University of Groningen

The power of standing up

Regterschot, Gerrit

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

CHAPTER 1

General introduction

G. R. H. (Ruben) Regterschot
Tobias Morat
Marjanne Folkersma
Wiebren Zijlstra

ABSTRACT

Leg muscle strength (LMS) and leg muscle power (LMP) are determinants of aspects of functional status and important parameters for measuring intervention effects in older adults. Field tests are often used for the evaluation of LMS and LMP in older persons. However, criteria important for the application of strength and power related field tests in older adults have not been systematically taken into account and are not yet fully listed and described in a single publication. Therefore, this chapter describes criteria important for the application of strength and power related field tests in older adults. In addition, strength and power related field tests commonly used in older adults are evaluated by using the described criteria. Based on this evaluation, this chapter provides a perspective on the further development of field tests. Criteria important for strength and power related field tests are: adequate accuracy, precision, concurrent validity, clinical validity, practical feasibility and pure strength or power outcomes. Commonly used strength and power related field tests do not meet all the aforementioned criteria. Therefore, further development of field tests is necessary. Mobile sensing systems are potentially useful for the evaluation of LMS and LMP in older adults. Mobile sensing systems do not have the limitations of commonly used field tests and provide important additional advantages. In particular, mobile sensing systems offer the opportunity of continuous monitoring during free-movement in the home-environment, thereby reducing the need of standardized assessments by health-care professionals. Future studies should examine the clinical validity of mobile sensing systems and evaluate the application of sensor technology in exercise-based interventions.

1

1. INTRODUCTION

Leg muscle strength (LMS) and leg muscle power (LMP) are major determinants of aspects of functional status in older people, such as mobility, activities of daily living and fall risk [1-8]. Therefore, LMS and LMP are considered important parameters for the identification of lower functioning individuals and the evaluation of intervention effects in older adults. Field tests are often used for the evaluation of LMS and LMP in older persons, because field tests are easy-to-use and applicable in clinical settings. Examples of commonly used field tests for the evaluation of LMS and LMP in older adults are hand-held dynamometry (HHD) [9], the Five-Times-Sit-to-Stand Test (FTSST) [10], stair walk tests [e.g. 11] and the Short Physical Performance Battery (SPPB) [12].

Strength and power related field tests need to fulfill methodological as well as feasibility criteria. Relevant criteria have been mentioned in different publications [e.g. 13,14]. However, criteria important for the application of strength and power related field tests in older adults have not been systematically taken into account and are not yet fully listed and described in a single publication. As a result, strength and power related field tests are applied in clinical settings without considering the essential methodological and/or practical feasibility criteria. For example, the FTSST has a limited ability to discriminate between higher and lower functioning individuals [10], and stair walk tests may not be feasible because at least half of the older adults between 75-79 years has difficulty with stair walking [15]. Therefore, the present chapter describes criteria important for the application of strength and power related field tests in older adults. In addition, this study evaluates strength and power related field tests commonly used in older adults by using the criteria described in the present chapter. Based on the results of this evaluation, the present chapter also provides a perspective on the further development of strength and power related field tests.

2. CRITERIA FOR THE APPLICATION OF STRENGTH AND POWER RELATED FIELD TESTS IN OLDER ADULTS

This section describes criteria important for the application of strength and power related field tests in older adults. The criteria can be categorized into general methodological criteria, criteria related to clinical validity and criteria related to feasibility. Whenever possible, the next sections will be short in addressing specific criteria by referring to further literature which more extensively treats the specific subject.

2.1 General methodological criteria

2.1.1 Accuracy

Accuracy is the closeness of agreement between a measured value and the true value or an accepted reference value [16]. For example, a new field test for the measurement of LMP may replace an accepted standard method if the new method shows high accuracy in comparison to an accepted gold standard method. Accuracy is usually quantified as the average difference between target values and measured values.

2.1.2 Precision

Precision is the closeness of agreement between repeated measurements [17]. The precision of an instrument that aims to measure a certain physical quantity is calculated by considering the variance of measured values [17]. When regarding the repeated execution of identical measurement procedures, precision consists of repeatability and reproducibility [17]. Repeatability (or test-retest reliability) is the precision when repeated measurements are performed in the same subjects under similar conditions [17,18]. Reproducibility is the precision when repeated measurements are performed in the same subjects but under different conditions, for example in different environments or with different testers [17].

Adequate repeatability is an important requirement for clinical validity of strength and power related field tests, because field tests should show small variation in outcome during repeated measurements under similar conditions in order to be useful for clinical assessments [13]. Measurement devices are often evaluated on two types of repeatability, namely relative repeatability and absolute repeatability [13,19]. Relative repeatability is defined as the consistency of the individuals rank in a sample over repeated measurements under similar conditions [19]. Absolute repeatability is defined as the extent to which individual scores vary during repeated measurements under similar conditions [19]. Relative repeatability is frequently evaluated with the intra-class correlation coefficient (ICC) [13,18-20]. The ICC is often interpreted according to the following criteria: ICC≥0.75 excellent repeatability, 0.40≤ICC<0.75 fair to good repeatability, and ICC<0.40 poor repeatability [21]. The standard error of measurement (SEM) [18,19] as well as Bland and Altman limits of agreement (LOA) [19,20,22] are often used to determine absolute repeatability of a measure. The smaller the SEM and the LOA, the better the absolute repeatability [19]. As to the acceptability of a specific size of SEM or LOA, this depends on specific measurement goals and the size of the differences one wants to measure. Hence, it is not possible to define absolute criteria for SEM and LOA.

2.1.3 Concurrent validity

A new measurement instrument demonstrates adequate concurrent validity when the new measurement instrument is associated with the outcomes of a previously validated measure [23]. For example, a new test for the measurement of LMS and LMP has adequate concurrent

validity when the new test is associated with the outcomes of a previously validated method for the measurement of LMS and LMP, such as isokinetic dynamometry [e.g. 24]. Concurrent validity of a measure is often evaluated using a correlation analysis [e.g. 24,25]. Correlations (r) may be interpreted as follows: little (if any correlation) when $0.00 < r \leq 0.25$; weak when $0.26 \leq r \leq 0.49$; moderate when $0.50 \leq r \leq 0.69$; strong when $0.70 \leq r \leq 0.89$; very strong when $0.90 \leq r \leq 1.00$ [26].

2.2 Criteria related to clinical validity

Discriminative ability and sensitivity to change are important criteria for clinical validity of strength and power related field tests, because they directly relate to diagnosis and evaluation of treatment effects. Both measurement properties are described in more detail below.

2.2.1 Discriminative ability

Discriminative ability is a measurement property that can be defined as the ability of a measure to correctly classify subjects into two different groups when true group belonging is known [13]. Adequate discriminative ability is important for clinical validity, because clinical measures should be able to discriminate between subjects with and without a certain condition [13]. For application in clinical settings strength and power related field tests should be able to discriminate between individuals with adequate strength and power, and individuals with insufficient strength and power. Individuals with insufficient LMS and LMP may be selected for participation in an exercise intervention aimed at improving LMS and LMP.

Discriminative ability is often evaluated by using the Receiver Operating Characteristic (ROC) curve [13,27]. ROC curves show the discriminative ability for different cut-off values of the measure (see fictitious data in Figure 1). The y-axis shows the percentage of subjects with a certain condition correctly classified by the measure as having this condition (sensitivity) and the x-axis shows the percentage of subjects wrongly classified as having the condition (1-specificity) [27]. The discriminative ability of a test is often evaluated by calculating the area under the ROC curve (AUC). The larger the AUC, the higher the discriminative ability of a test. The discriminative ability of a test is considered: non-informative when AUC=0.5, insufficiently accurate when $0.5 < AUC \leq 0.7$, moderately accurate when $0.7 < AUC \leq 0.9$, highly accurate when $0.9 < AUC < 1$, and perfect when AUC=1 [27].

2.2.2 Sensitivity to change

Sensitivity to change is a measurement property that can be defined as the ability of a measure to detect a change over time [13]. In order to be useful in clinical settings, strength and power related field tests should be able to detect changes over time. An adequate sensitivity to change of strength and power related field tests is essential for the evaluation of intervention effects. The change as the result of an intervention should be determined with validated measures before the change can be used to determine the sensitivity to change
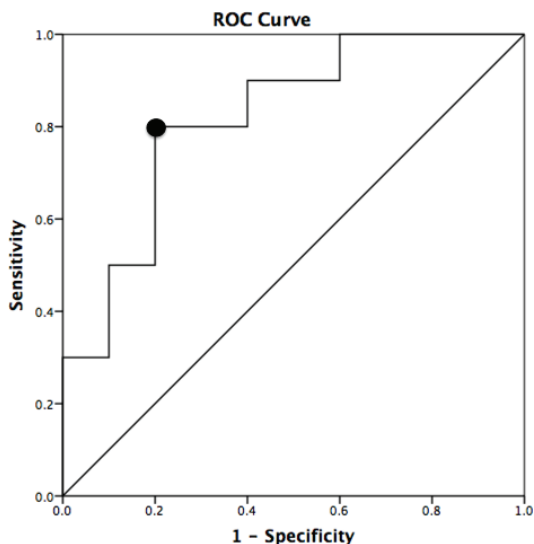
Figure 1 | Example of an ROC curve based on fictitious data from a leg power field test. The area under the ROC curve (AUC) is 0.82, indicating a moderately accurate discriminative ability of the field test. The optimal cut-off point is marked with a black circle. The sensitivity and specificity corresponding to the optimal cut-off point are respectively 0.80 and 0.80.

of a new measure. A high sensitivity to change is the result of a large mean change and a small standard deviation (SD) of change. Sensitivity to change is often evaluated using the standardized response mean (SRM) [13,28]. The SRM is an effect size measure and is calculated as: SRM = mean change/SD of change [28]. The SRM is considered small when $0.20 \leq SRM < 0.50$, moderate when $0.50 \leq SRM < 0.80$ and large when $SRM \geq 0.80$ [29].

2.3 Outcome measure

According to Horlings et al. (2008) strength measures can be divided into two groups: direct and indirect (surrogate) strength measures [4]. Direct strength measures are those that provide pure strength outcomes [4]. An example of a direct strength measure is HHD. Indirect strength measures are those that do not provide pure strength outcomes. Indirect measures evaluate strength by testing aspects of functional performance [4]. An example of an indirect strength measure is the FTSST [4]. The FTSST provides time duration in seconds as outcome measure and not pure strength or power outcomes in, for example, Newton (N) or Watt (W). Indirect measures of strength test not only muscle strength, but also other aspects, such as coordination [4]. Therefore, if strength or power is an important outcome measure of the study it is recommended to use direct measures of strength or power instead of indirect measures [4].

2.4 Criteria related to practical feasibility

Field tests for the evaluation of LMS and LMP in older adults are developed for use in environments outside the laboratory, such as clinical environments, home settings or community settings. Therefore, strength and power related field tests should be portable, lightweight and executable in a home environment [14,30]. However, also other factors should be taken into account when considering the practical feasibility of strength and power related field tests: the duration of the measurement, the amount of space required for the measurement, simplicity and acceptability of the test since older adults should be able to perform the test, safety as well as the risk on muscle soreness and injuries, ease of data acquisition and data analysis since clinicians without extensive experience with laboratory devices should be able to administer the test, costs associated with the test and specificity of testing since the movement pattern, contraction type (eccentric, concentric or isometric) and contraction velocity of leg muscles during daily life activities should be transferred to field tests in order to test as specific as possible [14,30].

3. EVALUATION OF STRENGTH AND POWER RELATED FIELD TESTS

In this section we will evaluate strength and power related field tests commonly used in older adults. First we will summarize findings of the few available studies that evaluated strength and power related field tests commonly used in older adults on aspects of repeatability, validity and practical feasibility. Subsequently we will evaluate field tests considered valid and feasible by these studies using the criteria described in the previous section of the present chapter.

A recent review study by Mijnarends et al. (2013) evaluated commonly used strength and power related field tests on aspects of repeatability, validity and practical feasibility in older adults [30]. The results of this review study showed that only HHD, the SPPB and gait speed (GS) over a short distance have adequate repeatability, concurrent validity, construct validity and practical feasibility in older adults. In addition, Stark et al. (2011) concluded that HHD is reliable, valid and practical for the measurement of muscle strength in young as well as older adults [9]. Furthermore, Freiberger et al. (2012) showed that the SPPB has adequate repeatability, responsiveness and validity in older adults [31]. Moreover, Rydwik et al. (2012) concluded that habitual GS is reliable and valid in older adults, however, this study also revealed that the responsiveness of GS is unclear [32]. Together these review studies indicate that HHD, SPPB and GS tests are reliable, valid and feasible in older adults.

However, a critical reflection on HHD, SPPB and GS tests is necessary since these tests have important limitations when the criteria described in the previous section of the present chapter are taken into account. First, as already noted, the responsiveness of GS tests is undetermined [32]. For this reason, the clinical validity of GS tests remains unclear

and needs further investigation. Second, the SPPB and GS tests do not provide pure strength or power outcome measures. This is a limitation of the SPPB and GS tests since field tests providing pure strength or power outcome measures are recommended, in particular when strength and power are important outcome measures of a study [4]. Third, repeatability of LMS assessments with HHD is inadequate when individuals produce high muscle forces [33]. When an individual produces high muscle forces it is difficult for the therapist to keep the hand-held dynamometer in position during the assessment, thereby reducing the repeatability of the measurement [33,34]. Fourth, practical feasibility of the SPPB and GS tests is limited, because both tests require a relatively large amount of space, which may be problematic in home settings [31]. Furthermore, the performance of the SPPB requires much time (10-15 minutes) [12] compared to the performance of other field tests (e.g. HHD), which further reduces the practical feasibility of the SPPB. Hence, HHD, SPPB and GS tests have serious limitations, which indicate the need for the development of alternative field tests for the evaluation of LMS and LMP in older adults.

4. FUTURE PERSPECTIVE AND REMAINING CHALLENGES

The preceding section demonstrated the need for the further development of field tests for the evaluation of LMS and LMP in older adults. In the present section we will provide a perspective on the further development of strength and power related field tests and we will identify remaining challenges.

With recent developments in technology, mobile sensing systems have become available for the measurement of motor functioning. For example, methods based on body-fixed motion sensors have been developed to assess aspects of mobility in older people [e.g. 35,36]. In addition, gaming systems (e.g. Xbox Kinect, Nintendo Wii) are currently available for the measurement and training of motor functioning in older adults [37]. Moreover, mobile sensing systems have been developed for the measurement of LMS and LMP in older adults. For example, body-fixed motion sensors have been applied to estimate leg power during the sit-to-stand (STS) movement in young and older adults [38]. For the estimation of power during STS body-fixed motion sensors have been attached to the right side of the hip and the chest (see Figure 2) [38]. The vertical acceleration signal measured with the body-fixed motion sensors can be used to estimate the vertical peak power of the body's center of mass during the STS movement [38]. In addition, methods based on force plates are currently available for the assessment of power during the chair rise transfer [39-42]. Force plate methods use the vertical ground reaction force measured with force plates beneath the feet of a person to estimate the vertical power during the STS movement [39-42].

Mobile sensing systems for the measurement of LMS and LMP do not have the limitations associated with commonly used strength and power related field tests. While the SPPB and GS tests do not provide pure strength or power outcomes, mobile sensing

1

Figure 2 | A motion sensor can be used to estimate power during the sit-to-stand movement. The small black box on the right side of the hip represents a body-fixed motion sensor. The vertical acceleration signal measured with the motion sensor can be used to estimate the vertical peak power of the body's center of mass during the sit-to-stand transfer [38].

systems with pure strength and power outcome measures are already available. For example, a sensor-based method for the evaluation of sit-to-stand performance provides pure power outcomes with adequate concurrent validity [38]. In addition, mobile sensing systems have been developed as a solution for the inadequate repeatability of HHD when high muscle forces are generated. For example, studies integrated force sensors in a fixed station for the assessment of isometric quadriceps strength [34,43]. By using a fixed station the quadriceps force measurement does not depend on the strength of the therapist. Studies showed that a force sensor integrated in a fixed station provides reliable assessments of quadriceps strength in young adults as well as older adults [34,44]. Furthermore, in contrast to commonly used strength and power related field tests (such as the SPPB and GS tests) mobile sensing systems for the measurement of LMS and LMP have a high practical feasibility. For example, sensor-based assessments of leg power production during sit-to-stand movements can easily be performed in home settings and take only a few minutes

[38]. Furthermore, body-fixed sensor systems are highly portable and lightweight [45].

Mobile sensing systems not only solve the issues associated with commonly used strength and power related field tests, but also offer important additional advantages. In particular, specific mobile sensing systems provide the possibility of assessments in the home environment on a day-to-day or week-to-week basis without a clinician being physically present. For this reason, the clinical relevance of mobile sensing systems can be enormous. This is in particular true for systems based on body-fixed motion sensors. Body-fixed motion sensors can be worn during free-movement in the home environment and are therefore ideal for the continuous monitoring of LMS and LMP in older adults [35,36,46]. Continuous monitoring with body-fixed motion sensors during daily life activities has the potential to provide early indications of functional decline and intervention effects. Outcomes of sensor-based assessments can be used by older adults for self-monitoring and by health-care professionals to individualize exercise programs through remote feedback [46]. As a result of this, the performance of standardized assessments by health-care professionals may become less needed. Health-care professionals may only perform a detailed examination of functioning when a functional decline has been detected based on daily life monitoring with sensor technology.

However, the clinical relevance of sensor-based technology for the evaluation of strength and power has not yet been fully demonstrated. With a few exceptions studies report on technical validity and feasibility of sensor technology for the measurement of motor functioning (with the majority of studies showing adequate technical validity and feasibility), but not on the clinical validity of sensor technology [35,47]. Therefore, future studies should focus on investigating the clinical validity of sensor-based technology in older persons.

5. CONCLUSIONS

The present chapter described criteria important for the application of strength and power related field tests in older adults. Based on these criteria, we demonstrated that strength and power related field tests commonly used in older adults each have their very specific limitations. Mobile sensing systems for the evaluation of LMS and LMP solve the issues associated with commonly used field tests and provide important additional advantages. In particular, mobile sensing systems offer the opportunity of continuous monitoring during free-movement in the home-environment, thereby limiting the need for standardized assessments by health-care professionals. Future research should investigate the clinical validity of mobile sensing systems.

6. AIMS AND OUTLINE OF THE THESIS

So far there is no clinically valid sensor-based method for the estimation of leg power in older adults. For this reason, the primary aim of this thesis is to develop a method based on body-fixed motion sensors for the estimation of the power that is produced to lift the body's center of mass (CoM) during the sit-to-stand (STS) movement in older adults, and to evaluate the method's clinical validity. A secondary aim is to compare the clinical relevance of sensor-based estimation of power during STS and the clinical relevance of sensor-based estimation of other quantities (e.g. maximal acceleration, maximal velocity) during STS in older adults. A tertiary aim is to compare the clinical relevance of sensor-based STS measures (e.g. peak power, maximal acceleration, maximal velocity) and the clinical relevance of field tests commonly used for the evaluation of leg strength, leg power and mobility (e.g. Five-Times-Sit-to-Stand Test, Timed-Up-and-Go Test) in older adults. To achieve these aims, measures of STS performance are developed based on data of body-fixed motion sensors collected in standardized conditions. Subsequently, sensor-based STS performance measures are evaluated on repeatability, sensitivity to change, discriminative ability, construct validity, accuracy and concurrent validity. The evaluation of the sensor-based STS measures and standard clinical measures on methodological criteria is based on different studies. The methods and results of these studies are presented in the following chapters of this thesis.

In **chapter 2**, the repeatability of sensor-based estimation of STS peak power is investigated and compared to the repeatability of other (non power) sensor-based STS measures in young and older adults. In addition, the repeatability of sensor-based STS measures is compared to the repeatability of several clinical measures commonly used for the evaluation of leg strength, leg power and mobility in older adults. In chapter 2 a body-fixed motion sensor on the right side of the hip is used.

In **chapter 3**, the sensitivity to change of sensor-based estimation of STS peak power is investigated and compared to the sensitivity to change of other (non power) sensor-based STS measures. Furthermore, the sensitivity to change of sensor-based STS measures is compared to the sensitivity to change of measures commonly used for the evaluation of leg strength, leg power and mobility in older adults. The sensitivity to change of sensor-based STS measures and standard clinical measures is investigated by applying an 8 week exercise-based intervention aimed at improving leg strength, leg power and balance. Assessments are performed before and after the intervention to determine sensitivity to change. A body-fixed motion sensor on the right side of the hip is used for the investigation of sensitivity to change.

In **chapter 4**, sensor-based measures of STS peak power are evaluated on the ability to discriminate between higher and lower functioning individuals. Furthermore, the construct validity of sensor-based measures of STS peak power is investigated and compared to the construct validity of other (non power) sensor-based STS measures. Construct validity is determined by investigating the associations between sensor-based STS measures and standard clinical measures of leg strength, leg

power, mobility, activity limitations and frailty. STS measures from two sensors are evaluated on discriminative ability and construct validity, namely STS measures from a sensor on the right side of the hip and STS measures from a sensor on the chest.

In **chapter 5**, accuracy and concurrent validity of sensor-based estimations of STS peak power are evaluated in older adults. In addition, the accuracy and concurrent validity of sensor-based measures of STS peak power are compared to the accuracy and concurrent validity of other (non power) sensor-based STS measures. Accuracy and concurrent validity are not essential for clinical validity. However, accuracy and concurrent validity are important for the interpretation of sensor-based STS measures and are therefore topics of investigation. Accuracy and concurrent validity of sensor-based STS measures are investigated by comparing the sensor outcomes to the outcomes of a measurement method based on force plates. In chapter 5, two sensors are applied for the assessment of STS performance, namely a sensor on the right side of the hip and a sensor on the chest.

In **chapter 6**, the findings of the thesis are summarized and discussed. In addition, it describes practical implications, critical reflections, recommendations for further research, and a future perspective on clinical assessments and exercise-based interventions. Furthermore, chapter 6 describes a general trend in the medical sciences toward the development of individualized treatment strategies based on continuous health monitoring. Also the relevance of health monitoring technology and personalized treatment strategies are described as well as the potential dangers of the application of health monitoring technology on a large scale in society. The chapter ends with the conclusions of this thesis.

REFERENCES

1. Bean JF, Kiely DK, Herman S, Leveille SG, Mizer K, Frontera WR, Fielding RA. The relationship between leg power and physical performance in mobility-limited older people. J Am Geriatr Soc 2002;50:461-467.

2. Bean JF, Leveille SG, Kiely DK, Bandinelli S, Guralnik JM, Ferrucci L. A comparison of leg power and leg strength within the InCHIANTI study: which influences mobility more? J Gerontol A Biol Sci Med Sci 2003;58:728-733.

3. Foldvari M, Clark M, Laviolette LC, Bernstein MA, Kaliton D, Castaneda C, Pu CT, Hausdorff JM, Fielding RA, Singh MA. Association of muscle power with functional status in community-dwelling elderly women. J Gerontol A Biol Sci Med Sci 2000;55:M192-199.

4. Horlings CG, van Engelen BG, Allum JH, Bloem BR. A weak balance: the contribution of muscle weakness to postural instability and falls. Nat Clin Pract Neurol 2008;4:504-515. doi: 10.1038/ncpneuro0886.

5. Perry MC, Carville SF, Smith IC, Rutherford OM, Newham DJ. Strength, power output and symmetry of leg muscles: effect of age and history of falling. Eur J Appl Physiol 2007;100:553-561.

6. Reid KF, Fielding RA. Skeletal muscle power: a critical determinant of physical functioning in older adults. Exerc Sport Sci Rev 2012;40:4-12. doi: 10.1097/JES.0b013e31823b5f13.

7. Skelton DA, Greig CA, Davies JM, Young A. Strength, power and related functional ability of healthy people aged 65-89 years. Age Ageing 1994;23:371-377.

8. Suzuki T, Bean JF, Fielding RA. Muscle power of the ankle flexors predicts functional performance in community-dwelling older women. J Am Geriatr Soc 2001;49:1161-1167.

9. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. PM R 2011;3:472-479. doi: 10.1016/j.pmrj.2010.10.025.

10. Whitney SL, Wrisley DM, Marchetti GF, Gee MA, Redfern MS, Furman JM. Clinical measurement of sit-to-stand performance in people with balance disorders: validity of data for the Five-Times-Sit-to-Stand Test. Phys Ther 2005;85:1034-1045.

11. Bean JF, Kiely DK, LaRose S, Alian J, Frontera WR. Is stair climb power a clinically relevant measure of leg power impairments in at-risk older adults? Arch Phys Med Rehabil 2007;88:604-609.

12. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, Wallace RB. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49:M85-94.

13. Moe-Nilssen R, Nordin E, Lundin-Olsson L; Work Package 3 of European Community Research Network Prevention of Falls Network Europe. Criteria for evaluation of measurement properties of clinical balance measures for use in fall prevention studies. J Eval Clin Pract 2008;14:236-240. doi: 10.1111/j.1365-2753.2007.00839.x.

14. Powers SK, Howley ET. Exercise physiology: theory and application to fitness and performance. New York: McGraw-Hill; 2004.

15. Malmberg JJ, Miilunpalo SI, Vuori IM, Pasanen ME, Oja P, Haapanen-Niemi NA. A health-related fitness and functional performance test battery for middle-aged and older adults: feasibility and health-related content validity. Arch Phys Med Rehabil 2002;83:666-677.

16. Working Group 2 of the Joint Committee for Guides in Metrology (JCGM/WG 2) (2008) International vocabulary of metrology — Basic and general concepts and associated terms (VIM). Available on www.bipm.org

17. Pryseley A, Mintiens K, Knapen K, Van der Stede Y, Molenberghs G..Estimating precision, repeatability, and reproducibility from Gaussian and non- Gaussian data: a mixed models approach. J Appl Stat 2010;37:1729–1747.

18. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 2005;19:231-240.

19. Bruton A, Conway JH, Holgate ST. Reliability: What is it and how is it measured? Physiotherapy 2000;86:94-99.

20. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil 1998;12:187-199.

21. Fleiss JL. The Design and Analysis of Clinical Experiments. New York: John Wiley; 1986.

22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-310.

23. Gouttebarge V, Wind H, Kuijer PP, Frings-Dresen MH. Reliability and validity of Functional Capacity Evaluation methods: a systematic review with reference to Blankenship system, Ergos work simulator, Ergo-Kit and Isernhagen work system. Int Arch Occup Environ Health 2004;77:527-537.

24. Taylor JD, Fletcher JP. Correlation between the 8-repetition maximum test and isokinetic dynamometry in the measurement of muscle strength of the knee extensors: A concurrent validity study. Physiother Theory Pract 2013;29:335-341. doi: 10.3109/09593985.2012.727529.

25. Campanini I, Merlo A. Reliability, smallest real difference and concurrent validity of indices computed from GRF components in gait of stroke patients. Gait Posture 2009;30:127-131. doi: 10.1016/j.gaitpost.2009.03.011.

26. Domholdt E. Physical Therapy Research: Principles and Applications. Philadelphia, PA: Saunders; 2000.

27. Greiner M, Pfeiffer D, Smith RD. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med 2000;45:23-41.

28. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459-468.

29. Cohen J. A power primer. Psychol Bull 1992;112:155-159.

30. Mijnarends DM, Meijers JM, Halfens RJ, ter Borg S, Luiking YC, Verlaan S, Schoberer D, Cruz Jentoft AJ, van Loon LJ, Schols JM. Validity and reliability of tools to measure muscle mass, strength, and physical performance in community-dwelling older people: a systematic review. J Am Med Dir Assoc 2013;14:170-178. doi: 10.1016/j.jamda.2012.10.009.

31. Freiberger E, de Vreede P, Schoene D, Rydwik E, Mueller V, Frändin K, Hopman-Rock M. Performance-based physical function in older community-dwelling persons: a systematic review of instruments. Age Ageing 2012;41:712-721. doi: 10.1093/ageing/afs099.

32. Rydwik E, Bergland A, Forsén L, Frändin K. Investigation into the reliability and validity of the measurement of elderly people's clinical walking speed: a systematic review. Physiother Theory Pract 2012;28:238-256. doi: 10.3109/09593985.2011.601804.

33. Agre JC, Magness JL, Hull SZ, Wright KC, Baxter TL, Patterson R, Stradel L. Strength testing with a portable dynamometer: reliability for upper and lower extremities. Arch Phys Med Rehabil 1987; 68:454-458.

34. Verkerke GJ, Lemmink KA, Slagers AJ, Westhoff MH, van Riet GA, Rakhorst G. Precision, comfort and mechanical performance of the Quadriso-tester, a quadriceps force measuring device. Med Biol Eng Comput 2003;41:283-289.

35. Zijlstra W, Aminian K. Mobility assessment in older people, new possibilities and challenges. European Journal of Ageing 2007;4:3–12.

36. Zijlstra W, Becker C, Pfeiffer K. Wearable systems for monitoring mobility related activities: from technology to application for healthcare services. In: Ziefle M, Röcker C (ed) E-Health, Assistive Technologies and Applications for Assisted Living. Hershey PA: IGI Global; 2011. pp 244-267.

37. Taylor MJ, McCormick D, Shawis T, Impson

R, Griffin M. Activity-promoting gaming systems in exercise and rehabilitation. J Rehabil Res Dev 2011;48:1171-1186.

38. Zijlstra W, Bisseling RW, Schlumbohm S, Baldus H. A body-fixed-sensor-based analysis of power during sit-to-stand movements. Gait Posture 2010;31:272-278. doi: 10.1016/j.gaitpost.2009.11.003.

39. Chen PY, Wei SH, Hsieh WL, Cheen JR, Chen LK, Kao CL. Lower limb power rehabilitation (LLPR) using interactive video game for improvement of balance function in older people. Arch Gerontol Geriatr 2012;55:677-682. doi: 10.1016/j.archger.2012.05.012.

40. Cheng YY, Wei SH, Chen PY, Tsai MW, Cheng IC, Liu DH, Kao CL. Can sit-to-stand lower limb muscle power predict fall status? Gait Posture 2014;40:403-407. doi: 10.1016/j.gaitpost.2014.05.064.

41. Lindemann U, Muche R, Stuber M, Zijlstra W, Hauer K, Becker C. Coordination of strength exertion during the chair-rise movement in very old people. J Gerontol A Biol Sci Med Sci 2007;62:636-640.

42. Zech A, Steib S, Sportwiss D, Freiberger E, Pfeifer K. Functional muscle power testing in young, middle-aged, and community-dwelling nonfrail and prefrail older adults. Arch Phys Med Rehabil 2011;92:967-971. doi: 10.1016/j.apmr.2010.12.031.

43. Nadler SF, DePrince ML, Hauesien N, Malanga GA, Stitik TP, Price E. Portable dynamometer anchoring station for measuring strength of the hip extensors and abductors. Arch Phys Med Rehabil 2000;81:1072-1076.

44. Douma KW, Regterschot GRH, Krijnen WP, Slager GEC, van der Schans CP, Zijlstra W. Reliability of the Q Force; a mobile instrument for measuring isometric quadriceps muscle strength. Submitted.

45. Van Acht V, Bongers E, Lambert N, Verberne R. Miniature wireless inertial sensor for measuring human motions. Conf Proc IEEE Eng Med Biol Soc 2007;2007:6279-6282.

46. Geraedts HA, Zijlstra W, Zhang W, Bulstra S, Stevens M. Adherence to and effectiveness of an individually tailored home-based exercise program for frail older adults, driven by mobility monitoring: design of a prospective cohort study. BMC Public Health 2014;14:570. doi: 10.1186/1471-2458-14-570.

47. de Bruin ED, Hartmann A, Uebelhart D, Murer K, Zijlstra W. Wearable systems for monitoring mobility-related activities in older people: a systematic review. Clin Rehabil 2008;22:878-895. doi: 10.1177/0269215508090675.

1