Reliability of the Test of Wheeled Mobility (TOWM) and the Short Wheelie Test

Osnat Fliess-Douer [a], Lucas H.V. Van der Woude [b], Yves C. Vanlandewijck [a]

[a] Faculty of Kinesiology and Rehabilitation Sciences, Department of Rehabilitation Sciences, Catholic University of Leuven, Leuven, Belgium

[b] Center for Human Movement Sciences, University Medical Center Groningen, University of Groningen, Groningen,The Netherlands.

Published version: http://dx.doi.org/10.1016/j.apmr.2012.09.023
Journal homepage: http://www.archives-pmr.org/

Abstract

Objective: To assess the reliability of the Test of Wheeled Mobility (TOWM) and the Wheelie test.

Design: Cohort study**.**

Setting: Gymnasium at the Katholieke Universiteit Leuven.

Participants: Thirty manual wheelchair-users (Age 23-53 years) with a spinal cord injury.

Intervention: Participants performed the 30 skills of the TOWM and 8 skills of the Wheelie test twice. Ability, time and anxiety scores were assessed on-field. Quality scores were assessed by video analysis.

Main outcome measure: Test-retest reliability was evaluated for the ability, time, anxiety and quality scores of both tests. Intrarater and interrater reliability were determined based on the quality scores of 20 participants. Intraclass coefficient and non-parametric statistics were applied, as well as standard error of measurement (SEM), method error (ME), coefficient variation of method error (CV_{ME}), minimal detectable change (MDC_{95}), and technical error of measurement (TEM).

Results: Test-retest reliability: no significant differences between t1 and t2 in the ability, quality and time scores, except to the anxiety scores. SEM, ME CV_{ME} and MDC_{95} values were low for the ability and quality total score and higher for the time and anxiety total score. Intrarater and interrater reliability ICC's of both tests ranged between 0.91-0.99. Interrater relative TEM for the TOWM and the Wheelie test total quality score was 3.7% and 6.3% respectively and intrarater relative TEM was 4.3% and 6.1%. ICC's per individual tasks, ranged between 0.88-1.00, except for 'level propulsion forward' that showed low ICC scores (Interrater: 0.49; intrarater: 0.44; test-retest: 0.43).

Conclusions: Based on ability and quality total scores, the TOWM and the Wheelie test are reliable when assessing wheeled mobility of manual wheelchair users with SCI. The quality criteria of one task from the TOWM and three tasks from the Wheelie test need to be refined.

Key words:

Wheeled mobility, spinal cord injury, wheelchair skills, reliability

A list of abbreviations:

TOWM – *Test of Wheeled Mobility,*

WM – *Wheeled mobility;*

SCI – *Spinal cord injury,*

VAS – *visual analog scale*

Wheeled mobility (WM) is a key to independence for a large group of people with a handicap, especially for those who will not return to walking again, such as individuals with a complete spinal cord injury (SCI).[1] Learning wheelchair skills performance is an important part of the rehabilitation process, since for many people with SCI, the wheelchair will serve as the main device for mobility.

Wheeled mobility (WM) is defined by the ICF as "Moving around using equipment: moving the whole body from place to place, on any surface or space, by using specific devices designed to facilitate moving or create other ways of moving around, such as moving down the street in a wheelchair or a walker". [2]

Participation is also an important rehabilitation outcome for persons with SCI. In the International Classification of Functioning, Disability and Health (ICF)[2] participation is defined as 'involvement in life situations' including, for example, work and school, social relationships and community organizations. Participation restrictions are the problems that an individual may have in involvement in life situations. From the literature it is known that persons with activity limitations experience participation restrictions in daily life and there is a positive relationship between WM skills and participation in persons with SCI.[3]

Therapists should have a valid, reliable and sensitive measuring tool at their disposal, in order to objectively and systematically assess their patient's level of WM performances, before, during and after interventions. Currently, several wheelchair skills tests are available, based on actual performance.[4-16]

In a recently published systematic literature review on wheelchair skill tests[17], results showed that only few tests focus explicitly on WM in persons with a SCI [8, 11, 13, 14]. Wheelchair skill tests that were aiming to the general wheelchair user's populations failed to differentiate between levels of performance and resulted in a "ceiling effect", mainly in individuals with paraplegia[17]. The review study revealed a lack of a broadly accepted wheelchair skills test, and disclosed large inconsistencies among the current available tests, which made comparison of study results impossible[17]. Furthermore, different scales were applied to express test scores. Some tests used qualitative scales, while others used quantitative scales.

A quantitative measurement can be, for example, the time necessary to complete a task or the percentage of a slope. The qualitative measurement has to be well defined. It is not sufficient to apply only a pass/fail scale, because the same level of difficulty can be completed with a different grade of maturity. In this case, it can be useful to incorporate the term "with difficulties", or to combine also a performance time assessment (quantitative and qualitative measurements within the same test). Still, these scales might not be sensitive enough to detect small changes in WM because they only record whether a person can perform a particular task or whether the task is performed independently. Small improvements in the quality of the skill performance can often not be scored. Nonetheless, these small changes can be of great importance. The way to overcome this problem would be through developing quality assessment criteria for each WM skill. These criteria (based on key components which compose the skill) will reflect the WM maturity proficiency.

Within the scope of a study aimed to promote a standardized broadly accepted and applicable WM test, the Test of Wheeled Mobility (TOWM) and the short Wheelie tests were developed.[18] The development of these two tests was based on experts' opinions and on the outcome of an international survey among users, aimed to create a sorted list of the most essential WM skills. [19] In addition, a systematic critical literature review of available wheeled mobility skill tests enabled the development of the new tests, relying on strengths of existing tools.[17] The purpose of the TOWM and Wheelie test is to assess WM skills in manual wheelchair users with SCI during and after clinical rehabilitation, allowing accurate monitoring and assessment of small changes in WM. The TOWM and Wheelie test are primarily designed for clinical purposes but may also be used in a research setting. Both assessment tools were tested for their feasibility and validity with respect to duration, costs, content, construct, convergent and predictive validity[18]. The protocols of the TOWM and the Wheelie test with description of the tasks as well as testing equipment and score sheets can be obtained at: www.scionn.nl/inhoudp28.htm.

The reliability of any new measurement instrument is critical to ensure that the measurement error is small enough to detect actual changes in what is being measured[20]. The nature of reality is such that due to instrument imprecision and human inconsistencies, measurements are not free of error (i.e. perfectly reliable). The aim of this study was to assess the reliability and response stability of the TOWM and the Wheelie test. Responsiveness refers to the ability of a measure to detect clinically meaningful change over time, and provides a means for determining if an individual's change in score is related to true recovery, or to natural variation in repeated performances. Scale responsiveness is an important concept for clinicians in this time of evidence-based practice, and understanding and interpreting the responsiveness of a scale enables clinicians to discriminate true change from measurement error[21].

Therefore, the objectives of this study were: (1) to determine the test–retest reliability and response stability of the TOWM and the Wheelie test based on the four scales total scores; (2) to assess the inter- and intrarater reliability of both tests' tasks, based on the agreement between and within raters, as well as on technical error of measurement index (accuracy assessment) ; and (3) To determine the minimum detectable change at 95% confidence for both tests in a group of manual wheelchair users with SCI.

Methods

Participants

Thirty wheelchair users with SCI (convenience sample) were recruited during the first 2 weeks of February 2011. Recruitment was performed by word of mouth, email and telephone calls. All participants were living in Belgium, were post clinical rehabilitation, used a hand-rim wheelchair, and were between 18 and 65 years of age. Potential participants were not included if they had a current cardiorespiratory disorder or orthopedic or other medical complications that restricted them in performing the tasks required for the TOWM and the Wheelie test. One participant did not attend the retest (t2) due to a recurrence of an old shoulder injury unrelated to the WM-testing; therefore, the results of 29 participants were included in the presented data analysis.

All procedures were performed in accordance with the guidelines of the Declaration of Helsinki. The study was approved by the medical ethics committee of the Catholic University of Leuven, Belgium. Prior to participation, all participants signed an informed consent. Participants were reimbursed for transportation costs.

Measuring instruments

The TOWM consists of 30 standardized tasks which are conditional to mobility in persons with SCI.[18] The short Wheelie test includes 8 tasks which are related to the ability to perform a mature pattern of a wheelie (balancing on the rear wheels). The TOWM and the Wheelie test tasks present different difficulty levels and are applied from the easiest to the most difficult. Including preparation and evaluation time, the estimated duration of both tests together is 40 minutes (Mean ± SD for the TOWM 24.7±5.93 min., and for the Wheelie test 12.62±5.08 min.), including equipment preparations, instructions for each task and performance of the test.[18]

The TOWM and the Wheelie Test have 4 scoring methods: The *ability score* refers to all the items that can be performed sufficiently and independently, and is being assigned as 1 point if the participant completes the task successfully in the first trial, 0.5 point if he succeeds on the second trial, and 0 score presents either a failure or avoid trying. The *performance time score* is the sum of the performance times (in sec.) of 2 tasks of the TOWM and 4 tasks of the Wheelie test.

In addition to these two scoring methods that were already used in previous tests [8,11], an *anxiety score* (tested by using a visual-analog scale (VAS 0-10) prior to each task performance), and a *qualitative score* (based on skill maturity criteria), were introduced. The latter are unique scoring methods and as far we know, were never used before in instruments aiming to assess wheelchair skills. The intentions of developing these methods were first, to include the psychological impact (anxiety) on WM performance, and secondly, to qualitatively assessed each participant in detail, according to the maturity of his WM proficiency.[18]

A specific focus of this study was targeting the reliability assessment of this newly developed *quality scale;* in this scale, scoring is based on a process-oriented assessment. For each skill of the TOWM and the

Wheelie test, five components were selected as the most important quality "performance criteria". If a participant perfectly accomplishes these 5 components while performing a task, it constitutes a 'mature pattern of the skill'. This method provides meaningful feedback to both, the therapist and the client, regarding how well one performed each WM task, what exactly were the mistakes, how mature was the overall quality of his WM proficiency, and it allows later on to use the score sheet as a teaching aid and rehabilitation guide.

Testing procedures

A test – retest procedure was undertaken. In test occasion 1 (t1), the research team explained about the study aims, asked the participants to complete a personal information form and to sign the consent form. Body dimensions (participant's weight, height and body mass index (BMI=weight (kg) / height^2 (m)), were also taken. Following these procedures, the research team introduced the course of the tests and ran both tests (TOWM followed by the Wheelie test). After 1 week from test occasion 1, all the participants were tested again (t2) at the same place, time of the day and by the same research team (Figure 1). Both tests were performed while using participants' own daily wheelchairs and the research team did not change sitting position or configuration at the time of the test. Inspection of the wheelchair configuration (e.g. the calculation of the wheelchair center of gravity) was performed at t1 and at t2, to ensure the use of the same wheelchair (configuration and condition) at both test occasions. Center of gravity was tested according to the formula $X = (F1\times d) \times (m\times g)^{-1}$; $Y = X \times \text{cotg}\ (angl_{incl})$[22] (with F1 as the weight on the front wheels, d as the horizontal distance between rear and front wheel axle, and m x g as the total weight (person + wheelchair)). The y coordinate of center of gravity (vertical coordinate) is calculated as: $Y = x \cdot \text{cotg}$ $(angl_{incl})$. $Angl_{incl}$ is determined by the angle over which the wheelchair-user system had to be inclined to decrease x to zero).

The participants were asked to refrain from smoking, drinking alcohol and caffeine products for at least 2 hours before each trial. For the quality assessment (done after the test), participants were videotaped by the same photographer, using the same camera (Canon Basler 100Hz), placed on a marked line (the test map is presented in Figure 2)

Reliability assessment

Three raters were involved in the reliability analysis (Figure 1). (All with master's degree in physical therapy, with at least 6 months of experience working in SCI unit, and they all received the same training on using the TOWM and the Wheelie test). For both the test-retest and intrarater reliability assessment, the exact interval between the first and second rating was one week.

For test-retest reliability, rater number 1 scored during "real time" the ability scores, performance time scores and anxiety scores at t1 and at t2. The quality scores of a random selection 20 participant's videos taken at t1 and at t2 were assessed after the test, by rater number 2 (Figure 1). Stratification was performed to ensure that a full range of lesion levels was represented in the sample (i.e. individuals with tetraplegia, and with high and low paraplegia).

Intrarater reliability was assessed by rater number 2 who scored twice (repeatedly, one week apart) the same participants' videotapes (taken at t1). On the second evaluation, the rater was not permitted to review the results of his initial evaluation.

Interrater reliability was determined using the quality scores obtained at t1 by two independent raters (rater number 2 and rater number 3) (Figure 1).

Statistical procedures

Descriptive statistic and reliability analysis were performed using SPSS (version 16.0) and Microsoft Excel 2010. For measuring test-retest reliability, Wilcoxon signed-rank test was used to verify the absence of significant differences (systematic errors) between the measures at the two different times (t1 and t2). The interclass correlation coefficient (ICC) with a 95% CI was used as a measure of reliability for both, test-retest reliability as well as for inter- and intrarater reliability. A priori an ICC of 0.80 or higher was defined as an indication of good reliability.[20]

In addition to measuring the reliability of instruments and raters, the consistency or stability of repeated participants' responses was assessed (response stability). Response stability is basic to establish all other types of reliability, because if the response variable varies from measurement to measurement, it will not be possible to separate out errors due to the rater or instrument. Response stability was expressed in terms of Method Error (ME), Standard Error of Measurement (SEM) and coefficient of variation (CV_{ME}).[23]

Method error (ME) is a measure of the discrepancy between two sets of repeated scores, or their difference scores. Large difference scores reflect greater measurement error. Method error is often used as an adjunct to test-retest correlation statistics, as it reflects the percentage of variation from trial to trial, which the correlation coefficient does not. In addition, unlike the correlation coefficient, method error is not affected by lack of variation in raw scores. Method error was calculated using the standard deviation of the difference scores between test and retest:

$$ME = \frac{Sd}{\sqrt{2}}$$

Method error should be interpreted relative to the size of the mean differences. Therefore, it was converted to percentages using the coefficient of variation:

$$CV_{ME} = \frac{2ME}{\bar{X}_1+\bar{X}_2} \times 100$$

The standard error of measurement (SEM) provides a value for measurement error in the same units as the measurement itself, i.e. it is an indication of absolute reliability.[23] This type of reliability is more clinically applicable on a day-to-day basis, rather than a relative reliability co-efficient value, such as an ICC, which is more difficult to interpret for clinical decision-making. The SEM was also calculated using the ICC values as reliability coefficient for the data.[24]

$$SEM = SD \times \sqrt{(1 - ICC)}$$

The SEM allows the calculation of the Minimal Detectable Change (MDC), which is an estimate of the smallest change in score that can be detected objectively for a client, i.e. the amount by which a patient's score needs to change to be sure the change is greater than measurement error.[24] The MDC was calculated to a 95% degree of confidence:

$$MDC_{95} = SEM \times \sqrt{2} \times 1.96$$

The accuracy of the measurements was analysed by means of the technical error of measurement (TEM). The TEM index allows verifying the accuracy degree when performing and repeating measurements by the same appraiser (intrarater accuracy) and when comparing measurement with measurements from different appraisers (interrater accuracy).[25] TEM was calculated using the formula:

$$TEM = \sqrt{\frac{\sum_D 2}{2N}}$$

Where: D is the difference between measurements and N is the total number of measurements taken. The TEM presents the same measurement unit (cm, mm, points) and at least 20 measurements are required.[26]

Results

The mean, SD and range scores of the TOWM and the Wheelie test are presented in Table 1. Because of various task difficulties, not all persons were able to perform all the tasks of the TOWM and the Wheelie test. If a participant was not able to perform a given task after two trials, an ability score of '0' was given and no time score was available. As a result, the group composition of participants differs slightly for the different four scales.

The mean age of the participants was 38.8 ±8.0 years and the time since injury was 12.4 ±10.5 years. The high values of these standard deviations (SD) demonstrate the large variability in the population group. SCI lesion level ranged from C5 – L1, (tetraplegia n=6, high paraplegia n=16, low paraplegia n=7). The mean Body Mass Index (BMI kg/m^2) was 24.2 ±3.9.

The mean age of the 20 participants that were qualitatively assessed was 39.9 ± 7.12 years (range 26-52 years) and time since injury was 13.2 ± 11.5 years (range 1-35). Mean and SD of the TOWM ability scores of that group was 19.9 ± 4.8 (range 13-28) and of the Wheelie test ability scores 5.6 ± 2.4 (range 0-8). No significant differences were found between the characteristics of the two subsample groups (i.e. in age, time since injury and lesion level).

The center of gravity did not differ between t1 and t2, ensuring the use of the same wheelchair setup at both tests occasions, and ICC's (consistency of the center of gravity examination done by the examiners at t1 and t2 for axis x and y) were high (0.92-0.98).

Test – retest reliability based on total scores

No significant differences were found between t1 and t2 in the ability, quality and performance time total scores. Only the anxiety scores showed significant differences between t1 and t2 (The VAS anxiety score was lower at the retest).

For the TOWM, the ICCs associated with the test–retest reliability varied from 0.91 to 0.99 and the 95% CIs from 0.80 to 0.99; (Table 1). For the Wheelie test, the ICCs associated with the test–retest reliability varied from 0.94 to 0.99 and the 95% CIs from 0.87 to 0.99.

Response stability indexes are presented in Table 1. In general, response stability was higher for the ability and quality scales compared to the time scale. Anxiety scale had very low response stability, as expected due to the significant differences found between t1 and t2 in the anxiety total score.

For the TOWM total scores, standard error of measurement (SEM) values were 0.63 point for the total ability score and 1.29 point for the total quality score. Time and anxiety total scores showed much higher SEM. Similarly, method error (ME) and coefficient variation of method error (CV_{ME}) values were low for both ability and quality total score and much higher for the time and anxiety total score. Minimal detectable change (MDC_{95}) values were 1.74 point for the total ability score and 3.59 point for the total quality score and much higher for the time and anxiety total score.

For the Wheelie test total scores, standard error of measurement (SEM) values were 0.44 point for the total ability score and 1.15 point for the total quality score. Time and anxiety total score showed higher SEM. Method error (ME) and coefficient variation of method error (CV_{ME}) values were low for the ability, quality and time total score and much higher for the anxiety total score. Minimal detectable change (MDC_{95}) values were 1.23 point for the total ability scores and 3.20 point for the total quality score and much higher for the time and anxiety total scores.

Published version: http://dx.doi.org/10.1016/j.apmr.2012.09.023
Journal homepage: http://www.archives-pmr.org/

Test – retest reliability based on quality scores per task

The results of the test-retest reliability for the quality scores, per task, are presented in Table 2. Statistical significant differences between t1 and t2 were only found for the quality scores of the 'wheelie forward' task, but 'ascend sidewalk with a run up 10 cm', 'up a slope with a run up' and 'stationary wheelie', were nearly significant. The ICCs associated with the test–retest reliability for the TOWM and Wheelie test quality scores, per task, varied from 0.88 to 0.99 and the 95% CIs from 0.73 to 0.99, except for 'level propulsion forward' that had a lower ICC (0.43).

Examining the per task response stability indexes, showed lower stability while preforming 'level propulsion forward' and 'one handed propulsion' of the TOWM, and 'one handed wheelie' of the Wheelie test. Minimal detectable change (MDC_{95}) was higher for the first two tasks of the TOWM and Wheelie forward 10 m of the Wheelie test.

Intrarater reliability

The results of the intrarater reliability for the quality scores per task are presented in Table 3. No significant differences were found between the first and the second evaluation done by the same rater in the total quality scores of both tests. The ICC associated with the intrarater reliability of both the TOWM and Wheelie test total quality score was 0.99.

Intrarater technical error of measurement (TEM) for the TOWM quality total score was 1.07, and 1.06 for the Wheelie test (Table 3), which corresponds to a relative TEM of 3.7% and 6.3% respectively.

The analysis of each task separately showed significant differences between the assessments of the 'uneven surface' and 'accelerate and stop in a wheelie' tasks. For all tasks, ICCs were above 0.95 except for the 'level propulsion forward' task, which had a lower ICC (0.49). CIs values varied from 0.88 to 1.0, except for 'level propulsion forward' (0.27-0.80). Per task analysis of the TEM index revealed task's average TEM of 0.30 for the TOWM, and 0.38 for the Wheelie test.

Interrater reliability

The results of the interrater reliability are presented in Table 4.

No significant differences were found between the two raters based on the total quality scores of both tests. The ICC associated with the interrater reliability of both the TOWM and Wheelie test total quality score was 0.99.

Interrater TEM of the TOWM quality total score was 1.23, and 1.04 for the Wheelie test, which corresponds to a relative TEM of 4.3% and 6.1% respectively.

The analysis of each task separately showed a significant difference between the two raters for the 'descend 10 cm sidewalk' task. ICCs per task varied from 0.89 to 0.99 in both tests, except to 'level propulsion forward' (0.44). Per task analysis of the TEM index revealed tasks' average TEM of 0.44 for the TOWM, and 0.39 for the Wheelie test.

Published version: http://dx.doi.org/10.1016/j.apmr.2012.09.023
Journal homepage: http://www.archives-pmr.org/

Discussion

The study's objectives to determine the reliability, response stability and minimal detectable change of the TOWM and the Wheelie test were accomplished.

Test – retest reliability based on total scores

According to standards suggested by Eliasziw et al.,[27] the test-retest reliability ICCs values for the total ability, time, and quality scores of both tests were excellent. It indicates that both tests were able to measure WM with consistency and no learning or training effects were demonstrated. The only significant difference between t1 and t2 was confirmed for the anxiety scores. Participants in the retest were significantly less anxious. It can be explained by the psychological habituation effect of the first test trial. Once a person is familiar with the environment, with the tasks, and he feels secured and trusts the guard, (safety person), it is only logical that his anxiety in the second trial will be reduced. To avoid this learning effect in future studies, it is advised to have one dummy session (before t1) for habituation. However, the fact that the participants were less anxious in the second test did not affect the performance scores.

Response stability indexes, found in this study, revealed that the ability and quality scales are more reliable than the time and anxiety scales. The low SEM and CV_{ME} values for the TOWM and Wheelie test total ability and quality scores represent a reasonable measurement error indicating precision and high response stability for the ability and quality scales. However, total time score showed high response variability (especially for the TOWM). It indicates that the participants may have put more emphasis on performing the task well rather than performing it fast, leading to high time variability from trial to trial. This finding is consistent with previous studies arguing the relevancy of time measurement when assessing daily-wheeled mobility skills.[17] Because response stability indexes are based on the variability within the difference scores, they do not account well for the anxiety scale which had significant difference (systematic bias) between t1 and t2.

The minimal detectable change (MDC_{95}) values, found in this study, indicate that in a future intervention based on repeated measurements of the TOWM and the Wheelie test, total ability score of both tests will need to exceed about 1.5 points (on a scale range 0-30 for the TOWM, and 0-8 for the Wheelie test), and total quality scores will need to exceed about 3.5 points (TOWM scale range 0-50, Wheelie test scale range 0-40), in order to be confident (in 95%) that the difference is due to a real wheeled mobility ability change and not due to a measurement error. Since MDC_{95} values for wheeled mobility test-retest assessment were not reported in previous studies, it is recommended to establish MDC_{95} values for the TOWM and the Wheelie test for specific SCI characteristics (e.g. for paraplegics, tetraplegics, complete, incomplete, upon discharge from rehabilitation, experienced wheelchair users etc.). ICC's values, found in the current study, were compared to values reported in three previously published wheeled mobility tests.[5, 7, 12] The test-retest ICC values for the TOWM and Wheelie test were similar to those reported for the overall score of the Obstacle Course Assessment of Wheelchair Users Performance test (OCAWUP)[5], and for the Wheelchair Physical Functional Performance test (WC-PFP)[12], and higher than the test-retest ICC's reported for the Wheelchair Skill Test (WST 2.4)[7].

None of the previous wheeled mobility tests, published in the international literature, included reliability assessment on measurement error parameters such as response stability indexes or minimal detectable change[17]; therefore, the parameters reported in this study could not be compared with previous findings.

Test – retest reliability based on quality scores per task

The quality assessment per task showed that 'level propulsion forward' had low ICC values. This can be explained by the low variance in the scores, since all participants obtained a high score on this particular task.

The Wheelchair Skills Test (WST version 2.4) showed a success rate of 100% for level propulsion forward[7], from which it can be concluded that this task is not challenging enough. However, level propulsion forward is the task most frequently tested in the available tests evaluating wheelchair skill performance[17] and was graded "an extremely essential WM skill" in an international survey among wheelchair users.[19]

The current study observation may suggest omitting this task from the TOWM due to a ceiling effect. However, the sample included only post rehabilitation wheelchair users; therefore decision to omit this task should be made only after testing the TOWM with participants during their early stage of rehabilitation.

The significant differences between t1 and t2 in the quality scores of the 'wheelie forward' task may suggest a learning effect for this particular skill. To overcome this learning effect, it is recommended to have habituation session before t1.

Response stability analysis per task showed that 'One-handed propulsion' of the TOWM, and 'one-handed wheelie' of the Wheelie test had the highest response variability, pointing that it was difficult for the participants to reproduce the same results in the repeated experiment. A possible explanation may be that one-handed tasks were not familiar to the participants, as they are less required in daily life situations, leading to performance variability from trial to trial.

Minimal detectable change (MDC_{95}) analysis per task indicated that for the first two tasks of the TOWM (level propulsion and one-handed propulsion), and for wheelie forward on 10m line, at least 2 points change between t1 and t2 is needed (quality scale range per task 0-5), in order to claim a real performance change, for these tasks. It should be noted that a common characteristic of these three tasks is that the camera-shooting angle is in constant change (following the participant) while in other task the camera is fixed. This may be an explanation for a higher measurement error while video assessing of these tasks. A field test "on spot" assessment analysis is recommended in order to verify if the inconsistency is indeed due to the video analysis difficulties or that the task criteria are not clear enough and need refining.

Intrarater reliability

The intrarater ICC of the total quality scores of 0.99 for both, the TOWM and the Wheelie test is a higher value compared to intrarater reliability reported in three previous wheeled mobility tests [7,9,11].

Relative TEM of 3.7% for the intrarater assessment of the TOWM, and 6.3% for the Wheelie test were found.

There are no acceptable ranges for technical error of measurement (TEM) for neither interrater nor intrarater wheeled mobility skill assessment (unlike other domains, such as anthropometry, where the relative TEM for beginner anthropometrist for intrarater skin folds measurement is known to be 7.5% and 5% for a skillful anthropometrist).[28] Future studies should aim to establish relative TEM acceptable ranges in wheeled mobility assessment.

Intrarater analysis per task showed that ICC values for all tasks, except for 'level propulsion forward', were excellent. The low ICC value for 'level propulsion forward' (0.50) was a result of the small variance between participants' scores (variability among subjects' scores must be large in order to demonstrate reliability).

Published version: http://dx.doi.org/10.1016/j.apmr.2012.09.023
Journal homepage: http://www.archives-pmr.org/

The significant differences that were found between the two quality measurements of the 'uneven surface' and 'accelerate and stop in a wheelie', might indicate that the skill maturity criteria in these cases were not as clearly formulated as for the other skills. "The major protection against tester bias (i.e. the observer can be influenced by his memory of the first score), is to develop grading criteria that are as objective as possible and to train the tester in the use of the instrument".[23] On that account, it is advisable to review the quality criteria of these specific skills and to make adjustments as necessary.

Interrater reliability

Similar to the intrarater reliability, the interrater reliability value for the total quality scores, and for all tasks except 'level propulsion forward', were excellent. The interrater ICC values of both tests total quality scores (0.99) are higher than the ICC values reported to the overall scores of four previously published wheeled mobility tests.[5, 7, 9, 11]

The significant difference between the two raters that was found for 'descend 10 cm sidewalk', indicates the need to review the maturity criteria to maximize the reliability when qualitatively assessing this skill.

Limitations and future work

Only twenty-nine wheelchair users participated in the study in total. This is a rather small sample, but it is comparable with sample sizes used in other studies assessing the reliability of new wheelchair skill tests.[5, 10-12, 14] Despite the small sample, representation of different lesion levels and time passed since injury was a goal; therefore, it represents a wide range of spinal cord injury levels and it includes participants shortly after clinical rehabilitation next to participants who have been using a wheelchair for a long time. Consequently, the result may be quite representative to the wider population of SCI who are manual wheelchair users. The one week interval between the two test's occasions was set to avoid learning or training effects during the testing period. This seems to have been quite adequate, although the anxiety scores showed a 'learning' trend.

As suggested by Routhier et al,[6] the procedure of quality assessment by video, which was used in this study, might be a limitation, since the examiners cannot see everything done by the wheelchair user as a result of a 2D film and restricted view produced by the video, compared to human observation of real life conditions and actions. In addition, Kirby et al.[7] agreed that scoring from a silent videotape is more challenging than doing so in person and that it might underestimate the true reliability. It is suggested that in future research, the quality assessment will be carried out on field, to allow comparison between both evaluation methods and to reassess the reliability of the quality scale.

The interrater reliability in this study was based on two raters' assessment and for the quality criteria only. It is recommended that in a future study, interrater will be assessed by 3 raters and will be tested also for the ability scores.

The time scale was found as less sensitive and the anxiety scale showed a learning effect; it is recommended to reassess WM with the TOWM and the Wheelie test following a habituation session, and to perceive if response stability findings will change. If instability recurred, it would be suggested not to include these scales in future studies.

In the intrarater and interrater analysis, TEMs demonstrated more errors when assessing the Wheelie test tasks compared to the TOWM. This reflects not only the need for refining the quality criteria of the Wheelie test tasks, but also the necessity of technical training of evaluators on the quality assessment of the Wheelie test, in order to minimize the variability verified.

Conclusion

Results suggest an excellent reliability of the total ability and quality scores of the TOWM and the Wheelie test when, assessing WM of manual wheelchair users with SCI. It may be considered to omit the 'level propulsion' task from the TOWM, if a "ceiling effect" will be detected also among participants during the early rehabilitation phase. The quality criteria of 'descend 10 cm sidewalk' and 'one-handed propulsion' (from the TOWM), and 'wheelie forward', 'uneven surface' and 'accelerate and stop in a wheelie' (from the Wheelie test), need to be reviewed and possibly adjusted. Reducing the number of tasks may also be considered.

Acknowledgments

We thank Joeri Verellen, PhD, Maarten Abeel, MA, Elien Baeten, MA, Melina Boeckmans, MA, and Thomas Witdouckand, MA, for their assistance and contribution for the development and assessment of the new instruments.

Published version: http://dx.doi.org/10.1016/j.apmr.2012.09.023
Journal homepage: http://www.archives-pmr.org/

REFERENCES

1. Field-Fote E. (2009). Spinal Cord Injury Rehabilitation. Philadelphia, F.A. Davis Company.
2. World Health Organization. Towards a common language of function, disability and health: ICF. Geneva: WHO, 2002, 1–23.
3. Kilkens OJ, Post MW, Dallmeijer AJ, van Asbeck FW, Van Der Woude LH. (2005). Relationship between manual wheelchair skill performance and participation of persons with spinal cord injuries 1 year after discharge from inpatient rehabilitation. J Rehabil Res Dev; 42(3 Suppl 1):65-73.
4. Middleton JW, Harvey LA, Batty J, Cameron I, Quirk R, Winstanley J. (2006). Five additional mobility and locomotor items to improve responsiveness of the FIM in wheelchair-dependent individuals with spinal cord injury. Spinal Cord; 44:495-504.
5. Routhier F, Desrosiers J, Vincent C, Nadeau S. (2005). Reliability and construct validity of an obstacle course assessment of wheelchair user performance. Int J rehabil res; 28: 49-56.
6. Routhier F, Vincent C, Desrosiers J, Nadeau S, Guerette C. (2004). Development of an obstacle course assessment of wheelchair user performances (OCAWUP): a content validity study. Technol Disabil; 16:19-31.
7. Kirby RL, Dupuis DJ, MacPhee AH, Coolen AL, Smith C, Best KL, Newton AM, Mountain AD, MacLeod DA, Bonaparte JP. (2004). The wheelchair skills test (version 2.4): Measurement properties. Arch Phys Med Rehabil; 85: 794-804.
8. Kilkens O, Dallmeijer A, de Witte L, van der Woude L, Post M. (2004). The Wheelchair Circuit: construct validity and responsiveness of a test to assess manual wheelchair mobility in persons with spinal cord injury. Arch Phys Med Rehabil; 85 (3):424-431.
9. Stanley RK, Stafford DJ, Rasch E, Rodgers MM. (2003). Development of a functional assessment measure for manual wheelchair users. J Rehabil Res Dev; 40:301-307.
10. Kirby RL, Swuste J, Dupuis DJ, MacLeod DA, Monroe R. (2002). The Wheelchair Skills Test: A pilot study of a new outcome measure. Arch Phys Med Rehabil; 83(1):10-18.
11. Kilkens O, Post M, van der Woude L, Dallmeijer A, van den Hauvel W. (2002). The Wheelchair Circuit: reliability of a test to assess mobility in persons with spinal cord injuries. Arch Phys Med Rehabil; 83 (12):1783-1788.
12. Cress ME, Kinne S, Patrick DL, Maher E. (2002). Physical functional performance in persons using a manual wheelchair. J Orthop Phys Ther; 32 (3): 104-113.
13. Taricco M, Apolone G, Colombo C, Filardo G, Telaro E, Liberati A. (2000). Functional status in patients with spinal cord injury: A new standardized measurement scale. Gruppo Interdisciplinare Valutazione Interventi Riabilitativi. Arch Phys Med Rehabil; 81(9):1173-1180.
14. Harvey LA, Batty J, Fahey A. (1998). Reliability of a tool for assessing mobility in wheelchair-dependent paraplegics. Spinal Cord; 36(6):427-431.
15. Gans BM, Haley SM, Hallenborg SC, Mann N, Inacio CA, Faas RM. (1988). Description and interobserver reliability of the Tufts Assessment of Motor Performance. Am J Phys Med Rehabil; 67(5):202-210.
16. Jebsen RH, Trieschmann RB, Mikulic MA, Hartley RB, McMillan JA, Snook ME. (1970). Measurement of time in a standardized test of patient mobility. Arch Phys Med Rehabil;51(3):170-175.
17. Fliess-Douer O, Vanlandewijck, YC, Lubel Manor G, van der Woude LH. (2010). A systematic review of wheelchair skills tests for manual wheelchair users with a spinal cord injury: Towards a standardized outcome measure, Clinic Rehabil; 24: 867–886.
18. Fliess-Douer O, van der Woude LH, Vanlandewijck YC. (2012). Development of the Test of Wheeled Mobility (TOWM) and a short Wheelie test: A feasibility and validity study of two new tests assessing wheeled mobility skills in persons with spinal cord injuries. Clin Rehabil., Submitted
19. Fliess-Douer O, Vanlandewijck, YC, van der Woude LH (2012). Most essential wheeled mobility skills for daily life: An international survey among Paralympic wheelchair athletes with SCI. Arch Phys Med Rehabil; 98: 629-635
20. Rankin G, Stokes M. (1998). Reliability of assessment tools in rehabilitation: An illustration of appropriate statistical analyses. Clin Rehabil; 12:187-99.
21. Williams G.Robertson V, Greenwood K, Goldie P, Morris ME. (2005). The high-level mobility assessment tool (HiMAT) for traumatic brain injury. Part 2: Content validity and discriminability. Brain Injury; 19(10): 833–843.
22. Vanlandewijck Y, Theisen D, Daly D. (2001). Wheelchair propulsion biomechanics: implications for wheelchair sports. Sports Med.;31(5):339-367.
23. Portney LG, Watkins, MP. (1993). Foundation of clinical research – application to practice. Appleton & Lange.
24. McDowell I. (2006). Measuring health: a guide to rating scales and questionnaires, 3rd edition, Oxford University Press, Inc.
25. Goto R, Mascie-Taylor CGN. (2007). Precision of measurement as a component of human variation. J Physiol Anthropol 2007; 26: 253-56.
26. Perini TA, de Oliveira GL, Ornellas JS, de Oliveira FP. (2005). Technical error of measurement in anthropometry. Rev Bras Med Esporte; 11: 86-90.

27. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K. (1994). Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. Phy Ther; 74:777–788.
28. Norton K, Olds T. (2000). Antropometrica. Argentina: Biosystem.

SUPPLIERS LIST

a) Canon Basler 100Hz camera, Canon Inc., Tokyo, Japan; Local supplier: Canon België NV/SA, Berkenlaan 3 831 Diegem (Machelen)
b) IBM SPSS Inc., 233 S Wacker Drive 1100, Chicago, IL60606-6412, USA
c) Microsoft Excel 2010, Microsoft Corporation, Redmond WA 98052-6399 USA
d) Veterinary weighing scales Kern EOS 150K100NXL, Metil industrie bvba, Emile Idiersstraat 17 1160 BRUSSEL, BELGIUM
e) Industrial 53mm Diameter Hole Cable Protector 1.5m, The Safety Supply Company Ltd, Unit 4 Atlas Business Centre, Oxgate Lane London NW2 7HJ UK
f) Aluminum Wheelchair Access Ramps, Altec GmbH, D-78224 Singen, Germany

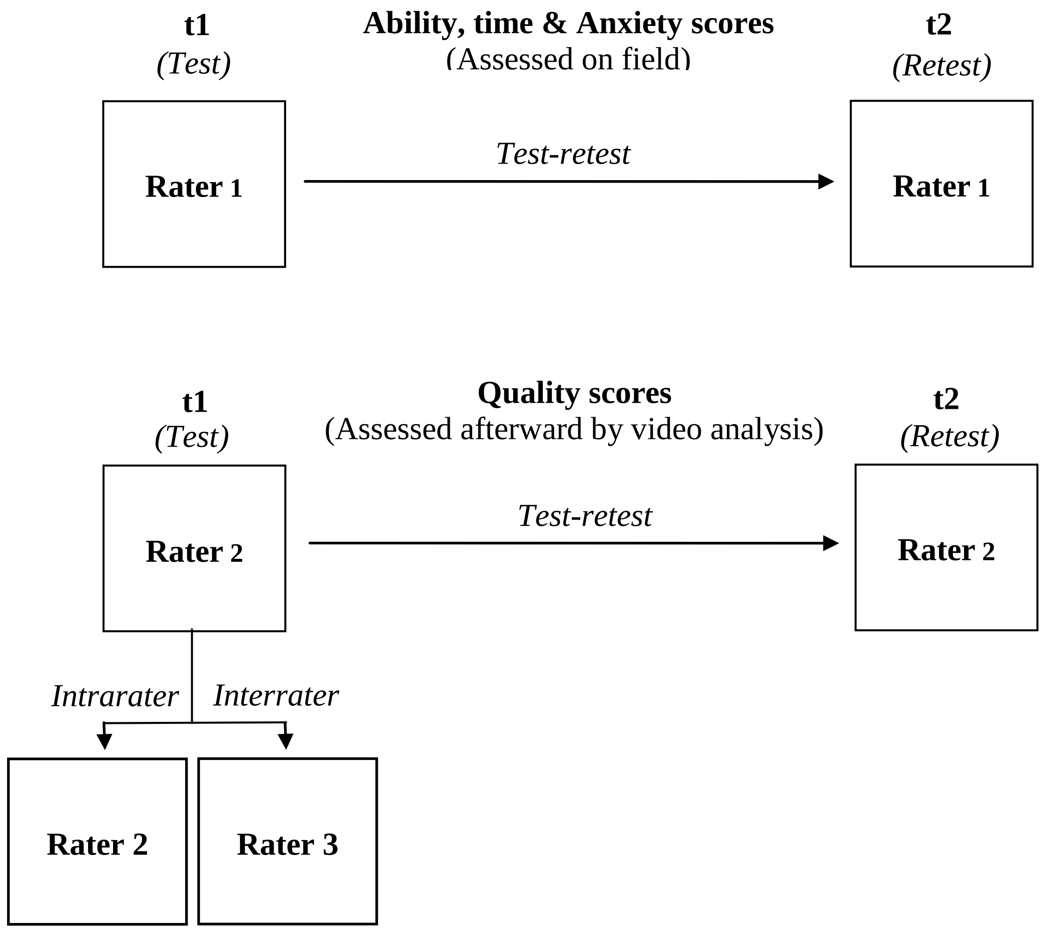

Figure 1: **test-retest, intrarater and interrater reliability**

Appendix 1: **Schematic test map (TOWM & Wheelie Test)**

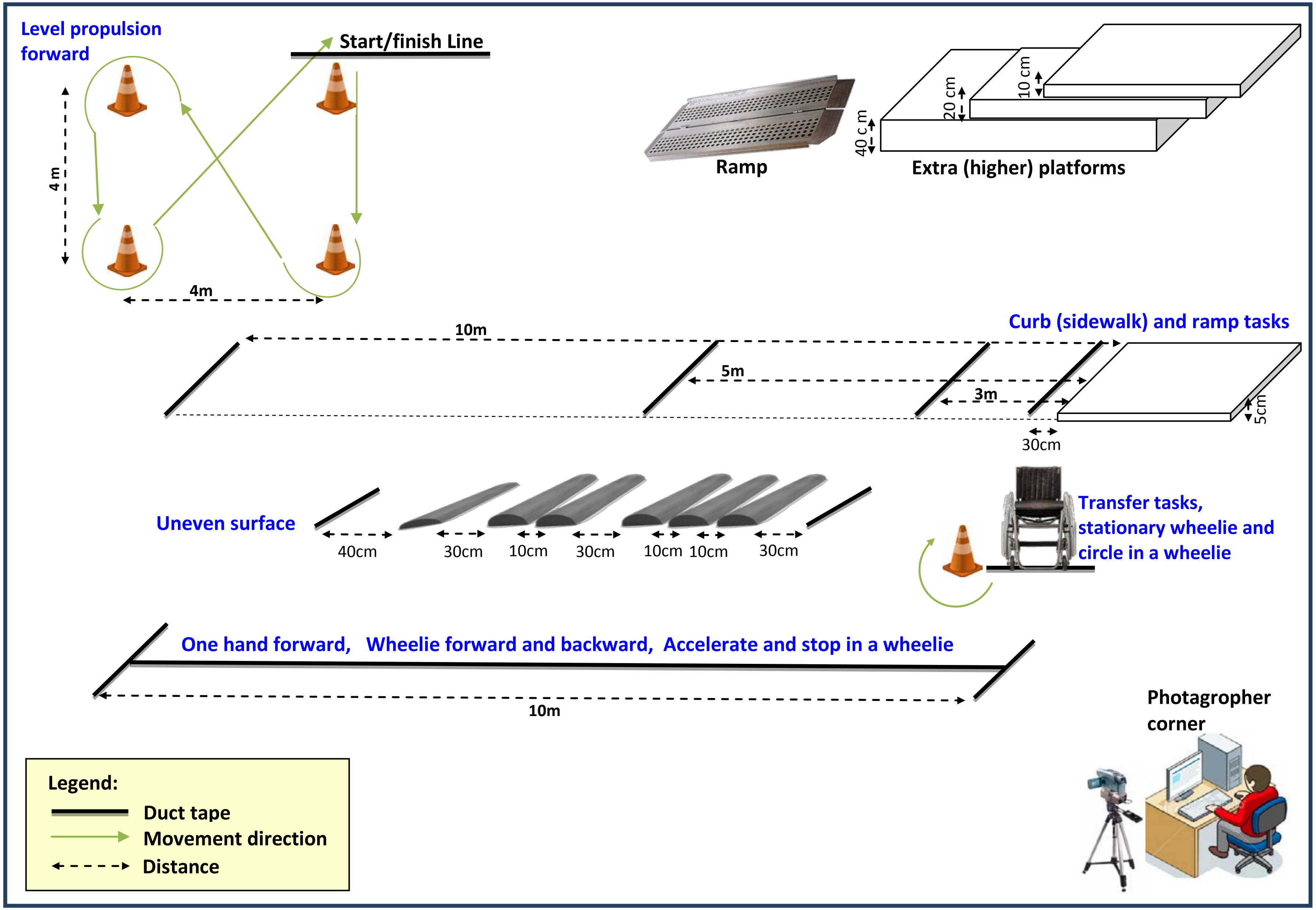

Table 1: **TOWM and the Wheelie test mean scores, range, test-retest reliability and responsiveness**

Scale	***n***	**t1 mean (SD) [range]**	**t2 mean (SD) [range]**	**t2-t1 Mean diff.**	**Sig.**	**ICC**	**95% CI**	**SEM**	**ME**	CV_{ME}	MDC_{95}
TOWM											
Ability score (*scale 0-30)*	29	20.6 (4.5) [11.5-28]	20.8 (4.5) [11.5-29]	0.2	0.34	0.98	.96-.99	0.63	0.88	4%	1.74
Quality score (scale 0-50)	20	28.8 (13.5) [7-48]	28.6 (12.7) [7-44]	0.2	0.81	0.99	.99-.99	1.29	1.16	4%	3.59
Time score (*in sec.)*	15	17.6 (8.3) [10.9-41]	17.5 (8.3) [9.6-39]	-0.1	0.82	0.94	.88-.97	5.73	9.08	26%	15.87
Anxiety score *(VAS 0-10 X 30 items)*	29	14.6 (21.5) [0-66]	8.7 (16.4) [0-71]	-5.9	0.01*	0.91	.80-.95	5.75	7.93	68%	15.93
Wheelie test											
Ability score (*scale 0-8)*	29	5.2 (2.3) [0-8]	5.5± (2.2) [0-8]	0.3	0.17	0.96	.91-.98	0.44	0.62	12%	1.23
Quality score (*scale 0-40)*	20	17 (11.9) [0-35]	16.5 (11.5) [0-37]	-0.5	0.15	0.99	.99-.99	1.15	1.38	8%	3.20
Time score (*in sec*.)	15	13.2 (5.5) [6.6-23.6]	12.6 (5.1) [6.9-23.9]	0.7	0.28	0.97	.92-.99	3.63	4.81	9%	10.07
Anxiety score *(VAS 0-10 X8 items)*	29	6.2 (10.7) [0-32]	4.7 (9) [0-36]	-1.5	0.08	0.94	.87-.97	2.41	3.27	62%	6.68

Wilcoxon signed-rank test was used to verify the absence of significant differences (systematic errors) between the measures at two different times (t1 and t2).
Sig: significant, *p*-value <0.05; ICC: Intraclass Correlation Coefficient; CI: Confidence Interval; SEM: Standard Error of Measurement; ME: Method Error; CV_{ME}: Coefficient variation of method error; MDC_{95}: Minimal Detectable Change (95% confidence)

Table 2: **Test-retest reliability of the quality scores per task (*n*=20)**

Task	**t1 Mean ± SD**	**t2 Mean ± SD**	**Sig.**	**ICC**	**95% CI**	**SEM**	**ME**	**CV_{ME}**	**MDC_{95}**
TOWM									
Level propulsion forward 4x4	4.3 ± 0.7	4.2 ± 0.6	0.56	0.43	-.43-.77	0.48	0.53	12%	1.34
One hand propulsion (10m)	2.8 ± 1.7	2.7 ± 1.7	0.41	0.88	.73-.95	0.57	0.45	17%	1.57
Ascend sidewalk 10 cm	0.8 ± 1.8	0.8 ± 1.8	1.00	1.00	1.00	0	0	0%	0
Descend sidewalk 10 cm	3.6 ± 1.7	3.7 ± 1.7	0.41	0.97	.93-.98	0.29	0.39	11%	0.80
Ascend sidewalk run up 10 cm	2.8 ± 2.4	3.0 ± 2.5	0.06	0.99	.98-.99	0.24	0.29	10%	0.67
Up a slope 15%	3.8 ± 1.6	3.9 ± 1.5	0.18	0.97	.93-.99	0.26	0.29	7%	0.73
Down a slope 15%	3.5 ± 1.6	3.6 ± 1.6	0.32	0.99	.98-.99	0.16	0.16	4%	0.43
Down a slope, stop in wheelie	2.1 ± 2.4	2.1 ± 2.4	1.00	1.00	1.00	0	0	0%	0
Up a slope with a run up	1.6 ± 2.1	1.5 ± 1.9	0.08	0.99	.97-.99	0.2	0.26	16%	0.55
Chair transfer stable	3.1 ± 1.8	2.9 ± 1.8	0.10	0.97	.94-.99	0.30	0.7	12%	0.84
Wheelie test									
Stationary wheelie	3.4 ± 1.3	3.5 ± 1.3	0.08	0.98	.95-.99	0.19	0.40	12%	0.52
One handed wheelie	1.2 ± 1.8	1.4 ± 2.0	0.18	0.98	.95-.99	0.27	0.53	40%	0.74
Wheelie forward 10 m	2.9 ± 1.9	2.4 ± 1.8	0.02*	0.94	.86-97	0.46	0.57	20%	1.27
Wheelie backward 10 m	1.9 ± 1.5	1.9 ± 1.6	0.56	0.98	.96-.99	0.22	0.36	19%	0.60
Circle forward	2.9 ± 2.0	2.9 ± 2.1	1.00	0.98	.95-.99	0.29	0.45	15%	0.81
Uneven surface	1.5 ± 1.7	1.2 ± 1.5	0.10	0.96	.88-.98	0.32	0.42	31%	0.87
Accelerate and stop in wheelie	1.7 ± 2.0	1.6 ± 1.9	0.32	0.98	.96-.99	0.26	0.42	28%	0.71
Backward over curb	1.4 ± 1.8	1.3 ± 1.7	0.56	0.98	.96-.99	0.26	0.39	27%	0.73

t1 = Test 1; t2 = Test 2 (retest); Sig: significant level *p*-value <0.05; ICC: Intraclass Correlation Coefficient; 95% CI: 95% Confidence; SEM: Standard Error of Measurement; ME: Method Error; CV_{ME}: Coefficient variation of method error; MDC_{95}: Minimal Detectable Change (95% confidence)

Table 3: **Intrarater reliability of the quality scores (*n*=20)**

Task	**t1 (1st)** **Mean ± SD**	**t1 (2nd)** **Mean ± SD**	**Sig.**	**ICC**	**95% CI**	**TEM**
TOWM						
Level Propulsion Forward 4x4	4.3 ± 0.7	4.2 ± 0.6	0.37	0.49	-.27-.80	0.55
One hand propulsion (10m)	2.8 ± 1.7	2.6 ± 1.6	0.10	0.96	.90-.98	0.57
Ascend sidewalk 10 cm	0.8 ± 1.8	0.8 ± 1.8	1.00	1.00	1.00	0
Descend sidewalk 10 cm	3.6 ± 1.7	3.7 ± 1.7	0.41	0.97	.93-.98	0.39
Ascend sidewalk run up 10 cm	2.8 ± 2.4	3.0 ± 2.5	0.07	0.99	.98-.99	0.32
Up a slope 15%	3.8 ± 1.6	4.0 ± 1.5	0.06	0.98	.95-.99	0.35
Down a slope 15%	3.5 ± 1.6	3.6 ± 1.6	0.32	0.99	.98-.99	0.16
Up a slope with a run up	1.6 ± 2.1	1.5 ± 1.9	0.08	0.99	.97-.99	0.27
Down a slope, stop in wheelie	2.1 ± 2.4	2.1 ± 2.4	1.00	1.00	1.00	0
Chair transfer stable	3.1 ± 1.8	2.9 ± 1.7	0.10	0.97	.94-.99	0.39
Total quality score TOWM	**28.7 ± 13.5**	**28.5 ± 12.7**	**0.66**	**0.99**	**.99-.99**	**1.07**
Wheelie test						
Stationary wheelie	3.4 ± 1.3	3.4 ± 1.3	1.00	0.95	.88-.98	0.27
One handed wheelie	1.2 ± 1.8	1.4 ± 2.0	0.41	0.96	.90-.98	0.35
Wheelie forward 10 m	2.9 ± 1.9	2.6 ± 1.8	0.11	0.95	.88-.98	0.67
Wheelie backward 10 m	1.9 ± 1.5	1.9 ± 1.6	0.65	0.97	.93-.98	0.27
Circle forward	2.9 ± 2.1	3.0 ± 2.1	0.48	0.97	.93-.99	0.39
Uneven surface	1.5 ± 1.8	1.1 ± 1.4	0.02*	0.96	.91-.98	0.47
Accelerate and stop in wheelie	1.7 ± 2.0	1.3 ± 1.6	0.01*	0.97	.93-.98	0.32
Backward over curb	1.4 ± 1.8	1.5 ± 1.9	0.41	0.97	.94-.99	0.27
Total quality score Wheelie test	**17.0 ± 11.9**	**16.4 ± 11.4**	**0.17**	**0.99**	**.98-.99**	**1.06**

t1 (1st) = Test occasion 1, first assessment; t1 (2nd) = Test occasion 1, second assessment done by the same rater one week after;
Sig.: Significant level *p*-value <0.05; ICC: Intraclass Coefficient; 95% CI: 95% Confidence Interval; TEM: Technical error of measurement
* Statistical significant differences

Table 4: **Interrater reliability of the quality scores (*n*=20)**

Task	**t1 R2 Mean ± SD**	**t1 R3 Mean ± SD**	**Sig.**	**ICC**	**95% CI**	**TEM**
TOWM						
Level Propulsion Forward 4x4	4.3 ± 0.7	4.4 ± 0.6	0.76	0.44	-.40-.78	0.52
One hand propulsion (10m)	2.8 ± 1.7	2.6 ± 1.6	0.38	0.89	.73-.95	0.71
Ascend sidewalk 10 cm	0.8 ± 1.8	0.7 ± 1.6	0.16	0.99	.97-.99	0.22
Descend sidewalk 10 cm	3.6 ± 1.7	3.2 ± 1.7	0.03*	0.94	.87-.98	0.59
Ascend sidewalk run up 10 cm	2.8 ± 2.4	2.7 ± 2.3	0.48	0.98	.95-.99	0.45
Up a slope 15%	3.8 ± 1.6	4.0 ± 1.4	0.06	0.98	.95-.99	0.42
Down a slope 15%	3.5 ± 1.6	3.6 ± 1.6	0.32	0.99	.98-.99	0.16
Up a slope with a run up	1.6 ± 2.1	1.7 ± 2.2	0.71	0.98	.95-.99	0.42
Down a slope, stop in wheelie	2.1 ± 2.4	1.9 ± 2.2	0.32	0.96	.91-.98	0.59
Chair transfer stable	3.1 ± 1.8	3.2 ± 1.8	0.65	0.98	.94-.99	0.35
Total quality score TOWM	**28.7 ± 13.5**	**28.1 ± 12.9**	**0.15**	**0.99**	**.99-.99**	**1.23**
Wheelie test						
Stationary wheelie	3.4 ± 1.3	3.3 ± 1.4	0.74	0.93	.83-.97	0.47
One handed wheelie	1.2 ± 1.8	1.1 ± 1.7	0.58	0.96	.88-.98	0.50
Wheelie forward 10 m	2.9 ± 1.9	2.8 ± 2.0	0.48	0.97	.93-.99	0.45
Wheelie backward 10 m	1.9 ± 1.5	1.9 ± 1.7	0.65	0.97	.93-.99	0.35
Circle forward	2.9 ± 2.0	2.9 ± 2.1	0.71	0.98	.94-.99	0.42
Uneven surface	1.5 ± 1.8	1.6 ± 1.9	0.41	0.98	.94-.99	0.39
Accelerate and stop in wheelie	1.7 ± 2.0	1.6 ± 1.9	0.32	0.99	.96-.99	0.32
Backward over curb	1.4 ± 1.8	1.5 ± 1.9	0.16	0.99	.98-.99	0.22
Total quality score Wheelie test	**17.0 ± 11.9**	**16.90 ± 12.51**	**0.57**	**0.99**	**.99-.99**	**1.04**

t1 R2 = Assessment of the quality scores obtained at test occasion 1 done by Rater 2; t1 R3 = Assessment of the quality scores obtained at test occasion 1done by Rater 3;
Sig: Significant level *p*-value <0.05; ICC: Intraclass Coefficient; 95% CI: 95% Confidence Interval; TEM: Technical error of measurement.
* Statistical significant differences

Research Coordination Office
Huis Bethlehem
Schapenstraat 34
B-3000 Leuven
Belgium
Tel.: +32 16 32 40 65
Fax: +32 16 32 41 98
onderzoek@kuleuven.be
www.kuleuven.be