Wouters EF, et al. Reproducibility and validity of the 6-Minute Walk Test using the Gait Real-Time Analysis Interactive Lab in

- patients with COPD and healthy elderly. *PLoS One*. 2016;11(9):e0162444. <https://doi.org/10.1371/journal.pone.0162444> PMID:27607426 PMCID:PMC5015964
26. Mänttari A, Suni J, Sievänen H, Husu P, Vähä-Ypyä H, Valkeinen H, et al. Six-minute walk test: a tool for predicting maximal aerobic power (VO<sub>2</sub> max) in healthy adults. *Clin Physiol Funct Imaging*. 2018. <https://doi.org/10.1111/cpf.12525> PMID:29851229
  27. McDonald CM, Henricson EK, Abresch RT, Florence J, Eagle M, Gappmaier E, et al. The 6-minute walk test and other clinical endpoints in duchenne muscular dystrophy: reliability, concurrent validity, and minimal clinically important differences from a multicenter study. *Muscle Nerve*. 2013;48(3):357-68. <https://doi.org/10.1002/mus.23905> PMID:23674289 PMCID:PMC3826053
  28. Moriello C, Mayo NE, Feldman L, Carli F. Validating the six-minute walk test as a measure of recovery after elective colon resection surgery. *Arch Phys Med Rehabil*. 2008;89(6):1083-9. <https://doi.org/10.1016/j.apmr.2007.11.031> PMID:18503803
  29. Morinder G, Mattsson E, Sollander C, Marcus C, Larsson UE. Six-minute walk test in obese children and adolescents: reproducibility and validity. *Physiother Res Int*. 2009;14(2):91-104. <https://doi.org/10.1002/pri.428> PMID:19003813
  30. Mossberg KA, Fortini E. Responsiveness and validity of the six-minute walk test in individuals with traumatic brain injury. *Phys Ther*. 2012; 92(5):726-33. <https://doi.org/10.2522/ptj.20110157> PMID:22282772 PMCID:PMC3345338
  31. Nasuti G, Stuart-Hill L, Temple VA. The Six-Minute Walk Test for adults with intellectual disability: a study of validity and reliability. *J Intellect Dev Disabil*. 2013;38(1):31-8. <https://doi.org/10.3109/13668250.2012.748885> PMID:23244742
  32. Nathan SD, du Bois RM, Albera C, Bradford WZ, Costabel U, Kartashov A, et al. Validation of test performance characteristics and minimal clinically important difference of the 6-minute walk test in patients with idiopathic pulmonary fibrosis. *Respir Med*. 2015;109(7):914-22. <https://doi.org/10.1016/j.rmed.2015.04.008> PMID:25956020
  33. Nsenga LA, Shephard RJ, Ahmaid S. Six-minute walk test in children with cerebral palsy gross motor function classification system levels I and II: reproducibility, validity, and training effects. *Arch Phys Med Rehabil*. 2012;93(12):2333-9. <https://doi.org/10.1016/j.apmr.2012.06.005> PMID:22721868
  34. Pankoff B, Overend T, Lucy D, White K. Validity and responsiveness of the 6 minute walk test for people with fibromyalgia. *J Rheumatol*. 2000;27(11):2666-70.
  35. Pecorelli N, Fiore JF, Gillis C, Awasthi R, Mappin-Kasirer B, Niculiseanu P, et al. The six-minute walk test as a measure of postoperative recovery after colorectal resection: further examination of its measurement properties. *Surg Endosc*. 2016;30(6):2199-206. <https://doi.org/10.1007/s00464-015-4478-1> PMID:26310528
  36. Phillips D, Tomazos LC, Moseley S, L'italien G, Silva HG, Lara SL. Reliability and validity of the 6-Minute Walk Test in hypophosphatasia. *JBM Plus*. 2019;3(6):e10131. <https://doi.org/10.1002/jbm4.10131> PMID:31346563 PMCID:PMC6636775
  37. Ries JD, Echternach JL, Nof L, Blodgett MG. Test-retest reliability and minimal detectable change scores for the Timed "Up & Go" Test, the Six-Minute Walk Test, and Gait Speed in people with Alzheimer Disease. *Phys Ther*. 2009;89(6):569-79. <https://doi.org/10.2522/ptj.20080258> PMID:19389792
  38. Sandberg A, Cider Å, Jivegård L, Nordanstig J, Wittboldt S, Bäck M. Test-retest reliability, agreement, and minimal detectable change in the 6-minute walk test in patients with intermittent claudication. *J Vasc Surg*. 2020;71(1):197-203. <https://doi.org/10.1016/j.jvs.2019.02.056> PMID:31147129
  39. Sanjak M, Langford V, Holsten S, Rozario N, Patterson CGM, Bravver E, et al. Six-Minute Walk Test as a measure of walking capacity in ambulatory individuals with amyotrophic lateral sclerosis. *Arch Phys Med Rehabil*. 2017;98(11):2301-7. <https://doi.org/10.1016/j.apmr.2017.04.004> PMID:28465220
  40. Schmidt K, Vogt L, Thiel C, Jäger E, Banzer W. Validity of the six-minute walk test in cancer patients. *Int J Sports Med*. 2013;34(7):631-6. <https://doi.org/10.1055/s-0032-1323746> PMID:23444095
  41. Scivoletto G, Tamburella F, Laurenza L, Foti C, Ditunno JF, Molinari M. Validity and reliability of the 10-m walk test and the 6-min walk test in spinal cord injury patients. *Spinal Cord*. 2011;49(6):736-40. <https://doi.org/10.1038/sc.2010.180> PMID:21221120
  42. Sinclair RCF, Batterham AM, Davies S, Cawthorn L, Danjoux GR. Validity of the 6 min walk test in prediction of the anaerobic threshold before major non-cardiac surgery. *Br J Anaesth*. 2012;108(1):30-5. <https://doi.org/10.1093/bja/aer322> PMID:21980122
  43. Stellmann JP, Neuhaus A, Götze N, Briken S, Lederer C, Schimpl M, et al. Ecological validity of walking capacity tests in multiple sclerosis. *PLoS One*. 2015;10(4):e0123822. <https://doi.org/10.1371/journal.pone.0123822> PMID:25879750 PMCID:PMC4399985
  44. Takken T, Engelbert R, Van Bergen M, Groothoff J, Nauta J, Van Hoek K, et al. Six-minute walking test in children with ESRD: discrimination validity and construct validity. *Pediatr Nephrol*. 2009;24(11):2217-23. <https://doi.org/10.1007/s00467-009-1259-x> PMID:19633871 PMCID:PMC2753769
  45. Uszko-Lencer NHMK, Mesquita R, Janssen E, Werter C, Brunner-La RHP, Pitta F, et al. Reliability, construct validity and determinants of 6-minute walk test performance in patients with chronic heart failure. *Int J Cardiol*. 2017;240:285-290. <https://doi.org/10.1016/j.ijcard.2017.02.109> PMID:28377186
  46. Vanhelst J, Fardy PS, Salleron J, Béghin L. The six-minute walk test in obese youth: reproducibility, validity, and prediction equation to assess aerobic power. *Disabil Rehabil*. 2013;35(6):479-82. <https://doi.org/10.3109/09638288.2012.699581> PMID:22779759
  47. Vis JC, Thoonsen H, Duffels MG, De Bruin-Bom RA, Huisman SA, Van Dijk AP, et al. Six-minute walk test in patients with Down syndrome: validity and reproducibility. *Arch Phys Med Rehabil*. 2009;90(8):1423-7. <https://doi.org/10.1016/j.apmr.2009.02.015> PMID:19651279
  48. Arjunan A, Peel NM, Hubbard RE. Gait speed and frailty status in relation to adverse outcomes in geriatric rehabilitation. *Arch Phys Med Rehabil*. 2019;100(5):859-64. <https://doi.org/10.1016/j.apmr.2018.08.187> PMID:30312596
  49. Bloem AEM, Veltkamp M, Spruit MA, Custers JWH, Bakker EWP, Dolk HM, et al. Validation of 4-meter-gait-speed test and 5-repetitions-sit-to-stand test in patients with pulmonary fibrosis: a clinimetric validation study. *Sarcoidosis Vasc Diffuse Lung Dis*. 2018;35(4):317-326. doi: <https://doi.org/10.36141/svld.v35i4.7035>
  50. Boonstra AM, Fidler V, Eisma WH. Walking speedmmsubjects and amputees: aspects of validity of gait analysis. *Prosthet Orthot Int*. 1993;17(2):78-82. <https://doi.org/10.3109/03093649309164360> PMID:8233772
  51. Combs AS, Diehl MD, Filip J, Long E. Short-distance walking speed tests in people with Parkinson disease: Reliability, responsiveness, and validity. *Gait Posture*. 2014;39(2):784-8. <https://doi.org/10.1016/j.gaitpost.2013.10.019> PMID:24246801
  52. Dolmage TE, Evans RA, Goldstein RS. Validity of measurement in patients with severe chronic lung disease, including prognostic and practical implications. *CHEST*. 2017;153(5):1101-5. <https://doi.org/10.1016/j.chest.2017.11.024> PMID:29222008
  53. Kim HJ, Park Y, Lee HJ. The reliability and validity of gait speed with different walking pace and distances against general health, physical function, and chronic disease in aged adults. *J Exerc Nutrition Biochem*. 2016;20(3):46-50. <https://doi.org/10.20463/jenb.2016.09.20.3.7> PMID:27757387 PMCID:PMC5067420
  54. Kon SS, Patel MS, Canavan JL, Clark AL, Jones SE, Nolan CM, et al. Reliability and validity of 4-metre gait speed in COPD. *Eur Respir J*. 2013;42(2):333-40. <https://doi.org/10.1183/09031936.00162712> PMID:23222875
  55. Nolan CM, Maddocks M, Maher TM, Canavan JL, Jones SE,

- Barker RE, et al. Phenotypic characteristics associated with slow gait speed in idiopathic pulmonary fibrosis. *Respirology*. 2018;23(5):498-506. <https://doi.org/10.1111/resp.13213> PMID:29139195
56. Peters DM, Fritz SL, Krotish DE. Assessing the reliability and validity of a shorter walk test compared with the 10-Meter Walk Test for measurements of gait speed in healthy, older adults. *J Geriatr Phys Ther*. 2013;36(1):24-30. <https://doi.org/10.1519/JPT.0b013e318248e20d> PMID:22415358
  57. Reychler G, Audag N, Dewulf S, Morale MN, Caty G. Validation of 6 min step test and 4-m gait speed in children: A randomized cross-over study. *Gait Posture*. 2018;61:19-24. <https://doi.org/10.1016/j.gaitpost.2017.12.011> PMID:29289866
  58. Van Loo MA, Moseley AM, Bosman JM, De Bie RA, Hassett L. Inter-rater reliability and concurrent validity of walking speed measurement after traumatic brain injury. *Clin Rehabil*. 2003;17(7):775-9. <https://doi.org/10.1191/0269215503cr6770a> PMID:14606745
  59. Arcuri JF, Borghi-Silva A, Labadessa IG, Sentanin AC, Candolo C, Pires Di Lorenzo VA. Validity and reliability of the 6-Minute Step Test in healthy individuals: a cross-sectional study. *Clin J Sport Med*. 2016;26(1):69-75. <https://doi.org/10.1097/JSM.000000000000190> PMID:25706661
  60. Bonnevie T, Allingham M, Prieur G, Combret Y, Debeaumont D, Patout M, et al. The six-minute stepper test is related to muscle strength but cannot substitute for the one repetition maximum to prescribe strength training in patients with COPD. *International journal of chronic obstructive pulmonary disease*, 2019;14:767-774. <https://doi.org/10.2147/COPD.S193585> PMID:30992662 PMCid:PMC6445245
  61. Chung MM, Chan RW, Fung, YK, Fong SS, Lam SS, Lai CW, et al. Reliability and validity of Alternate Step Test times in subjects with chronic stroke. *J Rehabil Med*. 2014;46(10):969-74. <https://doi.org/10.2340/16501977-1877> PMID:25167536
  62. Cooney JK, Moore JP, Ahmad YA, Jones JG, Lemmey AB, Casanova F, et al. A simple step test to estimate cardio-respiratory fitness levels of rheumatoid arthritis patients in a clinical setting. *Int J Rheumatol*. 2013;2013:174541. <https://doi.org/10.1155/2013/174541> PMID:24454385 PMCid:PMC3884971
  63. Hayes RM, Maldonado D, Gossett T, Shepherd T, Mehta SP, Flesher SL. Developing and validating a step test of aerobic fitness among elementary school children. *Physiother Can*. 2019;71(2):187-94. <https://doi.org/10.3138/ptc.2017-44.pp> PMID:31040514 PMCid:PMC6484966
  64. Hong SJ, Goh EY, Chua SY, Ng SS. Reliability and validity of step test scores in subjects with chronic stroke. *Arch Phys Med Rehabil*. 2012;93(6):1065-71. <https://doi.org/10.1016/j.apmr.2011.12.022> PMID:22475055
  65. Knight E, Stuckey MI, Petrella RJ. Validation of the Step Test and Exercise Prescription Tool for adults. *Can J Diabetes*, 2014;38(3):164-71. <https://doi.org/10.1016/j.jcjd.2014.03.007> PMID:24909087
  66. Miyamoto K, Takebayashi H, Takimoto K, Miyamoto S, Morioka S, Yagi F. A new simple performance test focused on agility in elderly people: The Ten Step Test. *Gerontology*. 2008;54(6):365-72. <https://doi.org/10.1159/000146787> PMID:18645242
  67. Pessoa BV, Arcuri JF, Labadessa IG, Costa JNF, Sentanin AC, Pires VA, et al. Validity of the six-minute step test of free cadence in patients with chronic obstructive pulmonary disease. *Braz J Phys Ther*. 2014;18(3):228-36. <https://doi.org/10.1590/bjpt-rbf.2014.0041> PMID:25003275 PMCid:PMC4183495
  68. Di Thommazo-Luporini L, Carvalho LP, Luporini RL, Trimer R, Pantoni CB, Martinez AF, et al. Are cardiovascular and metabolic responses to field walking tests interchangeable and obesity-dependent? *Disabil Rehabil*. 2016;38(18):1820-9. <https://doi.org/10.3109/09638288.2015.1107645> PMID:26681386
  69. Travençolo CF, Arcuri JF, Polito MD. Validity and reliability of the 6-min step test in individuals with coronary artery disease. *Physiother Res Int*. 2020;25(1):e1810. <https://doi.org/10.1002/pri.1810> PMID:31599079
  70. Bruun IH, Mogensen CB, Nørgaard B, Schiøttz-Christensen B, Maribo T. Validity and responsiveness to change of the 30-Second Chair-Stand Test in older adults admitted to an emergency department. *J Geriatr Phys Ther*. 2019;42(4):265-74. <https://doi.org/10.1519/JPT.000000000000166> PMID:29227412
  71. Crook S, Büsching G, Schultz K, Leibert N, Jelusic D, Keusch S, et al. A multicentre validation of the 1-min sit-to-stand test in patients with COPD. *Eur Respir J*. 2017;49(3):1601871. <https://doi.org/10.1183/13993003.01871-2016> PMID:28254766
  72. Goldberg A, Chavis M, Watkins J, Wilson T. The five-times-sit-to-stand test: validity, reliability and detectable change in older females. *Aging Clin Exp Res*. 2012;24(4):339-44. <https://doi.org/10.1007/BF03325265> PMID:23238309
  73. Jones JC, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Res Q Exerc Sport*. 1999;70(2):113-9. <https://doi.org/10.1080/02701367.1999.10608028> PMID:10380242
  74. Jones SE, Kon SS, Canavan JL, Patel MS, Clark AL, Nolan CM, et al. The five-repetition sit-to-stand test as a functional outcome measure in COPD. *Thorax*. 2013;68(11):1015-20. <https://doi.org/10.1136/thoraxjnl-2013-203576> PMID:23783372
  75. Kohlbrenner D, Benden C, Radtke T. The 1-minute sit-to-stand test in lung transplant candidates: an easy-to-use alternative to the 6-minute walk test? *Respir Care*. 2020;65(4):437-43. <https://doi.org/10.4187/respcare.07124> PMID:31641072
  76. Kumban W, Amatachaya S, Emasithi A, Siritarativat W. Five-times-sit-to-stand test in children with cerebral palsy: reliability and concurrent validity. *NeuroRehabilitation*. 2013;32(1):9-15. <https://doi.org/10.3233/NRE-130818> PMID:23422454
  77. Mcallister LS, Palombaro KM. Modified 30-second sit-to-stand test: reliability and validity in older adults unable to complete traditional sit-to-stand testing. *J Geriatr Phys Ther*. 2020;43(3):153-8. <https://doi.org/10.1519/JPT.000000000000227> PMID:30807554
  78. Møller AB, Bibby BM, Skjærbaek AG, Jensen E, Sørensen H, Stenager E, et al. Validity and variability of the 5-repetition sit-to-stand test in patients with multiple sclerosis. *Disabil Rehabil*. 2012;34(26):2251-8. <https://doi.org/10.3109/09638288.2012.683479> PMID:22612360
  79. Mong Y, Teo TW, Ng SS. 5-repetition sit-to-stand test in subjects with chronic stroke: reliability and validity. *Arch Phys Med Rehabil*. 2010;91(3):407-13. <https://doi.org/10.1016/j.apmr.2009.10.030> PMID:20298832
  80. Nilsagård Y, Andreasson M, Carling A, Vesterlin H. Examining the validity and sensitivity to change of the 5 and 10 sit-to-stand tests in people with multiple sclerosis. *Physiother Res Int*. 2017;22(4). <https://doi.org/10.1002/pri.1681> PMID:28044398
  81. Staartjes VE, Schröder ML. The five-repetition sit-to-stand test: evaluation of a simple and objective tool for the assessment of degenerative pathologies of the lumbar spine. *J Neurosurg Spine*. 2018;29(4):380-7. <https://doi.org/10.3171/2018.2.SPINE171416> PMID:29957147
  82. Wang TH, Liao HF, Peng YC. Reliability and validity of the five-repetition sit-to-stand test for children with cerebral palsy. *Clin Rehabil*. 2012;26(7):664-71. <https://doi.org/10.1177/0269215511426889> PMID:22080526
  83. Botolfson P, Helbostad JL, Moe-Nilssen R, Wall JC. Reliability and concurrent validity of the Expanded Timed Up-and-Go test in older people with impaired mobility. *Physiother Res Int*. 2008;13(2):94-106. <https://doi.org/10.1002/pri.394> PMID:18288773
  84. Chan PP, Si Tou JI, Tse MM, Ng SS. Reliability and validity of the Timed Up and Go Test with a motor task in people with chronic stroke. *Arch Phys Med Rehabil*. 2017;98(11):2213-20. <https://doi.org/10.1016/j.apmr.2017.03.008> PMID:28392324
  85. Evensen NM, Kvåle A, Brækken IH. Convergent validity of the Timed Up and Go test and Ten-metre Timed Walk Test in pregnant women with pelvic girdle pain. *Manual Therapy*,

- 2015; 21:94-9. <https://doi.org/10.1016/j.math.2015.06.002> PMID:26235131
86. Gautschi OP, Smoll NR, Corniola MV, Joswig H, Chau I, Hildebrandt G, et al. Validity and reliability of a measurement of objective functional impairment in lumbar degenerative disc disease: The Timed Up and Go (TUG) Test. *Neurosurgery*. 2016;79(2):270-7. <https://doi.org/10.1227/NEU.0000000000001195> PMID:26702840
87. Hwang R, Morris NR, Mandrusiak A, Mudge A, Suna J, Adsett J, et al. Timed Up and Go test: a reliable and valid test in patients with chronic heart failure. *J Card Fail*. 2016; 22(8):646-50. <https://doi.org/10.1016/j.cardfail.2015.09.018> PMID:26456063
88. Kalron A, Dolev M, Givon U. Further construct validity of the Timed Up-and-Go Test as a measure of ambulation in multiple sclerosis patients. *Eur J Phys Rehabil Med*. 2017;53(6):841-7. <https://doi.org/10.23736/S1973-9087.17.04599-3>
89. Kojima T, Ishikawa H, Tanaka S, Haga N, Nishida K, Yukioka M, et al. Validation and reliability of the Timed Up and Go test for measuring objective functional impairment in patients with long-standing rheumatoid arthritis: a cross-sectional study. *Int J Rheum Dis*. 2018;21(10):1793-800. <https://doi.org/10.1111/1756-185X.13237> PMID:29210204
90. Mesquita R, Janssen DJ, Wouters EF, Schols JM, Pitta F, Spruit MA. Within-day test-retest reliability of the Timed Up & Go test in patients with advanced chronic organ failure. *Arch Phys Med Rehabil*. 2013;94(11):2131-8. <https://doi.org/10.1016/j.apmr.2013.03.024> PMID:23583345
91. Nicolini-Panisson RD, Donadio MV. Normative values for the Timed 'Up and Go' test in children and adolescents and validation for individuals with Down syndrome. *Dev Med Child Neurol*. 2014;56(5):490-7. <https://doi.org/10.1111/dmcn.12290> PMID:24206424
92. Rolenz E, Reneker JC. Validity of the 8-Foot Up and Go, Timed Up and Go, and Activities-Specific Balance Confidence Scale in older adults with and without cognitive impairment. *J Rehabil Res Dev*. 2016;53(4):511-518. <https://doi.org/10.1682/JRRD.2015.03.0042> PMID:27532337
93. Schoppen T, Boonstra A, Groothoff JW, De Vries J, Göeken LN, Eisma WH. The Timed "up and go" test: reliability and validity in persons with unilateral lower limb amputation. *Arch Phys Med Rehabil*. 1999;80(7):825-8. [https://doi.org/10.1016/S0003-9993\(99\)90234-4](https://doi.org/10.1016/S0003-9993(99)90234-4)
94. Sebastião E, Sandroff BM, Learmonth YC, Motl RW. Validity of the Timed Up and Go Test as a measure of functional mobility in persons with multiple sclerosis. *Arch Phys Med Rehabil*. 2016;97(7):1072-7. <https://doi.org/10.1016/j.apmr.2015.12.031> PMID:26944709
95. Ito T. Comprehensive physical function assessment in elderly people. *Clin Phys Ther*. 2017. <https://doi.org/10.5772/67528>
96. Christakou A, Papadopoulos E, Patsaki I, Sidiras G, Nanas S. Functional Assessment Scales in a General Intensive Care Unit. A Review. *Hospital Chronicles*. 2013;8(4):164-170.
97. Heresi GA, Dweik RA. Strengths and limitations of the Six-Minute-Walk Test: a model biomarker study in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2011;183(9):1122-4. <https://doi.org/10.1164/rccm.201012-2079ED> PMID:21531951
98. Novaes RD, Miranda AS, Dourado VZ. Velocidade usual da marcha em brasileiros de meia idade e idosos. *Rev Bras Fisioter*. 2011;15(2):117-22. <https://doi.org/10.1590/S1413-35552011000200006> PMID:21789361
99. Keat BM, Guck TP, McGaha AL. Timed Up and Go (TUG) Test: normative reference values for ages 20 to 59 years and relationships with physical and mental health risk factors. *J Prim Care Community Health*. 2017;8(1):9-13. <https://doi.org/10.1177/2150131916659282> PMID:27450179 PMID:PMC5932649
100. Sociedade Brasileira de Pneumologia e Tisiologia (SBPT) [Internet]. Teste da Caminhada e do Degrau. Temas em revisão, 2011 [uploaded Dec 2011; cited 2020 Oct 31]. Available from: [http://itarget.com.br/newclients/sbpt.org.br/2011/downloads/arquivos/Revisoes/REVISAO\\_06\\_TESTE\\_CAMINHADA\\_DEGRA\\_U.pdf](http://itarget.com.br/newclients/sbpt.org.br/2011/downloads/arquivos/Revisoes/REVISAO_06_TESTE_CAMINHADA_DEGRA_U.pdf)

**Conflitos de interesse:** Os autores informam não haver conflitos de interesse relacionados a este artigo.

**Contribuição individual dos autores:**

Concepção e desenho do estudo: NCS

Análise e interpretação dos dados: NCS, DBC

Coleta de dados: NCS, DBC

Redação do manuscrito: NCS

Revisão crítica do texto: NSS, JMA, BSM

Aprovação final do manuscrito\*: NSS, JMA, BSM, NCS, DCB

Análise estatística: Não se aplica

Responsabilidade geral pelo estudo: NCS, NSS, JMA

\*Todos os autores leram e aprovaram a versão final do manuscrito submetido para publicação da Rev Cienc Saude.

**Informações sobre financiamento:** não se aplica.