Wheelchair Skills tests: a systematic review
Kilkens, O.J.E.; Post, M.W.M.; Seelen, H.A.M.; Dallmeijer, A.J.; van der Woude, L.H.V.

published in
Clinical Rehabilitation
2003

DOI (link to publisher)
10.1191/0269215503cr633oa

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

Link to publication in VU Research Portal

citation for published version (APA)

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 27. Sep. 2023
Wheelchair skills tests: a systematic review

Olga JE Kilkens, Marcel WM Post iRv, Institute for Rehabilitation Research, Hoensbroek, Annet J Dallmeijer Institute for Fundamental and Clinical Human Movement Sciences, Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, Henk AM Seelen iRv, Institute for Rehabilitation Research, Hoensbroek and Rehabilitation Foundation Limburg, Hoensbroek and Lucas HV van der Woude Institute for Fundamental and Clinical Human Movement Sciences, Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

Received 11th May 2002; returned for revisions 11th November 2002; revised manuscript accepted 14th December 2002.

Objective: To describe and compare the content, feasibility, outcome parameters, and clinimetric properties of the manual wheelchair skills tests reported in the literature.

Design: A systematic literature search was conducted in MEDLINE, EMBASE, PsychINFO and Current Contents. Tests were selected if they were observational tests, designed for subjects using hand-rim wheelchairs and were intended to assess wheelchair skill performance at the activity level.

Results: The search resulted in 34 papers, in which 24 different wheelchair skills tests were described. The skill most frequently included was wheelchair propulsion, consecutively followed by transferring, negotiating kerbs, ascending slopes, traversing tracks, sprinting and performing a wheelie. The three most frequently used outcome parameters were task performance time, independency of task performance, and physical strain during skill performance. Sensitivity to change was evaluated in three tests, validity in 10 tests, and reliability in nine tests.

Conclusions: Many tests are applied to measure wheelchair skill performance using different tasks and outcome measures. This makes it difficult to compare study results. Consensus among researchers as to which skills must be included as well as to standardization of the use of measurement instruments will reduce this problem and will additionally lead to a better insight in the quality of tests.

Introduction

The achievement of independent mobility is vital in the rehabilitation of physically disabled individuals. When ambulation is impaired, a hand-rim wheelchair provides a relatively fast and effective means of mobility for people with lower limb disabilities. A hand-rim wheelchair can provide the necessary access to social, vocational and recreational activities that are conditional to a productive and rewarding life. To function independently, people who use manual wheelchairs for mobility must possess a variety of skills. The ability to propel their wheelchairs over even surfaces brings the freedom to move about within a wheelchair-accessible environment. Independent
mobility within a greater variety of environments requires obstacle negotiation skills. These skills can make the difference between dependence and independence in daily life.  

Assessment of wheelchair skills can provide useful information concerning a person's current wheelchair skill performance. In clinical situations, wheelchair skills tests can help to define rehabilitation goals concerning mobility, and can also be used to evaluate the progression made regarding wheelchair mobility during rehabilitation. In research settings, measurement of wheelchair skills can be used to study the effect of an intervention aimed at wheelchair mobility or to study the relation between wheelchair skills and, for example, level of activity and/or participation.

At present there is no systematic overview of wheelchair skills tests available in the literature. It is therefore difficult to decide which test is most suitable in research or in clinical practice.

The objective of this review is to systematically document and describe the content, the target population, the study group, the test feasibility, the outcome parameters and the clinimetric properties of those hand-rim wheelchair skills tests that are currently reported in the literature. Such an overview may make it easier to choose the most suitable test to assess wheelchair skills in both clinical and research settings.

Methods

Search strategy

To locate wheelchair skills tests, a computerized literature search of MEDLINE (1966–2001), EMBASE (1989–2001), PsychINFO (1967–2001) and Current Contents (1998–2001) was conducted. The keywords used were: mobility and wheelchair combined with skill, task, measurement, test, ADL, functional, instrument, performance, clinimetrics, psychometrics, pathology, behaviour, activity, disability and assessment. The search strategy is described in the Appendix. In addition, the references given in relevant publications were further examined. Only studies written in English that were published in scientific journals were taken into consideration.

Selection criteria

A test was selected if it was an observational test, if it was constructed for subjects using hand-rim wheelchairs and when it intended to assess wheelchair-assisted mobility skills at the activity level as described in terms of the International Classification of Functioning, Disability and Health (ICF). In the ICF, mobility is defined as: 'moving by changing body position or location or by transferring from one place to another'. Consequently, this review focuses on tests that aim to assess the ability to propel and manoeuvre a wheelchair under standardized and/or simulated conditions of daily living. Tests aimed at measuring physical capacity were not selected.

The first author performed the selection of the tests, by reading the abstracts of all the initially identified articles. When necessary the full article was obtained and studied. In case of doubt on selection of a test, the other authors were consulted.

Assessment of selected tests

The wheelchair skills tests were systematically described and compared with respect to the following aspects:

- Content: the skills included in the test.
- Target population: the diagnostic groups for which the test was developed.
- Population at study: the diagnostic groups in which the test was used or studied.
- Feasibility: the amount of time and equipment needed to perform the test.
- Test outcomes: the outcome parameters used to reflect wheelchair skill performance and the complexity and interpretation of the scoring method.
- Clinimetric properties: sensitivity to change, validity and reliability of the test.

Results

Selection of tests

The selection process produced 34 papers in which 24 different wheelchair skills tests were described. Table 1 provides an overview of the selected tests, arranged alphabetically, according to the name of the first author of the paper in which the test was mentioned. Of the 24 tests
Table 1  General overview of the selected wheelchair skills tests

<table>
<thead>
<tr>
<th>Author</th>
<th>Name of test</th>
<th>N</th>
<th>Target population</th>
<th>Study population</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agre⁵</td>
<td></td>
<td>33</td>
<td>No information</td>
<td>Patients with spina bifida</td>
<td>Wheelchair propulsion velocity</td>
</tr>
<tr>
<td>Bolin⁶</td>
<td></td>
<td>4</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Task performance times</td>
</tr>
<tr>
<td>Capodaglio⁷</td>
<td></td>
<td>8</td>
<td>Patients with SCI</td>
<td>Patients with SCI</td>
<td>Physical strain (HRpeak)</td>
</tr>
<tr>
<td>Taricco³,⁴</td>
<td></td>
<td>47, 94</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Subjective rating of performance</td>
</tr>
<tr>
<td>VFM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Independence of task performance</td>
</tr>
<tr>
<td>Dallmeijer⁸,⁹</td>
<td></td>
<td>20, 19</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Task performance times</td>
</tr>
<tr>
<td>Duffill¹⁰</td>
<td></td>
<td>17</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Physical strain (%HR)</td>
</tr>
<tr>
<td>Dunkerley¹¹</td>
<td></td>
<td>11</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Ability to perform tasks</td>
</tr>
<tr>
<td>Durán¹²</td>
<td></td>
<td>13</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Task performance times</td>
</tr>
<tr>
<td>Findley¹³</td>
<td></td>
<td>40</td>
<td>No information</td>
<td>Patients with spina bifida, healthy individuals</td>
<td>Ability to perform tasks</td>
</tr>
<tr>
<td>Gans¹⁴</td>
<td></td>
<td>40</td>
<td>No information</td>
<td>Patients with spina bifida, healthy individuals</td>
<td>Wheelchair propulsion velocity</td>
</tr>
<tr>
<td>Haley¹⁵,¹⁷</td>
<td></td>
<td>206, 206</td>
<td>No information</td>
<td>Patients with musculoskeletal and neuromuscular disorders</td>
<td>Task performance times</td>
</tr>
<tr>
<td>TAM⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Independence of task performance</td>
</tr>
<tr>
<td>Gouvier¹⁵</td>
<td></td>
<td>2</td>
<td>Patients with stroke</td>
<td>Patients with stroke, healthy individuals</td>
<td>Technique</td>
</tr>
<tr>
<td>Webster³⁶-³⁸</td>
<td>WOC</td>
<td>72, 87, 55</td>
<td>No information</td>
<td></td>
<td>Movement pattern</td>
</tr>
<tr>
<td>Harvey¹⁸</td>
<td></td>
<td>20</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Movement control and co-ordination</td>
</tr>
<tr>
<td>Hutzler¹⁹</td>
<td></td>
<td>9</td>
<td>Wheelchair athletes</td>
<td>Patients with SCI</td>
<td>Number of collision errors</td>
</tr>
<tr>
<td>Janssen²⁰</td>
<td></td>
<td>37</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Independence of task performance</td>
</tr>
<tr>
<td>Janssen²¹</td>
<td></td>
<td>44</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Task performance times</td>
</tr>
<tr>
<td>Janssen²²</td>
<td></td>
<td>37</td>
<td>No information</td>
<td>Patients with SCI</td>
<td>Distance covered</td>
</tr>
<tr>
<td>Jebsen²³</td>
<td></td>
<td>118</td>
<td>No information</td>
<td>Patients with stroke, amputations, SCI neuropathy, hip fractures and healthy individuals</td>
<td>Physical strain (HRpeak, HRmean)</td>
</tr>
</tbody>
</table>

Tasks included: Wheelchair propulsion velocity, Task performance times, Physical strain (HRpeak, HRmean), Subjective rating of performance, Independence of task performance, Task performance times, Physical strain (%HR), Task performance times.
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>Group Description</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kirby²⁴</td>
<td>42</td>
<td>No information</td>
<td>Patients with SCI, amputations and healthy individuals</td>
</tr>
<tr>
<td>Kirby²⁶</td>
<td>24</td>
<td>No information</td>
<td>Patients with stroke, amputations, SCI, musculoskeletal and neuromuscular disorders</td>
</tr>
<tr>
<td>WST</td>
<td></td>
<td>No information</td>
<td>Patients with SCI</td>
</tr>
<tr>
<td>Lehmann²⁶</td>
<td>12</td>
<td>No information</td>
<td>Patients with SCI</td>
</tr>
<tr>
<td>Mattison²⁷</td>
<td>26</td>
<td>No information</td>
<td>Patients with stroke, vascular diseases and healthy individuals</td>
</tr>
<tr>
<td>Mizukami²⁸</td>
<td>109</td>
<td>No information</td>
<td>Patients with SCI</td>
</tr>
<tr>
<td>Panikoff²⁹</td>
<td>80</td>
<td>No information</td>
<td>Patients with head injury</td>
</tr>
<tr>
<td>Schnelle³⁰,³¹</td>
<td>97, 76</td>
<td>No information</td>
<td>Nursing home residents</td>
</tr>
<tr>
<td>Simmons³²</td>
<td>65</td>
<td>No information</td>
<td>轮椅篮球运动员有脊髓损伤、SCI、痉挛性脑瘫、中风和健康个体</td>
</tr>
<tr>
<td>Vanlandewijk³⁵</td>
<td>46</td>
<td>Wheelchair athletes</td>
<td>Wheelchair basketball players with spina bifida, SCI, spastic diplegia, polio, amputations and healthy individuals</td>
</tr>
</tbody>
</table>

SCI, Spinal cord injury; VFM, Valutazione Funzionale Mielolesie; TAMP, Tufts Assessment of Motor Performance; WOC, Wheelchair Obstacle Course; WST, Wheelchair Skills Test; HRpeak, peak heart rate reached during task performance; %HRR, percentage heart rate reserve; HRmean, mean heart rate reached during task performance.
found, seven were presented as measurement instruments and were extensively described in terms of development, content and use.\textsuperscript{7,14,18,23,25,35,36} In all other papers the aim was to evaluate an intervention or to detect differences between groups. These tests were only briefly described in the Methods section of the article. Only four tests had been given a name: the Valutazione Funzionale Mielolesie (VFM),\textsuperscript{7} Tufts Assessment of Motor Performance (TAMP),\textsuperscript{14} the Wheelchair Skills Test (WST)\textsuperscript{25} and the Wheelchair Obstacle Course (WOC).\textsuperscript{36}

**Assessment of selected tests**

**Content of tests**

Table 2 displays the types of wheelchair skills included in the different tests. Wheelchair propulsion is the most frequently included skill (in 14 tests). It is assessed in different ways: a set period of time,\textsuperscript{6,19} a fixed distance,\textsuperscript{5,6,14,18,25,36} or the longest distance possible.\textsuperscript{30,35} Following wheelchair propulsion, transfer from and to the wheelchair is the most commonly included skill (in 11 tests). Most tests require the performance of several different transfers.\textsuperscript{7,14,18,20-23,28,29} The negotiation of kerbs, ascending slopes and traversing tracks are third in line of most frequently used skills (each in 10 tests). The height of the kerbs used ranged from 0.025 to 0.15 m. Two tests\textsuperscript{7,12} require both ascending and descending of the kerb. All other tests only assess the ascending of the kerb. In all but three tests,\textsuperscript{7,10,14} the slopes used are defined in terms of inclination and length, inclinations ranging from 1 to 11 degrees, length ranging from 3.05 to 21 m. Some examples of tracks used are: slalom,\textsuperscript{6,19,25} figure of eight\textsuperscript{11} and obstacle course.\textsuperscript{12,36} In six tests a sprint is included. Nearly all tests use a sprint over a fixed distance (length ranging from 6.5 to 30 m). Although performing a wheelie is an important skill in achieving wheelchair mobility, this skill is only included in four tests.

Eleven tests include, in addition to the skills already mentioned, other specific wheelchair skills, e.g., managing brakes, negotiating doors and loading the wheelchair into a car. Fifteen tests consist entirely of the performance of wheelchair skills. In eight tests wheelchair skills are a part of a broader measure of ADL skills; these tests encompass other ADL tasks such as eating, bed mobility skills and washing hands.\textsuperscript{7,8,14,18,21,23,28,29}

**Target population and population at study**

Although only four tests were designed for a specific target population, 16 tests have only been used in study groups with one specific diagnosis, most often spinal cord injury. Four tests were used for subjects with varying medical conditions (Table 1).

**Feasibility**

On the one hand, tests should include enough elements to obtain an in-depth insight into wheelchair skill performance; on the other hand, tests have to be efficient and as short as possible. The completion time was mentioned for only six tests. The VFM,\textsuperscript{7} the TAMP\textsuperscript{14} and the test of Jebsen \textit{et al.}\textsuperscript{23} take up to 1 hour to complete. However, these tests contain other ADL tasks as well as specific wheelchair skills. The performance of Harvey’s test\textsuperscript{18} requires approximately 15 minutes, the time needed to complete the WST\textsuperscript{25} is 30 minutes, and the mean test duration of the wheelchair basketball field test of Vanlandewijck \textit{et al.}\textsuperscript{35} is 1 hour and 22 minutes.

Ideally, tests should not require much space or special equipment. In most studies, the materials needed for test performance are not addressed. In their paper, Jebsen \textit{et al.}\textsuperscript{23} dedicated a section to test equipment (a hospital bed, standardized wheelchair and straight chair). Harvey \textit{et al.}\textsuperscript{18} stated that no special equipment is required to perform their test. To assess physical strain during wheelchair skill performance a heart rate monitor is required. Twelve studies provide information on the wheelchairs used during test performance.\textsuperscript{6,8,11-13,20-26} Three studies used standardized wheelchairs.\textsuperscript{23,24,26} In eight studies, subjects used their daily use wheelchairs.\textsuperscript{8,11-13,20-22,25} Bolin \textit{et al.}\textsuperscript{6} aimed to improve the individual fit of the wheelchair in their subjects. The subjects performed a wheelchair skills test twice: first in their daily use wheelchair and later in an adapted or new wheelchair.

The outcome measures of the different tests are displayed in Table 1. The most common outcome measure is task performance time. Independence in wheelchair skill performance is
assessed by taking into account the use of assistive devices or the amount of help needed from another person. Six tests measure physical strain during skill performance, four tests evaluate distance covered during wheelchair propulsion, four tests rate the velocity of wheelchair propulsion, and three tests assess subjective ratings regarding skill performance.

A test should preferably have an uncomplicated scoring system that is convenient to use and that can be analysed easily. The scoring of the TAMP14 is extremely complex: one hundred and thirteen skills have to be rated on six measurement dimensions and rating this test requires extensive training.

**Clinimetric properties**

Nine tests6,10,11,13,22,24,26,28,30 were not evaluated on any of the clinimetric properties. Three tests provide information on sensitivity to change.7,8,12 Only two tests, the VFM7 and the WST,25 were explicitly subjected to a validation study. For eight other tests5,12,19,21,24,26,28,30,36 information on validity could be retrieved from the articles. The validity of these tests, however, was not explicitly evaluated. Information concerning reliability was given for 10 tests.7,14,15,18-20,23,25,27,35 Only five tests19,23,25,27,36 provided data on both reliability and validity. Table 3 displays the 11 tests from which the sensitivity to change and/or the validity have been evaluated. In Table 4, the 10 tests that have been assessed on the topic of reliability are shown.

From Table 4 it can be seen that all available test–retest and inter-rater reliability figures are satisfactory up to excellent; the data on validity are less unequivocal (Table 3).

**Discussion**

A literature search resulted in the selection of 24 different wheelchair skills tests. This collection may be incomplete, since only English-written studies, published in scientific journals were taken into account. However, we feel that we have provided a critical and useful overview of tests in which wheelchair skills are assessed.
<table>
<thead>
<tr>
<th>Author</th>
<th>Sensitivity to change</th>
<th>Content validity</th>
<th>Construct validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agre(^5)</td>
<td>No information</td>
<td>No information</td>
<td>Failed to find a relationship between motor function and propulsion speed</td>
</tr>
<tr>
<td>Capodaglio(^9) Taricco(^{34})</td>
<td>Two groups were tested before and after a 6-week rehabilitation programme: conventional or enhanced. No significant differences between the groups(^6) Subjects were tested at the start and the end of their rehabilitation period. Scores had significantly improved for subjects with quadriplegia and high-level paraplegia. There were no changes in the scores of subjects with low-level paraplegia(^{13})</td>
<td>The content of the VFM was compared to the content of the FIM, Barthel index and QIF. The VFM includes all basic ADL domains and has the largest number of tasks in the domains transfers and wheelchair use which are particularly relevant for wheelchair-dependent people(^{23})</td>
<td>No information</td>
</tr>
<tr>
<td>Dallmeije(^{6,9})</td>
<td>Subjects were tested at the start and at the end of their rehabilitation period. There were no changes in physical strain during the slope ascend and the transfers. The performance time of the slope ascend significantly decreased. The performance time of the transfer did not change(^{6})</td>
<td>No information</td>
<td>No information</td>
</tr>
<tr>
<td>Durán(^{12})</td>
<td>Subjects were tested before and after a 16-week exercise programme. Task performance times significantly decreased. The ability to negotiate kerbs improved</td>
<td>No information</td>
<td>Mean FIM scores were associated with an improved ability to negotiate kerbs, and inversely associated with task performance times</td>
</tr>
<tr>
<td>Gouvier(^{15}) Webster(^{36-38})</td>
<td>No information</td>
<td>No information</td>
<td>Right-sided stroke patients with neglect hit significantly more objects in left- than in right-space, and made significantly more errors than stroke patients without neglect and healthy control subjects(^{20,37}) Inverse correlation between sprint performance time and the subjects' wheelchair basketball classification (r(<em>{\text{rs}}) = -0.64). No relation between body weight and task performance times A lower lesion level was associated with a lower level of physical strain during wheelchair skill performance. No relation between lesion-completeness and physical strain during skill performance. Physical strain during skill performance was inversely related to strength, sprint power, peak oxygen uptake, and maximum power output (r(</em>{\text{rs}}) = -0.41 to -0.73)</td>
</tr>
<tr>
<td>Hutzler(^{19})</td>
<td>No information</td>
<td>No information</td>
<td></td>
</tr>
<tr>
<td>Janssen(^{21})</td>
<td>No information</td>
<td>No information</td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>No information</td>
<td>No information</td>
<td>No information</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Jebsen</td>
<td>No information</td>
<td>No information</td>
<td>No information</td>
</tr>
<tr>
<td>Kirby</td>
<td>No information</td>
<td>Occupational therapists evaluated the content of the WST, and unanimously approved with 30 of the 33 skills</td>
<td></td>
</tr>
<tr>
<td>Mattison</td>
<td>No information</td>
<td>No information</td>
<td>No information</td>
</tr>
<tr>
<td>Panikoff</td>
<td>No information</td>
<td>No information</td>
<td>No information</td>
</tr>
</tbody>
</table>

Healthy people performed the tasks concerning wheelchair mobility faster than patients with mobility problems. Occupational therapists rated whether patients' wheelchair skills had improved, not changed or worsened between two test trials. They rated that 13 subjects had improved, 8 had not changed, and none had worsened. Accordingly, the mean improvement in the total score was larger for the 'improved subjects' than for the 'unchanged subjects'. OT's also completed a VAS reflecting the subjects' wheelchair skill performance. Mean improvement for the 'improved subjects' was significantly higher than for the 'unchanged subjects'. Test scores significantly related to age, and wheelchair experience. No relation was found between test scores and diagnosis. There was no relation between physical strain during skill performance and age or diagnosis. Physical strain was significantly related to the perceived exertion during test performance ($r_{Pearson}=0.84$). A significant relation between the ability to perform transfers and length of coma.

FIM, Functional Independence Measure; QIF, Quadriplegia Index of Function.
There is limited consensus as to the content of wheelchair skills tests (Table 2). Even skills frequently used in tests (wheelchair propulsion, transfer, kerb, slope) show a large variation in, for instance, driving distance, objects to transfer to, height of the kerbs and angle of inclination of slopes. The number of skills included in wheelchair skills tests also shows a large variation, ranging from one\(^3\) to 113.\(^{14}\) More research is needed to identify a limited number of skills that together best reflect wheelchair skill performance in people who depend on a manual wheelchair for their mobility.

Although just four tests were designed for a specific target group, 16 tests were only used in a specific group (Table 1). The latter tests might also be capable of assessing wheelchair skills in subjects with other characteristics, but further research on validation and reliability in other subject groups is necessary to test this expectation.

Many different outcome parameters are used in the selected tests (Table 1). The choice for a particular outcome measure depends on the objectives of the study. Tests can be used to determine the feasibility of manual wheelchair propulsion, to measure the level of independence in wheelchair ADL, or to evaluate the effects of interventions. Outcome measurements in the categories time, distance and physical strain are very useful to provide information on the practicability of manual wheelchair mobility in daily life. When a subject needs, for example, five minutes to propel his or her wheelchair over a distance of 50 metres, an electric wheelchair may be a more suitable means of mobility. If the goal of a study is to describe the level of independent mobility, a scale of independence in performing wheelchair tasks is an obvious outcome measure.

The level of independence in performing well-chosen wheelchair skills is expected to be directly related to independent mobility in daily life. People who cannot perform wheelchair skills independently will not achieve independent mobility in all environmental circumstances. For the assessment of (changes in) wheelchair skill performance in completely independent individuals, outcome measures such as time, distance and physical strain should be applied. This is also shown by the results of Taricco et al.\(^3\) who measured wheelchair skill performance using a scale of independence. They showed good sensitivity to change in subjects with high-level spinal cord injury, but no sensitivity to change in subjects with low-level spinal cord injury.

Other outcome measures are relevant, but not so easy to interpret. The test of Dallmeijer et al.\(^8\)\(^,\)\(^9\) evaluates both physical strain during wheelchair skill performance and performance time of each skill. These two parameters are, however, interdependent. A decrease in performance time, reflecting better test performance, may result in a higher level of physical strain, indicating worse test performance. This interdependency may obfuscate the interpretation of test results.

Further, wheelchair skill performance relies on both technique and physical capacity. Repeated measurements can, for instance, show that maximal wheeling endurance time has increased over a certain period, which may be the result of an increase in physical capacity, an improved technique resulting in higher mechanical efficiency of wheelchair propulsion, or a result of both. For a correct interpretation of changed outcomes in longitudinal studies, the performance of a wheelchair skills test is best combined with an exercise test that provides information about physical capacity.

The WST\(^{25}\) leads to one overall score of wheelchair skill performance, expressed as the sum of the scores obtained on each skill. Such a total score might be very useful for research purposes,
<table>
<thead>
<tr>
<th>Author</th>
<th>Test-retest reliability</th>
<th>Inter-rater reliability</th>
<th>Intra-rater reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capodaglio/1</td>
<td>No information</td>
<td>The inter-rater reliability of the VFM was demonstrated in previous studies. These studies were only published in Italian and not retrieved.</td>
<td>No information</td>
</tr>
<tr>
<td>Tanicco/33, 34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gans/14</td>
<td>No information</td>
<td>Subjects' test performances were videotaped, and independently evaluated by three raters. Regarding wheelchair skills, inter-rater reliability was excellent (kappa = 0.65–0.83)</td>
<td>No information</td>
</tr>
<tr>
<td>Haley/16, 17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gouvier/16</td>
<td>No information</td>
<td>Two raters scored the number and type of errors two subjects made during test performance. Their agreement on the occurrence of errors and the type of error made was successively 95% and 83%</td>
<td>No information</td>
</tr>
<tr>
<td>Webster/18, 19, 20</td>
<td></td>
<td>Two raters scored the tests. They agreed 100% on the occurrence of errors, 80–90% on the occurrence of direct hits (r_{spearman}=0.90–0.97), and 85–90% on the occurrence of sideways errors (r_{spearman}=0.92–0.95)</td>
<td>No information</td>
</tr>
<tr>
<td>Harvey/18</td>
<td>No information</td>
<td>Two raters scored the test performances resulting in high inter-rater reliability (kappa=0.82–0.98)</td>
<td>No information</td>
</tr>
<tr>
<td>Hutzler/19</td>
<td>Excellent test-retest reliability for task performance time and propelled distance (r_{spearman}=0.87–0.99) on two test performances</td>
<td>No information</td>
<td></td>
</tr>
<tr>
<td>Janssen/20</td>
<td>Subjects performed three test trials: two on the same day (trial 1 and 2) and one a week later (trial 3). Correlations for heart rate were good for trial 1 versus 2, and trial 1 versus 3 (1 vs 2: r_{spearman}=0.79–0.97, ICC=0.73–0.97, 1 vs 3: r_{spearman}=0.69–0.95, ICC=0.52–0.92). Good correlations for the performance time of the kerb ascend for trial 1 versus 2 (r_{spearman}=0.68, ICC=0.92). Low correlations for trial 1 versus 3 (r_{spearman}=0.31, ICC=0.20)</td>
<td>No information</td>
<td></td>
</tr>
<tr>
<td>Jebsen/23</td>
<td>Performance times correlations were excellent (r_{spearman}=0.85–0.99)</td>
<td>No information</td>
<td></td>
</tr>
<tr>
<td>Kirby/25</td>
<td>Good correlation between the scores on two different test occasions (r_{spearman}=0.65)</td>
<td>Two raters independently scored the same videotapes of test performances, resulting in an excellent correlation coefficient (r_{spearman}=0.95)</td>
<td>One rater scored the same videotapes of subjects test performances twice, resulting in an excellent correlation coefficient (r_{spearman}=0.96)</td>
</tr>
<tr>
<td>Mattison/27</td>
<td>Subjects performed two test trials. Good correlations for distance traveled, physical strain during wheelchair propulsion, and propulsion velocity (r_{spearman} = 0.96, 0.84, and 0.70)</td>
<td>No information</td>
<td>No information</td>
</tr>
<tr>
<td>Vanlandewijck/26</td>
<td>Excellent correlations for all tasks (r_{spearman}=0.80–0.97), only one being good (r_{spearman}=0.65)</td>
<td>No information</td>
<td>No information</td>
</tr>
</tbody>
</table>
but can only lead to valid information if all test items measure the same phenomenon. The authors did not assess this.

Clinimetric properties

Clinimetric properties of nine tests were not described at all. Only two tests: the VFM\textsuperscript{7} and the WST\textsuperscript{25} have been extensively validated. The WST is the only test that has been adequately tested on both validity and reliability. More research is needed to assess the clinimetric qualities of the other tests described in the current review before these tests can be recommended for use. Performance time and physical strain are outcome measures that can be measured objectively. Ordinal scales of dependence, frequently used in wheelchair skills tests, are subject to interpretation. Raters need to assess the amount of help, often expressed in a few number of categories of assistance needed. Therefore, the raters should receive appropriate training. The good inter-rater reliability figures of tests using ratings of independence are promising, but these figures are, in part, obtained in very small study groups. Although also rarely investigated, test-retest analyses of time, distance, velocity and physical strain revealed satisfying results.

The measurement of wheelchair skills will, at least in part, support validity of the tests due to the close resemblance with daily life activities. However the choice of tested tasks, outcome measures and the applicability in different subject groups may influence validity. One aspect of validity that is often ignored is the influence of wheelchair configuration on wheelchair skill performance outcomes. Subjects will perform best in a wheelchair that is optimally adjusted to their personal characteristics. To ensure that variations in wheelchair skill performance were not due to changes in wheelchair configuration, some tests were performed in standardized wheelchairs.\textsuperscript{23,24,26} However, most tests were executed in daily use wheelchairs.\textsuperscript{8,12,20,22,25} This may have resulted in subjects using different wheelchairs on different test occasions, which may have affected sensitivity to change and test-retest reliability or may bias comparisons between subjects having wheelchairs of different quality. Use of the daily use wheelchair may, however, improve the validity of the test. Subjects are not troubled by an unfamiliar wheelchair, and their test results will be more representative for their wheelchair skill performance in daily life. Therefore a careful choice for, or against standardization of wheelchair configuration has to be made, dependent on the purpose and the design of the study.

In conclusion, this review shows that there is, as yet, no standard test to measure wheelchair skill performance. Only seven out of the 24 tests found were extensively described in terms of development, content and use\textsuperscript{7,14,18,23,25,35,36} and only two tests have been extensively validated.\textsuperscript{7,25}

In addition, most tests have only been used in one study. Without further research on validity and reliability, these tests should be used with caution. The use of many different tests makes it difficult, if not impossible, to compare study results. Standardization of the skills tested and the use of measurement instruments are needed to enable comparisons between studies and to give a better insight in the quality of the tests used.

Future research could best concentrate on further validation of existing tests instead of developing even more tests. The selection of the best and most relevant items of these tests and combining elements of various tests might eventually lead to a superior test. However, it might not be possible to compose the ideal test for all patient groups and purposes. A distinction between a clinical instrument (containing all relevant items for assessment and evaluation of individual treatment) and a research instrument (containing a selection of items of varying difficulty) might be useful.

Acknowledgements

The study was supported by the Dutch Health Research and Development Council, ZON-MW, Rehabilitation program, grant no. 1435.0003.

References

3 Bussmann JBJ, Stam HJ. Techniques for measurement and assessment of mobility in rehabilitation: a
Wheelchair skills tests

429

33 Taricco M, Colombo C, Adone R, Chiesa G, Di


Appendix – Search strategy

#1 mobility and wheelchair
#2 #1 and skill*
#3 #1 and task*
#4 #1 and measurement*
#5 #1 and test*
#6 #1 and ADL
#7 #1 and functional*
#8 #1 and instrument*
#9 #1 and performance
#10 #1 and clinimetrics
#11 #1 and psychometrics
#12 #1 and pathology
#13 #1 and behaviour
#14 #1 and activit*
#15 #1 and disabilit*
#16 #1 and assessment
#17 #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16