




Assessment of Muscle Function and Physical Performance in Daily Clinical Practice

A position paper endorsed by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO)

Charlotte Beudart^{1,2} · Yves Rolland³ · Alfonso J. Cruz-Jentoft⁴ · Jürgen M. Bauer⁵ · Cornel Sieber⁶ · Cyrus Cooper^{7,8} · Nasser Al-Daghri⁹ · Islene Araujo de Carvalho¹⁰ · Ivan Bautmans¹¹ · Roberto Bernabei¹² · Olivier Bruyère^{1,2}  · Matteo Cesari^{13,14} · Antonio Cherubini¹⁵ · Bess Dawson-Hughes¹⁶ · John A. Kanis^{17,18} · Jean-Marc Kaufman¹⁹ · Francesco Landi¹² · Stefania Maggi²⁰ · Eugene McCloskey^{17,18,19,20,21} · Jean Petermans²² · Leocadio Rodriguez Mañas²³ · Jean-Yves Reginster^{1,2} · Regina Roller-Wirnsberger²⁴ · Laura A. Schaap²⁵ · Daniel Uebelhart²⁶ · René Rizzoli²⁷ · Roger A. Fielding²⁸

Received: 7 November 2018 / Accepted: 29 March 2019
© The Author(s) 2019

Abstract

It is well recognized that poor muscle function and poor physical performance are strong predictors of clinically relevant adverse events in older people. Given the large number of approaches to measure muscle function and physical performance, clinicians often struggle to choose a tool that is appropriate and validated for the population of older people they deal with. In this paper, an overview of different methods available and applicable in clinical settings is proposed. This paper is based on literature reviews performed by members of the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO) working group on frailty and sarcopenia. Face-to-face meetings were organized afterwards where the whole group could amend and discuss the recommendations further. Several characteristics should be considered when choosing a tool: (1) purpose of the assessment (intervention, screening, diagnosis); (2) patient characteristics (population, settings, functional ability, etc.); (3) psychometric properties of the tool (test–retest reliability, inter-rater reliability, responsiveness, floor and ceiling effects, etc.); (4) applicability of the tool in clinical settings (overall cost, time required for the examination, level of training, equipment, patient acceptance, etc.); (5) prognostic reliability for relevant clinical outcomes. Based on these criteria and the available evidence, the expert group advises the use of grip strength to measure muscle strength and the use of 4-m gait speed or the Short Physical Performance Battery test to measure physical performance in daily practice. The tools proposed are relevant for the assessment of muscle weakness and physical performance. Subjects with low values should receive additional diagnostic workups to achieve a full diagnosis of the underlying condition responsible (sarcopenia, frailty or other).

Keywords Muscle function · Muscle strength · Physical performance · Daily practice · Sarcopenia

Introduction

In 2015, WHO proposed a new vision for healthy ageing in its *World Report on Ageing and Health*. The report defines the goal of healthy ageing as helping people in “developing and maintaining the functional ability that enables well-being”. Functional ability is determined by the intrinsic capacity of

the individual, relevant environmental characteristics and the interactions between them. Although actions at all levels of the society are vital to foster healthy ageing, realigning health systems towards building and maintaining the intrinsic capacity and functional abilities of adults in the second half of life has been identified as an immediate priority.

Significant loss of intrinsic capacity (functioning) in older adults is characterized by the manifestation of common problems, such as difficulties walking at usual pace, loss of muscle mass and strength and mobility impairments

✉ Olivier Bruyère
olivier.bruyere@ulg.ac.be

Extended author information available on the last page of the article

[1]. At present, most health care professionals lack guidance or training to recognize and manage declines in physical capacities in older age. However, research evidence suggests that even in less resourced health care settings, health care professionals can be trained to detect declines in physical capacities (clinically expressed as mobility impairments) and deliver effective interventions to prevent and delay progression [2].

In 2017 WHO launched the Guidelines on community level interventions to manage declines in intrinsic capacity. The primary audience for these WHO-ICOPE guidelines on community-level interventions is health care providers working in primary and secondary care settings (<http://www.who.int/ageing/publications/guidelines-icope/en/>). Guidance on the assessment of physical performance in daily clinical practice can be helpful for geriatricians and for physicians working with frail elderly subjects.

Considering the large number of tools available to measure physical function in older adults, including muscle mass and strength, WHO requested support from the World Health Organization, Collaborating Center for Public Health Aspects of Musculoskeletal Health and Aging to undertake an initial review and consultation towards the identification of the most appropriated tools. As measures of muscle strength and physical performance are increasingly used for research and practice, a review and identification of appropriated tools to evaluate them is timely. The findings presented in this paper are result of a fruitful collaboration between experts of the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the World Health Organization, Collaborating Center for Public Health Aspects of Musculoskeletal Health and Aging. In this paper, an overview of different methods available and applicable in clinical settings is proposed. The WHO Strategy aims to encourage clinicians and general practitioners (GPs) working in primary health care settings to timely identify declines in muscle strength and physical performance on in older adults. This process should be followed by a precise diagnosis and the start of interventions in affected individuals [3]. The respective measures can be categorized as primary or secondary prevention in older persons that are at high-risk of future care dependence.

It is important to clarify the terms muscle function and physical performance. Muscle function first is underlined by the three concepts of muscle strength, muscle power and muscle endurance. Muscle strength refers to “*the amount of force a muscle can produce with a single maximal effort*”. Muscle strength should be differentiated to muscle power which is defined by “*the ability to exert a maximal force in as short a time as possible, as in accelerating, jumping and throwing implements*” and to muscle endurance which is defined as “*the ability of muscles to exert force against resistance over a sustained period of time*”. The concept of physical performance has become confusing along time and merits a new vision. When first defined, physical performance measures were used to objectively assess how an individual performed different activities of daily living (ADLs) or physical tasks in the clinic, as opposed to ADL scales based on asking questions about the ability to perform such tasks. However, in the last decades, this old concept has gradually changed, and measures of physical performance are now mostly related to ambulation and transfers. In fact, some measures of physical performance (i.e. gait speed) have become part of the definitions of frailty and sarcopenia. No definition of this updated concept of physical performance exists. After long discussion, the group developed a definition of physical performance definition which describes it as “*an objectively measured whole body function related with mobility*”. Physical performance goes beyond muscle function measures, as it involves many other body organs and systems (bones, balance and other neurological inputs, cardiovascular aspects, motivation...) being a multidimensional concept (Fig. 1). It is thus linked to the WHO concept of intrinsic capacity in its aspects related with mobility. Impairments in physical performance may be evident way before disability starts (as defined by inability to perform ADLs), so it allows for detection of vulnerability in the first steps of the disabling cascade. Physical performance is different from physical capacity where the notion of capacity to achieve is primordial.

In order to make this paper easy to consult for clinicians, the authors chose to focus on tests that clinicians are susceptible to use in their daily practice instead of presenting a review of all available tests in the literature, including some not suitable for routine assessments. The instruments included in this paper were a consensus decision of the

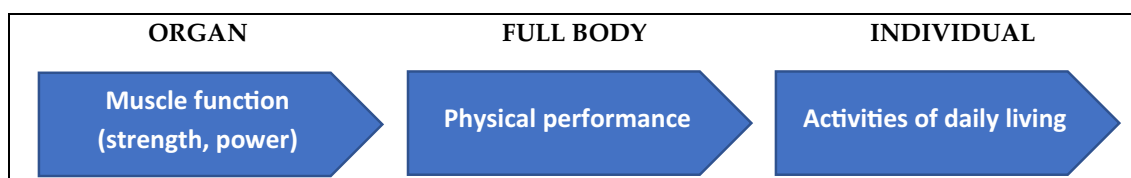


Fig. 1 Hierarchy of loss of physical function

authors after reviewing all the evidence on the tools available. They took into account multiple criteria, such as the availability of the tool, the presence of validated protocols for measurement and cut-offs, etc. Ultimately, they agreed to present the most applicable tests for clinical assessments.

Methods

As in previous initiatives [4–6], the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO) working group on frailty and sarcopenia consists of clinical scientists and experts in the field of musculoskeletal diseases. Different members of the ESCEO working group were asked to review the literature on (1) tools to assess muscle strength and power in daily practice—strengths and weaknesses (JB); (2) objective assessment of muscle strength and power (CS); (3) tools to assess physical performance in daily practice—strengths and weaknesses (ACJ); and (4) can reference standards in the assessment of muscle strength, muscle power and physical performance be developed? (RF). Randomized controlled trials, prospective studies, systematic reviews and meta-analyses published before September 2017 were sought on PubMed and Scopus using the following search terms: (1) Mesh terms searched: Aged, Muscle skeletal, Muscle strength, Physical functional performance, Gait, Walk test, Walking, Sarcopenia, Reference standards, Reference Values; (2) Additional terms searched in either title, abstract or keywords: muscle power, physical function, endurance test. Additional studies were identified through the expertise of the members of the working group but also by a manual search of the reference sections of relevant articles and existing reviews. Each member prepared a list of the most important papers based on their review of the literature and then created a set of preliminary recommendations. The subsequent step was a face-to-face meeting of the whole group to make amendments and discuss results. Shared conclusions were reached and finalized during the creation of the current manuscript whose drafts were circulated by mail (Fig. 2).

Measurement of Muscle Function

Muscle Strength

There are few well-validated techniques to measure muscle strength. Among them, strength of both upper limbs and lower limbs can be assessed; both of them having been shown to be highly correlated [7–10]. In clinical settings and for the diagnosis of sarcopenia and frailty, grip strength is the measure of choice for the assessment of overall muscle strength, as it has been shown to be a surrogate for lower

extremity muscle strength and as it is easier to measure [5, 10, 11]. Indeed, grip strength measurement only requires the holding of a handheld dynamometer, has standardized protocols of measurement and has robust and validated cut-offs values available, which is not the case for all lower limb muscle strength values. The common use of handgrip strength in clinical daily practice has been highlighted recently in a survey based on the experience of 255 clinicians from 55 countries showing a large part of the sample using handgrip strength (66.4% of the sample) compared to lower limb muscle assessments (e.g. leg press 24.2%, chest press 9.4%, isokinetic parameters 7.4% of the sample) [12]. For these reasons, this consensus paper focused only on grip strength as the measure of muscle strength, following the recommendation of the experts involved in this paper.

Grip Strength

Definition

Handgrip strength is the most widely used method for the measurement of muscle strength. It is recognized to be easily applicable both in research and in clinic settings [5]. Handgrip strength is usually measured during muscular isotonic contraction.

How to Measure Grip Strength in Clinical Practice?

- *Procedure* The measurement is easy to perform, the device is portable, with an acceptable cost, and does not require a specialist trained user. Because of the ease of its application, grip strength measurement can be used in clinical practice, and thus, can be applied in a large sample of older adults, symptomatic or asymptomatic, to identify those with low muscle strength. It is recommended to use standardized measurement protocols such as the Southampton protocol, proposed by Roberts et al. [13, 14] for the measurement of grip strength. Briefly, standardized conditions for the test include seating the subject in a standard chair with forearms resting flat on the chair arms. The testing nurse or physician should demonstrate the use of the dynamometer and show that gripping very tightly registers the best score. Six measures should be taken, 3 with each arm. Ideally, the patients should be encouraged to squeeze as hard and as tightly as possible during 3–5 s during each of the 6 trials; usually the highest reading of the 6 measurements is reported as the final result.
- *Time of administration* 5 min.
- *Equipment* A well-calibrated handheld dynamometer. Different features exist such as hydraulic dynamometer [e.g. Jamar, which is usually considered as the gold

reference for this measurement, with units in kilograms (kg) or pounds (lbf), pneumatic [e.g. Martin vigorimeter with units in millimetres of mercury (mmHg) or pounds per square inch (psi)] which measure grip pressure, mechanical [e.g. Harpenden dynamometer with units in kg or lbf) and strain dynamometer (with units in Newtons of force (N)]. Dynamometers have to be calibrated appropriately by manufacturer prior using it [15].

Performance Characteristics

- *Highly feasible* [16].
- *Test–retest reliability* in older adults is good (ICC ≥ 0.85) [17, 18].
- *Inter-rater reliability* in older adult is excellent (ICC 0.95–0.98) [16, 19].
- *Responsiveness* One study proposed a minimal change of 6 kg (13.2 lb) to be considered as significant [20]. Data on sensitivity of change in grip strength to interventions are still rather limited and inconsistent [18, 21].
- *Floor effects* For patients with upper extremity impairment and/or affected by rheumatoid arthritis, hand osteoarthritis or carpal tunnel syndrome, grip strength measure may not be an accurate reflection of muscle strength and may lead to underestimations. The design of the Jamar dynamometer may be the reason for this [22] and a pneumatic dynamometer may be a good alternative for these patients. With the Martin vigorimeter for example, patients try to squeeze rubber balls (available in three sizes) following the same protocol as that described for the Jamar dynamometer. However, the comparison between these two devices is limited given the different unit of measure provided by them.
- *Note* that the absolute values and the precision of grip strength measurement are influenced by hand position, hand size and dominance [23], body position [24] verbal encouragement [25] and patient motivation. For this reason, the application of a standardized protocol is highly recommended [13].

Reference Range

- Variety of normative data have been proposed, ranging from 16 to 21 kg for women and 26 to 30 kg for men [10, 26–29]. It should also be noted that values adjusted for BMI or height also exists. For example, the EWGSOP [27] proposed general values for grip strength (< 30 kg for men and < 20 kg for women) but also BMI dependent value (e.g. ≤ 29 kg for men with a BMI ≤ 24 kg/m²; ≤ 30 kg for men with a BMI between 24.1 and 28 kg/m²; ≤ 32 kg for men with a BMI > 28 kg/m²; ≤ 17 kg for women with a BMI ≤ 23 kg/m²; ≤ 17.3 kg for women

with a BMI between 23.1 and 26 kg/m²; ≤ 18 kg for women with a BMI between 26.1 and 29 kg/m² and finally ≤ 21 kg for women with a BMI > 29 kg/m²).

- Low grip strength is consistently associated with poor outcomes; care dependence, falls, fractures, mortality [30–35].

Muscle power

Compared to muscle strength, power concerns work rate (work done per unit time). In healthy older people, muscle power declines earlier and faster compared to muscle mass and strength [36]. Leg power has been shown to be highly correlated with physical performance tests such as gait speed, chair stand test and stair-climb time [37], and several comparative studies have found that muscle power is a better predictor of mortality compared to muscle strength [38]. Therefore, muscle power is often proposed as the primary therapeutic target for resistance training interventions in older adults. Muscle power can be assessed across a range of muscle groups, but most often the leg press and knee extension exercises are used to measure muscle power. Maximal strength is quantified through the 1RM (1 repetition maximum resistance), wherein the evaluation is carried out at the highest resistance for which the subject can complete the exercise once. To find the 1RM, the exercise is repeated several times at increasing resistance until failure to complete a single repetition. However, there is a distinct lack of standardization across studies, and different equipment and measurement techniques have been used [39]. The most common muscle power exercises, leg press and knee extension, show good reliability and validity [40]. Currently, a reference range for the different measures of muscle power is not yet available. Because measuring muscle power requires complex and sometimes expensive machines but also because the applicability of this measure is compromised in clinical settings (necessity of training for both clinicians and subjects, time-consuming, no standardized protocols available, etc.) and finally, because standardized cut-off points have not been agreed to define a low muscle power, muscle power assessments can hardly be considered useful in daily practice. For this reason, based on an agreement of the experts involved in this paper, measurement of muscle power will not be exhaustively discussed in the present paper.

Measurement of Physical Performance

Many tests are described in the literature to measure physical performance of older adults. Considering the applicability of these tests in clinical practice but also their performance characteristics, the experts involved in this paper agreed to

focus on some of the available physical performance tests. They therefore chose to present and encourage the measurement of gait speed, chair stand test, short physical performance battery (SPPB) test and the timed up and go (TUG) test.

Gait Speed Test

Two main types of gait speed tests exist; the short-distance walk tests (2.4-m distance, 4-m distance, 6-m distance and 10-m distance) and the long-distance walk tests (400-m walk test and 6-min walk test). Long-distance walk tests require a corridor of at least 20 m as well as a minimum time for execution of 15 min. These tests may be extremely useful in discriminating different categories of risk among older individuals in healthy conditions. Besides measuring physical performance in older adults, these tests also evaluate endurance of the subjects. Nevertheless, short walk tests can be used as surrogates for long-distance walk tests to measure the overall functional status in older adults [41]. Indeed, the 4-m gait speed test, for example, has been shown to be highly predictive of the ability to perform the 400-m walk test in older adults [41, 42]. Taking this into account and because short-walk tests are clearly more readily applicable in clinical settings, the experts of this group agreed to focus on short walk tests for the current review and strongly recommend these tests to assess physical performance of older adults in routine practice.

Short Walk Measures of Usual Gait Speed

Definition

The gait speed measurement is probably one of the most widely used tools in clinical practice for the assessment of physical performance. It is generally measured in a short distance (2.4-m distance, 4-m distance, 6-m and 10-m distance) with the 4-m distance as being the most commonly used short-walk test validated in older adults.

How to Measure Gait Speed Test in Clinical Practice?

- *Procedure* short-distance walk tests are applicable in clinics and in GPs offices but some training for the testing staff is required. This measure is also highly acceptable for participants and health professionals [43, 44]. Some tentative recommendations for a protocol of administration have been proposed following a systematic review [45]: (1) use a static start with timing commencing when the foot touches the floor the first time after the line; (2) usual or comfortable pace to be used as the standard, with fast pace used as appropriate for specific research questions. (3) Walking protocol to be reported in detail

including pace instructions, verbal or other encouragement, and specific timing procedures.

- *Time of administration* 95 ± 20 s.
- *Equipment* 4-m long flat floor devoid of obstacles and a chronograph. This test can therefore be administered in restricted areas.

Performance Characteristics

- *Test–retest reliability* excellent test–retest reliability on 4- and 10-m distance has been shown for healthy older adults (ICC values ranging from 0.96 to 0.98) [46]. Test–retest reliability of gait speed has also been assessed in populations with comorbidities such as in patients with stroke [47], COPD [48], in cardiac rehabilitation [49], etc. It has been demonstrated to be highly reliable in each of these populations.
- *Inter-rater reliability* very strong inter-rater reliability reported in older patients with COPD ICC value of 0.99 (95% CI 0.98–0.99) [48, 50].
- *Responsiveness* the 4-m gait speed is responsive to clinically meaningful changes with 0.05 m/s denoting a small change (i.e. clinically detectable, potentially important) and 0.1 m/s indicating a substantial change (clinically detectable, definitely important) [51].
- *Floor effect* A floor effect is obvious in subjects unable to walk. However, as it has been shown in a systematic review, there is a broad range of people for whom timed walking is a valid and sensitive outcome measure (patients with cancer, neurological problems, osteoarthritis, fractures, etc.) [45]. A ceiling effect has also been reported in certain populations. For example, it is doubtful whether a potential increase of muscle mass will result in an improvement in gait speed in well-functioning subjects with a high baseline walking speed.

Reference Range

- *Timed usual gait* has been shown to be highly predictive for future care dependence [52], for other adverse health events such as severe mobility limitation or mortality [53]. But it has also been demonstrated that poor performance in other tests of lower extremity function (standing balance test and chair rise test) had comparable prognostic value [54].
- *Cut-off < 0.8 m/s for 4-m distance* identifies subjects with poor physical performance [10]. A systematic review including 3 studies with 3261 participants revealed a very high sensitivity for the < 0.8 m/s cut-off for identifying frailty (Se = 0.99), as well as high negative predictive value (NPV = 0.99), but also a moderate specificity (Sp = 0.64) and a low positive predictive value (PPV = 0.26) [55].

- *Cut-off* < 1 m/s for a 6-m distance identifies older persons at high risk of health-related negative events [54].
- *Other cut-offs*, adapted for example, to the height of participants has also been proposed by Fried et al. on a distance of 15-ft (4.572 m) [17, 73].

Chair Stand Test

Definition

The 30-second (30-s) Chair stand test (CST) developed by Rikli and Jones [56] is one of the most important physical performance clinical tests because it measures lower body power, balance and endurance and relates it to the most demanding daily life activities. The 30-s CST has been widely used in many studies not only to evaluate functional fitness levels [57] but also to monitor training [58–60] and rehabilitation [61]. Another version of the chair stand test is also well known (as it is embedded in the SPPB) and consists of recording the amount of time to complete five sit-to-stand manoeuvres. The test has been shown as a predictor of falls but has some limitations. This test has indeed a restricted capacity to assess a wide variation in ability, which is relevant in older people, since some older adults cannot complete the five attempts and are therefore not assigned a score (floor effect). The utility of this test is therefore limited in subjects suffering from moderate to severe mobility limitations. Consequently, for older populations, the literature favours time-based protocols such as the 30-s chair stand test.

How to Measure the 30-s CST Test in Clinical Practice?

- *Procedure*: Classically, the 30-s CST consists of manually counting the number of total sit-stand-sit cycles completed during the 30 s of the test. This test is highly feasible in clinical practice and may therefore be recommended as a measure of physical performance.
- *Time of administration* 1–2 min.
- *Equipment* A chair with a straight back without arm rests and a stopwatch.

Performance Characteristics

- *Test–retest reliability* good test–retest reliability (0.84 for men to 0.92 for women) in healthy older adults [62] as well as excellent reliability for subjects with knee arthroplasty [63], for subjects with mild-to moderate dementia [64] or hospitalized patients with stroke [65].
- *Inter-rater reliability* limited data available but very strong inter-rater reliability on older adults in nursing home with mild to moderate dementia (perfect ICC of 1) [64].

- *Responsiveness* Very limited data available. Only a minimal detectable change (MDC) value of 3.49 in older people with dementia [66] has been defined. Then, an improvement in more than 3.49 sit-stand-sit cycles during the 30-s chair stand test is considered to be a true change in performance with 95% confidence.
- *Floor effect* this test is not impacted by floor effect, as it is the case of the five sit-to-stand test, since subjects unable to perform it are attributed a score of 0. Some authors argue that the 30-s CST protocol makes it possible to assess wider variations in ability levels of individuals compared to the five-time sit-to-stand test as it avoid potential floor effects [67].

Reference Range

- Limited data in literature.
- One author report an association between the test and the risk of falling in a population of older nursing home residents [68].
- Normative values have been proposed for two particular populations: Hong Kong older adults and US older adults. Normative value for Hong Kong older adults 70–74 years for example is a mean of 10.1 ± 3.8 stands during 30-s and 13 stands during 30 s for US norms [69].

Short Physical Performance Battery (SPPB)

Definition

The short physical performance battery (SPPB) originally developed at the National Institute on Aging for use in the Established Population for the Epidemiologic Studies of the Elderly (EPESE) [70] is the most widely used physical performance test battery that has been applied in clinical and research settings including randomized controlled trials. This battery of tests has been validated in large-scale epidemiological studies and evaluates lower extremity functional performance using timed measures of standing balance, gait speed and lower extremity strength. Moreover, both the global score and its individual elements may be analysed separately in different clinical or research settings.

How to Measure SPPB Test in Clinical Daily Practice?

- *Time of administration* 10 min.
- *Equipment* a 4-m track, ground marks, a chronometer and straight-backed chair.
- *Procedure* applicable in research and in clinics as well as in GPs offices but training is required. The full detailed instructions can be downloaded for free from the web (<https://sppbguide.com/>) but are summarized here below. This test is longer to apply compared to chair rising test

alone or to gait speed test alone but is nevertheless feasible and can be recommended as a screening test for poor physical performance and risk of sarcopenia.

The three following tests should be administrated successively in the same order as they are presented below:

1. Balance test

For the balance test, the subject is asked to hold three increasingly challenging standing positions for 10 s each: (1) a side-by-side position; (2) semi-tandem position (the heel of one foot beside the big toe of the other foot); (3) tandem position (the heel of one foot in front of and touching the toes of the other foot).

2. Walking speed test

For the walking speed test, the subject is asked to walk at his/her usual pace over a 4-m course (originally 8 feet or 2.4 m). He/she is instructed to stand with both feet touching the starting line and to start walking after a verbal command. The subject is allowed to use walking aids (cane, walker, or other walking aid) if necessary, but no assistance by another person can be provided. Timing begins when the starting command is given, and time in seconds needed to complete the entire distance is recorded. The faster of two walks is usually considered in computing the SPPB score.

3. Repeated chair stands test

The repeated chair stands test is performed using a straight-backed chair, placed with its back against a wall. The subject is first asked to stand from a sitting position without using their arms. If he/she is able to perform the task, he/she is then asked to stand up and sit down five times, as quickly as possible with arms folded across their chests. The time to complete five stands is recorded.

These three physical performance subtasks are used to calculate a summary score. The score of each of the three tests ranges from 0 to 4, where 4 indicates the best result and 0 the worst result. Therefore, a summary score ranging from 0 (worst performers) to 12 (best performers) is calculated by adding the categorical results derived from the three timed physical performance subtasks.

Performance Characteristics

- *Test–retest reliability* has been shown to be good to excellent (ICC ranging from 0.83 to 0.92 for measures made 1 week apart) [71–73]. The reproducibility of the SPPB, already very good, can nevertheless be enhanced through the use of standardized equipment and an appropriate standard operating procedure.
- *Inter-rater reliability* has been shown to be excellent (ICC 0.91) among acutely admitted older medical patients [16]. Data on healthy populations are limited.
- *Responsiveness* the SPPB is responsive to clinically meaningful changes [71, 74] with 0.5 points denoting a small change (i.e. clinically detectable, potentially important) and 1 point denoting a substantial change (clinically detectable, definitely important) [51].
- *Ceiling effects* it may have ceiling effects for high functioning and very fit older adults (who will score 12 points). For research, a more challenging SPPB has been proposed, for example, in the Health, Aging and Body Composition (Health ABC) Study, with time to walk 400 m measured instead of 4-m distance, and by extending the times from 10 to 30 s for the three standard balance tests [54]. This approach is relevant to assess the wide range of functioning at baseline among fit older participant of cohort studies but is unlikely to be useful in clinical practice.
- *Floor effects* it is still not clear how to score a subject unable to walk and therefore unable to perform the SPPB test correctly. Floor effects of 0 points may be observed in this specific case.

Reference Range

- *Scores obtained on a 12-point summary scale* provides a gradient of functional decline that has been shown to be valid and reliable in predicting the future risk of mobility impairment, care dependence, institutionalization, hospital admission and mortality [71, 75, 76].
- *Cut-off point of ≤ 10 points* has been shown to be a strong predictor of the loss of ability to walk 400 meters with a sensitivity of 0.69 and a specificity of 0.84 [42].
- *Cut-off point of ≤ 8 points* has been showed to be associated with mobility-related disability [52].
- *A very low SPPB test (0–6 points)* has been shown to be associated with increased risk of death [75, 77].

Timed-Get-Up-and-Go Test

Definition

The timed-get-up-and-go (TUG) test is a physical performance measure that has been mainly used to assess gait and dynamic balance [78]. It is a single test measuring the time a person takes to complete a complex series of different motor tasks.

How to Measure the TUG Test in Clinical Daily Practice?

- *Procedure* Applicable in clinical settings and GPs offices and on different populations (with frailty, Parkinson's

disease, cognitive impairment, recent joint surgery, osteoarthritis, etc.); little training is necessary.

- TUG requires the subject to stand up from a chair, walk three meters, turn around, return and sit down again. The stopwatch is started on the word “go” and stopped at the moment the subject sits down. The person wears regular footwear, uses his/her customary walking aid and walks at her/his usual pace. No physical assistance is given. A practice trial is given first, and the average time of two consecutive trials is recorded.
- Complete instructions can be downloaded in the following website: http://foxrehab.org/uploads/pdf/2008_AssitedLivingConsult_TUGTest.pdf. A video provided by the Center for Disease Control and Prevention (CDC) is also available on the following website: https://www.youtube.com/watch?v=BA7Y_oLEIGY.
- *Time of administration* 2–3 min.
- *Equipment* Chair with armrest, 3-m track, ground marks, stopwatch.

Performance Characteristics

- *Test–retest reliability* moderate (ICC 0.53) [79] to good test–retest reliability (ICC 0.97–0.99) [78, 80, 81] across studies.
- *Inter-rater reliability* excellent in the older population (ICC = 0.99 [78], ICC = 0.98 [82]).
- *Responsiveness* not well defined in older populations. A reduction in time greater than or equal to 0.8, 1.4 and 1.2 s on the TUG has been determined to be the Minimal Important Clinical Change score (MCID) for patients suffering from hip osteoarthritis [83].
- *Floor and ceiling effects* the TUG does not suffer from ceiling effects [84]. Floor effects are observed, as for each test involving a person’s walking ability.

Reference Range

- The time to perform the task is compared to normative values for age, gender, and research-based guidelines that measure increased risk of falls and functional decline. The performance of the patient can also be summarized into a five-point scaled score [85].
- A TUG score of 14 s is sensitive (87%) and specific (87%) for identifying older individuals who are at risk for falls [82].
- In the systematic review of Clegg et al. [55], assessing, among others, the accuracy of the TUG test for identifying frailty, a sensitivity of 0.93 has been shown for the cut-off of > 10 points (with a negative predictive value of 0.99), with a specificity of 0.62 (with a positive predictive value of only 0.17).

Practice Recommendations

After a review of the literature and discussion during the working group meeting, the following statements and recommendations on the use of measures of muscle strength and physical performance are proposed:

1. Muscle strength and physical performance measures are rarely assessed in daily practice compared to other clinical or biochemical parameters. However, limited strength and performance result in physical limitations which are strong predictors of adverse negative health outcomes such as care dependence, falls, fractures, hospitalization and death. Therefore, the experts involved in this position paper strongly encourage clinicians to routinely assess strength and physical performance in older adults. Unfortunately, data do not yet reveal what is the ideal frequency of repeated measures. Evidence is still needed in this regard.
2. Measurements of muscle strength and function are facilitated by their non-invasiveness and their effective limited time for application.
3. Many different tests to measure muscle strength and physical performances have been described, although not all have the same level of evidence or are easy to use in routine clinical practices. The experts involved in this paper summarized the most widely applicable tests in daily practice. For measuring muscle strength, the experts advise the use of a handheld dynamometer to assess handgrip strength. For measuring physical performance, the experts advise the assessment of 4-m gait speed, Short Physical Performance Battery test or Timed Up and Go test.
4. The choice of the tool to use should be guided according to different parameters:
 - a. The purpose of the assessment should be essential in the choice of the clinician. For an intervention, it is more important to have a tool with a responsiveness to test the impact of an intervention on muscle strength and function. *Based on the experts’ review of the literature, only grip strength, SPPB test and gait speed show a sufficient responsiveness in general older population. Good responsiveness of other tools has also been shown in specific population (e.g. population suffering from dementia, COPD population, etc.).* For a practical purpose, it is essential to have a cheap and easy-to-administer tool applicable to the majority of older adults. The presence of validated normative values for a large group of patients, and from different countries, is also essential. *Based on the experts’ review of the literature, validated normative values are mainly available for grip strength, gait speed and SPPB*

test. For diagnosis purposes, the objective is somewhat different and requires therefore a tool with robust validity and reliability. *Based on the experts' review of the literature, all the tools presented in this paper show satisfactory validity and reliability.*

- b. The population's characteristics will also determine the choice of test. Given the heterogeneity of the population in daily practice, it is important to choose a tool that is highly practicable in various patients' groups. For older people with significant losses of capacity or ability or robust, clinicians should be aware of the presence of floor and ceiling effects. *Based on the experts' review of the literature, each test involving a person's walking ability is prone to floor effects. It is also the case for grip strength, with limitations particularly in some patients with local disease, for example advanced arthritis. Some adaptations of the tests have been proposed in order to characterize the wide range of functioning in these patients.*
 - c. Applicability in clinical settings (cost, required time for the examination, necessary of training, complex equipment, etc.). *Based on the experts' review of the literature, the two most applicable tools in the clinical settings are grip strength measured by a handheld dynamometer and the 4-m gait speed. Both methods do not require a complex and expensive equipment, are cheap, quick and easy to administer, do not require a special training of clinicians and, for the grip strength measurement mainly, standardized validated protocols of assessment are available.*
 - d. The performance characteristics of the tools (test-retest reliability, inter-rater reliability, responsiveness, floor and ceiling effects, etc.). *Based on the experts' review of the literature, all the proposed tools seem reliable but only some of them have already been evaluated for responsiveness. The SPPB test and the 4-m gait speed test seem to be the most robust tools in terms of responsiveness.*
 - e. The prognostic values of the tests for adverse clinical outcomes. Numerous normative data and cut-off points for predicting outcomes have been proposed for the different tools available, both in healthy older population and in specific comorbidities. *Based on the experts' review of the literature, currently, there is substantial evidence for threshold values for handgrip strength, SPPB and 4-m gait speed as being strong predictors of adverse negative health outcomes.*
5. Based on the available evidences, the applicability of the tools in clinical practice, the required time for the

test, the required equipment, the performance characteristics of the tool but also the availability of robust cut-off points, the experts advised grip strength to measure muscle strength and 4-m gait speed or SPPB test to measure physical performance in daily practice. Subjects with low strength or performance should receive an additional diagnostic workup to achieve a full diagnosis of the condition that causes such problems (sarcopenia, frailty or other problems).

Acknowledgement Authors would like to express their gratitude to Dr. Amuthavalli Thiyagarajan Jotheeswaran for his valuable help with this manuscript.

Authors' Contributions JYR organized the meeting. ACJ, JB, CS and RF performed the literature review. CB has drafted the manuscript. All authors have taken part in the discussion and meeting and have critically revised, and given inputs to the final manuscript.

Compliance with Ethical Standards

Conflict of interest Whereas this paper reflects a collaboration between ESCEO and the WHO Collaborating Centre for Public Health Aspects of Musculoskeletal Health and Aging, in Liège, it cannot be seen as an official position of WHO as a global organization.

Dr. Rolland reports grants from Biophytis, Novartis outside the submitted work. Dr. Juergen M. reports grants and personal fees from Nestlé, grants and personal fees from Danone Nutricia, personal fees from Pfizer, personal fees from Novartis, personal fees from Bayer, outside the submitted work. Dr. Sieber reports personal fees from Abbott, personal fees from Braun, personal fees from Danone, personal fees from Fresenius, personal fees from Nutricia, personal fees from AMGEN, personal fees from Berlin-Chemie, personal fees from MSD, personal fees from Novartis, personal fees from Roche, personal fees from Servier, personal fees from Vifor, personal fees from Nestle, outside the submitted work. Professor Cooper reports personal fees from Alliance for Better Bone Health, Amgen, Eli Lilly, GSK, Medtronic, Merck, Novartis, Pfizer, Roche, Servier, Takeda and UCB. Dr. Bruyère reports grants from Biophytis, IBSA, MEDA, Servier, SMB, Theramex, outside the submitted work. Dr. Cherubini reports personal fees from Nestle outside the submitted work. Dr. Kanis reports grants from Amgen, grants from Lilly, grants from Radius Health, grants from UCB, outside the submitted work. Dr. Maggi reports grants from GSK, grants from Merck, grants from Pfizer, grants from Takeda outside the submitted work. Dr. Rizzoli reports personal fees from Nestlé, Danone, TEVA/Theramex, Radius Health, Sandoz, Effryx, Pfizer outside the submitted work. Dr. Fielding reports grants from National Institute of Health USA, during the conduct of the study; grants, personal fees and non-financial support from Axcella Health, grants and personal fees from Biophytis, grants and personal fees from Astellas, personal fees from Cytokinetics, personal fees from GSK, personal fees from Amazentis, personal fees from Nestle, outside the submitted work. Charlotte Beaudart, Nasser Al-Daghri, Islene Araujo de Carvalho, Ivan Bautmans, Roberto Bernabei, Matteo Cesari, Bess Dawson-Hughes, Jean-Marc Kaufman, Francesco Landi, Eugene McCloskey, Jean Petermans, Leocadio Rodriguez Mañas, Jean-Yves Reginster, Regina Roller-Wirnsberger, Laura A. Schaap, Daniel Uebelhart have nothing to disclose.

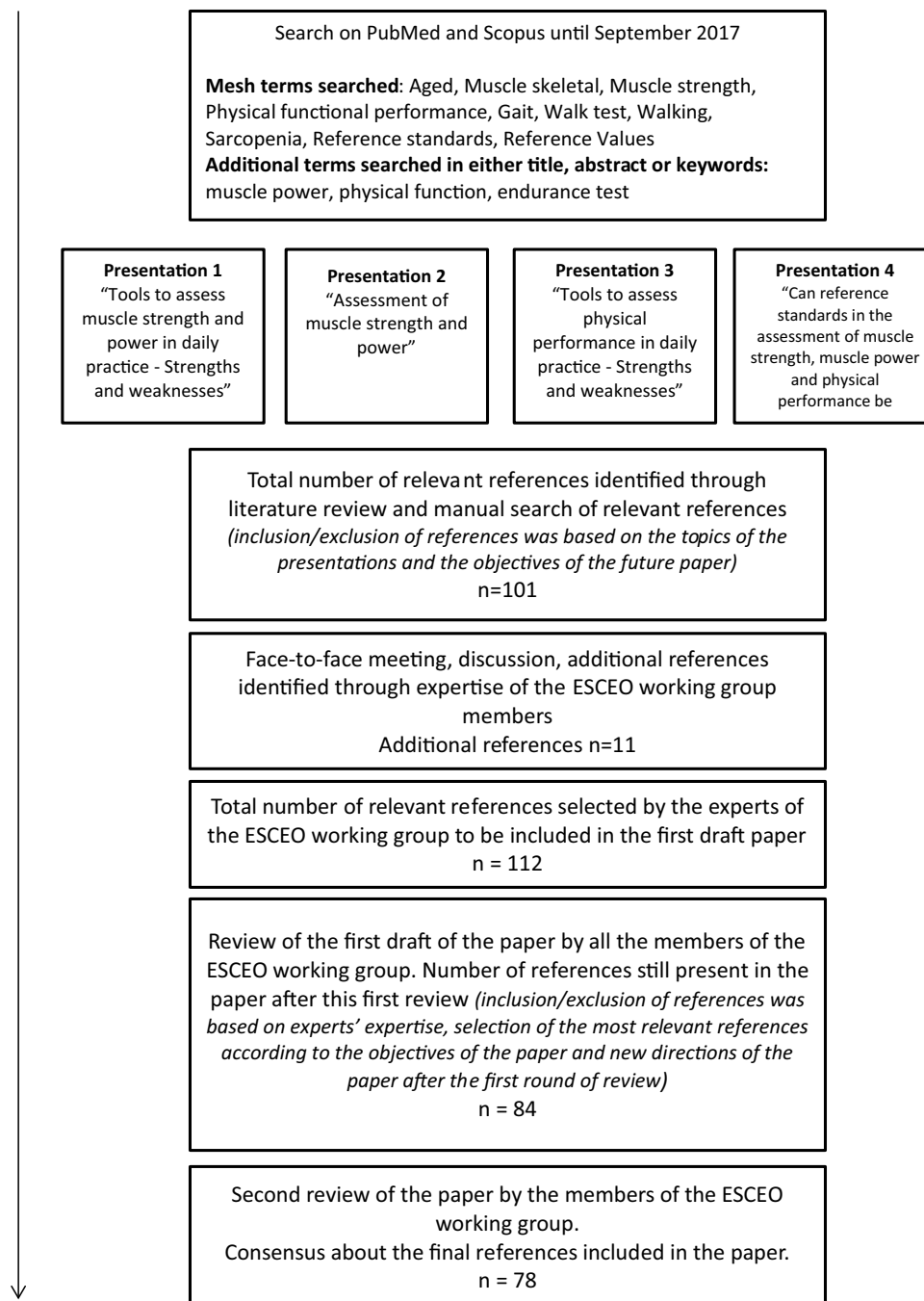
Human and Animal Rights and Informed Consent This article does not contain any studies with human/animal participants performed by any of the authors. Therefore, no consent form was needed.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1

See Fig. 2.

Fig. 2 Flowchart of the search strategy



References

- National Collaborating Centre for Acute Care (2006) Nutrition support in adults Oral nutrition support, enteral tube feeding and parenteral nutrition. <https://www.nice.org.uk/guidance/cg32/evidence/full-guideline-194889853>
- Araujo de Carvalho I, Epping-Jordan J, Pot AM et al (2017) Organizing integrated health-care services to meet older people's needs. *Bull World Health Organ* 95:756–763. <https://doi.org/10.2471/BLT.16.187617>
- Denison HJ, Cooper C, Sayer AA, Robinson SM (2015) Prevention and optimal management of sarcopenia: a review of combined exercise and nutrition interventions to improve muscle outcomes in older people. *Clin Interv Aging* 10:859–869. <https://doi.org/10.2147/CIA.S55842>
- Reginster J-Y, Cooper C, Rizzoli R et al (2016) Recommendations for the conduct of clinical trials for drugs to treat or prevent sarcopenia. *Aging Clin Exp Res* 28:47–58. <https://doi.org/10.1007/s40520-015-0517-y>
- Beaudart C, McCloskey E, Bruyère O et al (2016) Sarcopenia in daily practice: assessment and management. *BMC Geriatr* 16:170. <https://doi.org/10.1186/s12877-016-0349-4>
- Buckinx F, Landi F, Cesari M et al (2018) Pitfalls in the measurement of muscle mass: a need for a reference standard. *J Cachexia Sarcopenia Muscle* 9:269–278. <https://doi.org/10.1002/jcsm.12268>
- Bohannon RW, Magasi SR, Bubela DJ et al (2012) Grip and knee extension muscle strength reflect a common construct among adults. *Muscle Nerve* 46:555–558. <https://doi.org/10.1002/mus.23350>
- Aadahl M, Beyer N, Linneberg A et al (2011) Grip strength and lower limb extension power in 19–72-year-old Danish men and women: the Health 2006 study. *BMJ Open* 1:e000192. <https://doi.org/10.1136/bmjopen-2011-000192>
- Cooper C, Fielding R, Visser M et al (2013) Tools in the assessment of sarcopenia. *Calcif Tissue Int* 93:201–210. <https://doi.org/10.1007/s00223-013-9757-z>
- Lauretani F, Russo CR, Bandinelli S et al (2003) Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. *J Appl Physiol* 95:1851–1860. <https://doi.org/10.1152/jappphysiol.00246.2003>
- Fried LP, Tangen CM, Walston J et al (2001) Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 56:M146–M156
- Bruyère O, Beaudart C, Reginster J-Y et al (2015) Assessment of muscle mass, muscle strength and physical performance in clinical practice: an international survey. *Eur Geriatr Med* 7:243–246
- Roberts HC, Denison HJ, Martin HJ et al (2011) A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age Ageing* 40:423–429. <https://doi.org/10.1093/ageing/afr051>
- Schaap LA, Fox B, Henwood T et al (2016) Grip strength measurement: towards a standardized approach in sarcopenia research and practice. *Eur Geriatr Med*. <https://doi.org/10.1016/j.eurger.2015.11.012>
- Fess EE (1987) A method for checking Jamar dynamometer calibration. *J Hand Ther* 1:28–32. [https://doi.org/10.1016/S0894-1130\(87\)80009-1](https://doi.org/10.1016/S0894-1130(87)80009-1)
- Bodilsen AC, Juul-Larsen HG, Petersen J et al (2015) Feasibility and inter-rater reliability of physical performance measures in acutely admitted older medical patients. *PLoS ONE* 10:e0118248. <https://doi.org/10.1371/journal.pone.0118248>
- Wang C-Y, Chen L-Y (2010) Grip strength in older adults: test-retest reliability and cutoff for subjective weakness of using the hands in heavy tasks. *Arch Phys Med Rehabil* 91:1747–1751. <https://doi.org/10.1016/j.apmr.2010.07.225>
- Bohannon RW (2017) Test-retest reliability of measurements of hand-grip strength obtained by dynamometry from older adults: a systematic review of research in the PubMed database. *J frailty aging* 6:83–87. <https://doi.org/10.14283/jfa.2017.8>
- Peolsson A, Hedlund R, Oberg B (2001) Intra- and inter-tester reliability and reference values for hand strength. *J Rehabil Med* 33:36–41
- Nitschke JE, McMeeken JM, Burry HC, Matyas TA (1999) When is a change a genuine change? A clinically meaningful interpretation of grip strength measurements in healthy and disabled women. *J Hand Ther* 12:25–30
- Santanasto AJ, Glynn NW, Lovato LC et al (2017) Effect of physical activity versus health education on physical function, grip strength and mobility. *J Am Geriatr Soc* 65:1427–1433. <https://doi.org/10.1111/jgs.14804>
- Bean JF, Kiely DK, Herman S et al (2002) The relationship between leg power and physical performance in mobility-limited older people. *J Am Geriatr Soc* 50:461–467
- Bohannon RW (2003) Grip strength: a summary of studies comparing dominant and nondominant limb measurements. *Percept Mot Skills* 96:728–730. <https://doi.org/10.2466/pms.2003.96.3.728>
- Hillman T, Nunes Q, Hornby S et al (2005) A practical posture for hand grip dynamometry in the clinical setting. *Clin Nutr* 24:224–228. <https://doi.org/10.1016/j.clnu.2004.09.013>
- Jung M-C, Hallbeck MS (1999) The effects of instruction, verbal encouragement, and visual feedback on static handgrip strength. *Proc Hum Factors Ergon Soc Annu Meet* 43:703–707. <https://doi.org/10.1177/154193129904301215>
- Dodds RM, Syddall HHE, Cooper R et al (2014) Grip strength across the life course: normative data from twelve British studies. *PLoS ONE* 9:e113637. <https://doi.org/10.1371/journal.pone.0113637>
- Cruz-Jentoft AJ, Baeyens JP, Bauer JM et al (2010) Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on Sarcopenia in Older People. *Age Ageing* 39:412–423. <https://doi.org/10.1093/ageing/afq034>
- Anjum SN, Choudary P, Dimri R, Ankarath S (2012) Comparative evaluation of grip and pinch strength in an Asian and European population. *Hand Ther* 17:11–14. <https://doi.org/10.1258/ht.2011.011023>
- Studenski SA, Peters KW, Alley DE et al (2014) The FNIH sarcopenia project: rationale, study description, conference recommendations, and final estimates. *J Gerontol A Biol Sci Med Sci* 69:547–558. <https://doi.org/10.1093/gerona/glu010>
- Wu Y, Wang W, Liu T, Zhang D (2017) Association of grip strength with risk of all-cause mortality, cardiovascular diseases, and cancer in community-dwelling populations: a meta-analysis of prospective cohort studies. *J Am Med Dir Assoc* 18:551.e17–551.e35. <https://doi.org/10.1016/j.jamda.2017.03.011>
- Sallinen J, Stenholm S, Rantanen T et al (2010) Hand-grip strength cut points to screen older persons at risk for mobility limitation. *J Am Geriatr Soc* 58:1721–1726. <https://doi.org/10.1111/j.1532-5415.2010.03035.x>
- Cheung C-L, Tan KCB, Bow CH et al (2012) Low handgrip strength is a predictor of osteoporotic fractures: cross-sectional and prospective evidence from the Hong Kong Osteoporosis Study. *Age (Dordr)* 34:1239–1248. <https://doi.org/10.1007/s11357-011-9297-2>
- Giampaoli S, Ferrucci L, Cecchi F et al (1999) Hand-grip strength predicts incident disability in non-disabled older men. *Age Ageing* 28:283–288
- Rantanen T, Guralnik JM, Foley D et al (1999) Midlife hand grip strength as a predictor of old age disability. *JAMA* 281:558–560

35. Leong DP, Teo KK, Rangarajan S et al (2015) Prognostic value of grip strength: findings from the Prospective Urban Rural Epidemiology (PURE) study. *Lancet* (London, England) 386:266–273. [https://doi.org/10.1016/S0140-6736\(14\)62000-6](https://doi.org/10.1016/S0140-6736(14)62000-6)
36. Reid KF, Pasha E, Doros G et al (2014) Longitudinal decline of lower extremity muscle power in healthy and mobility-limited older adults: influence of muscle mass, strength, composition, neuromuscular activation and single fiber contractile properties. *Eur J Appl Physiol* 114:29–39. <https://doi.org/10.1007/s00421-013-2728-2>
37. Bassey EJ, Fiatarone MA, O'Neill EF et al (1992) Leg extensor power and functional performance in very old men and women. *Clin Sci (Lond)* 82:321–327
38. Metter EJ, Talbot LA, Schrager M, Conwit RA (2004) Arm-cranking muscle power and arm isometric muscle strength are independent predictors of all-cause mortality in men. *J Appl Physiol* 96:814–821. <https://doi.org/10.1152/jappphysiol.00370.2003>
39. Alcazar J, Guadalupe-Grau A, García-García FJ et al (2017) Skeletal muscle power measurement in older people: a systematic review of testing protocols and adverse events. *J Gerontol Ser A*. <https://doi.org/10.1093/gerona/glx216>
40. Mijnders D, Meijers J, Halfens R et al (2013) 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* 14:170–178
41. Rolland YM, Cesari M, Miller ME et al (2004) Reliability of the 400-M usual-pace walk test as an assessment of mobility limitation in older adults. *J Am Geriatr Soc* 52:972–976. <https://doi.org/10.1111/j.1532-5415.2004.52267.x>
42. Vasunilashorn S, Coppin AK, Patel KV et al (2009) Use of the short physical performance battery score to predict loss of ability to walk 400 meters: analysis from the InCHIANTI Study. *J Gerontol Ser A Biol Sci Med Sci* 64A:223–229. <https://doi.org/10.1093/gerona/gln022>
43. Karpman C, Lebrasseur NK, Depew ZS et al (2014) Measuring gait speed in the out-patient clinic: methodology and feasibility. *Respir Care* 59:531–537. <https://doi.org/10.4187/respcare.02688>
44. Working Group on Functional Outcome Measures for Clinical Trials (2008) Functional outcomes for clinical trials in frail older persons: time to be moving. *J Gerontol A Biol Sci Med Sci* 63:160–164
45. Graham JE, Ostir GV, Fisher SR, Ottenbacher KJ (2008) Assessing walking speed in clinical research: a systematic review. *J Eval Clin Pract* 14:552–562. <https://doi.org/10.1111/j.1365-2753.2007.00917.x>
46. Peters DM, Fritz SL, Krotish DE (2013) Assessing the reliability and validity of a shorter walk test compared with the 10-meter walk test for measurements of gait speed in healthy, older adults. *J Geriatr Phys Ther* 36:24–30. <https://doi.org/10.1519/JPT.0b013e318248e20d>
47. Green J, Forster A, Young J (2002) Reliability of gait speed measured by a timed walking test in patients one year after stroke. *Clin Rehabil* 16:306–314. <https://doi.org/10.1191/0269215502cr4950a>
48. Kon SSC, Patel MS, Canavan JL et al (2013) Reliability and validity of 4-metre gait speed in COPD. *Eur Respir J* 42:333–340. <https://doi.org/10.1183/09031936.00162712>
49. Puthoff ML, Saskowski D (2013) Reliability and responsiveness of gait speed, five times sit to stand, and hand grip strength for patients in cardiac rehabilitation. *Cardiopulm Phys Ther J* 24:31–37
50. Bisca GW, Fava LR, Morita AA et al (2017) 4-meter gait speed test in chronic obstructive pulmonary disease. *J Cardiopulm Rehabil Prev*. <https://doi.org/10.1097/hcr.0000000000000297>
51. Perera S, Mody SH, Woodman RC, Studenski SA (2006) Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc* 54:743–749. <https://doi.org/10.1111/j.1532-5415.2006.00701.x>
52. Guralnik JM, Ferrucci L, Pieper CF et al (2000) Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *J Gerontol A Biol Sci Med Sci* 55:M221–M231
53. Abellan van Kan G, Rolland Y, Andrieu S et al (2009) Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. *J Nutr Health Aging* 13:881–889
54. Cesari M, Kritchevsky SB, Newman AB et al (2009) Added value of physical performance measures in predicting adverse health-related events: results from the health, aging and body composition study. *J Am Geriatr Soc* 57:251–259. <https://doi.org/10.1111/j.1532-5415.2008.02126.x>
55. Clegg A, Rogers L, Young J (2015) Diagnostic test accuracy of simple instruments for identifying frailty in community-dwelling older people: a systematic review. *Age Ageing* 44:148–152. <https://doi.org/10.1093/ageing/afu157>
56. Rikli RE, Jones CJ (1999) Development and validation of a functional fitness test for community-residing older adults. *J Aging Phys Act* 7:129–161. <https://doi.org/10.1123/japa.7.2.129>
57. Okamoto N, Nakatani T, Okamoto Y et al (2010) Increasing the number of steps walked each day improves physical fitness in Japanese community-dwelling adults. *Int J Sports Med* 31:277–282. <https://doi.org/10.1055/s-0029-1234057>
58. Ozkaya GY, Aydin H, Toraman FN et al (2005) Effect of strength and endurance training on cognition in older people. *J Sports Sci Med* 4:300–313
59. Sáez S, De Villarreal E, Requena B, Arampatzi F, Salonikidis K (2010) Effect of plyometric training on chair-rise, jumping and sprinting performance in three age groups of women. *J Sports Med Phys Fitness* 50:166–173
60. Mato L, Wattanathorn J, Muchimapura S et al (2011) *Centella asiatica* improves physical performance and health-related quality of life in healthy elderly volunteer. Evidence-Based Complement Altern Med 2011:1–7. <https://doi.org/10.1093/ecam/nep177>
61. Nyland J, Frost K, Quesada P et al (2007) Self-reported chair-rise ability relates to stair-climbing readiness of total knee arthroplasty patients: a pilot study. *J Rehabil Res Dev* 44:751–759
62. Jones CJ, Rikli RE, Beam WC (1999) A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Res Q Exerc Sport* 70:113–119. <https://doi.org/10.1080/02701367.1999.10608028>
63. Unver B, Kalkan S, Yuksel E et al (2015) Reliability of the 50-foot walk test and 30-sec chair stand test in total knee arthroplasty. *Acta Ortopédica Bras* 23:184–187. <https://doi.org/10.1590/1413-78522015230401018>
64. Telenius EW, Engedal K, Bergland A (2015) Inter-rater reliability of the Berg Balance Scale, 30 s chair stand test and 6 m walking test, and construct validity of the Berg Balance Scale in nursing home residents with mild-to-moderate dementia. *BMJ Open* 5:e008321. <https://doi.org/10.1136/bmjopen-2015-008321>
65. Lyders Johansen K, Derby Stistrup R, Madsen J et al (2015) The timed up and go test and 30 second Chair-Stand Test are reliable for hospitalized patients with stroke. *Physiotherapy* 101:e918. <https://doi.org/10.1016/j.physio.2015.03.1760>
66. Blankevoort CG, van Heuvelen MJG, Scherder EJA (2013) Reliability of six physical performance tests in older people with dementia. *Phys Ther* 93:69–78. <https://doi.org/10.2522/ptj.20110164>
67. Martin DP, Engelberg R, Agel J, Swiontkowski MF (1997) Comparison of the musculoskeletal function assessment questionnaire with the short form-36, the western Ontario and McMaster

- Universities Osteoarthritis Index, and the sickness impact profile health-status measures. *J Bone Joint Surg Am* 79:1323–1335
68. Applebaum EV, Breton D, Feng ZW et al (2017) Modified 30-second Sit to Stand test predicts falls in a cohort of institutionalized older veterans. *PLoS ONE* 12:e0176946. <https://doi.org/10.1371/journal.pone.0176946>
 69. Macfarlane DJ, Chou KL, Cheng YH, Chi I (2006) Validity and normative data for thirty-second chair stand test in elderly community-dwelling Hong Kong Chinese. *Am J Hum Biol* 18:418–421. <https://doi.org/10.1002/ajhb.20503>
 70. Cornoni-Huntley J, Ostfeld AM, Taylor JO et al (1993) Established populations for epidemiologic studies of the elderly: study design and methodology. *Aging (Milano)* 5:27–37
 71. Ostir GV, Volpato S, Fried LP et al (2002) Reliability and sensitivity to change assessed for a summary measure of lower body function: results from the Women's Health and Aging Study. *J Clin Epidemiol* 55:916–921
 72. Gómez JF, Curcio C-L, Alvarado B et al (2013) Validity and reliability of the Short Physical Performance Battery (SPPB): a pilot study on mobility in the Colombian Andes. *Colomb medica (Cali. Colomb)* 44:165–171
 73. Freire AN, Guerra RO, Alvarado B et al (2012) Validity and reliability of the short physical performance battery in two diverse older adult populations in Quebec and Brazil. *J Aging Health* 24:863–878. <https://doi.org/10.1177/0898264312438551>
 74. LIFE Study Investigators, Pahor M, Blair SN et al (2006) Effects of a physical activity intervention on measures of physical performance: results of the lifestyle interventions and independence for Elders Pilot (LIFE-P) study. *J Gerontol A Biol Sci Med Sci* 61:1157–1165
 75. Pavasini R, Guralnik J, Brown JC et al (2016) Short physical performance battery and all-cause mortality: systematic review and meta-analysis. *BMC Med* 14:215. <https://doi.org/10.1186/s12916-016-0763-7>
 76. Volpato S, Cavalieri M, Sioulis F et al (2011) Predictive value of the short physical performance battery following hospitalization in older patients. *J Gerontol A Biol Sci Med Sci* 66:89–96. <https://doi.org/10.1093/gerona/glq167>
 77. Rolland Y, Lauwers-Cances V, Cesari M et al (2006) Physical performance measures as predictors of mortality in a cohort of community-dwelling older French women. *Eur J Epidemiol* 21:113–122. <https://doi.org/10.1007/s10654-005-5458-x>
 78. Podsiadlo D, Richardson S (1991) The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 39:142–148
 79. Rockwood K, Awalt E, Carver D, MacKnight C (2000) Feasibility and measurement properties of the functional reach and the timed up and go tests in the Canadian study of health and aging. *J Gerontol A Biol Sci Med Sci* 55:M70–M73
 80. Schoppen T, Boonstra A, Groothoff JW et al (1999) The Timed ‘up and go’ test: reliability and validity in persons with unilateral lower limb amputation. *Arch Phys Med Rehabil* 80:825–828
 81. Steffen TM, Hacker TA, Mollinger L (2002) Age- and gender-related test performance in community-dwelling elderly people: six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Phys Ther* 82:128–137
 82. Shumway-Cook A, Brauer S, Woollacott M (2000) Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther* 80:896–903
 83. Wright AA, Cook CE, Baxter GD et al (2011) A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. *J Orthop Sport Phys Ther* 41:319–327. <https://doi.org/10.2519/jospt.2011.3515>
 84. Herman T, Giladi N, Hausdorff JM (2011) Properties of the “timed up and go” test: more than meets the eye. *Gerontology* 57:203–210. <https://doi.org/10.1159/000314963>
 85. Mathias S, Nayak US, Isaacs B (1986) Balance in elderly patients: the ‘get-up and go’ test. *Arch Phys Med Rehabil* 67:387–389

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Charlotte Beudart^{1,2} · Yves Rolland³ · Alfonso J. Cruz-Jentoft⁴ · Jürgen M. Bauer⁵ · Cornel Sieber⁶ · Cyrus Cooper^{7,8} · Nasser Al-Daghri⁹ · Islene Araujo de Carvalho¹⁰ · Ivan Bautmans¹¹ · Roberto Bernabei¹² · Olivier Bruyère^{1,2} · Matteo Cesari^{13,14} · Antonio Cherubini¹⁵ · Bess Dawson-Hughes¹⁶ · John A. Kanis^{17,18} · Jean-Marc Kaufman¹⁹ · Francesco Landi¹² · Stefania Maggi²⁰ · Eugene McCloskey^{17,18,19,20,21} · Jean Petermans²² · Leocadio Rodriguez Mañas²³ · Jean-Yves Reginster^{1,2} · Regina Roller-Wirnsberger²⁴ · Laura A. Schaap²⁵ · Daniel Uebelhart²⁶ · René Rizzoli²⁷ · Roger A. Fielding²⁸

¹ Department of Public Health, Epidemiology and Health Economics, University of Liège, Avenue Hippocrate 13, CHU B23, 4000 Liège, Belgium

² WHO Collaborating Centre for Public Health Aspects of Musculo-Skeletal Health and Aging, Liège, Belgium

³ Gérontopôle of Toulouse, University of Toulouse III, CHU Purpan, Toulouse, France

⁴ Servicio de Geriátria, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain

⁵ Center for Geriatric Medicine, University of Heidelberg, Agaplesion Bethanien Hospital, Heidelberg, Germany

⁶ Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Germany

⁷ MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, England, UK

⁸ NIHR Musculoskeletal Biomedical Research Unit, University of Oxford, Oxford, UK

⁹ Prince Mutaib Chair for Biomarkers of Osteoporosis, Biochemistry Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

¹⁰ World Health Organization, Geneva, Switzerland

- ¹¹ Gerontology and Frailty in Ageing Research Department, Vrije Universiteit Brussel (VUB), Brussels, Belgium
- ¹² Department of Geriatrics, Neurosciences and Orthopedics, Catholic University of the Sacred Heart Rome, Milan, Italy
- ¹³ Fondazione IRCCS Ca' Granda–Ospedale Maggiore Policlinico, Milan, Italy
- ¹⁴ Department of Clinical and Community Sciences, University of Milan, Milan, Italy
- ¹⁵ Geriatria, Accettazione Geriatrica e Centro di Ricerca per l'invecchiamento, IRCCS-INRCA, Ancona, Italy
- ¹⁶ Bone Metabolism Laboratory, Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA, USA
- ¹⁷ Mary McKillop Health Institute, Australian Catholic University, Melbourne, Australia
- ¹⁸ Centre for Metabolic Bone Diseases, University of Sheffield, Sheffield, UK
- ¹⁹ Department of Endocrinology, Ghent University Hospital, Ghent, Belgium
- ²⁰ CNR Neuroscience Institute, Aging Branch, Padua, Italy
- ²¹ MRC and Arthritis Research UK Centre for Integrated Research in Musculoskeletal Ageing (CIMA), Sheffield, UK
- ²² Geriatric Department, CHU Sart-Tilman, Bât B35, 4000 Liège, Belgium
- ²³ Department of Geriatrics, Hospital Universitario de Getafe, Madrid, Spain
- ²⁴ Department of Internal Medicine, Medical University of Graz, Graz, Austria
- ²⁵ Department of Health Sciences, Faculty of Science, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
- ²⁶ Centre Hospitalier du Valais Romand, Centre Valaisan de Pneumologie, Crans-Montana, Switzerland
- ²⁷ Service of Bone Diseases, Faculty of Medicine, Geneva University Hospitals, Geneva, Switzerland
- ²⁸ Nutrition, Exercise Physiology and Sarcopenia Laboratory, Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, USA