Experimental Gerontology 145 (2021) 111202

Contents lists available at ScienceDirect

Experimental Gerontology

journal homepage: www.elsevier.com/locate/expgero

Validity, reliability, and measurement error of a sit-to-stand power test in older adults: A pre-registered study

Anoop T. Balachandran [a,*], Andrew D. Vigotsky [b], Norberto Quiles [a], Lidwine B. Mokkink [c], Mark A. Belio [a], Jordan McKenzie Glenn [d,e]

[a] Department of Family, Nutrition, and Exercise Sciences, Queens College, The City University of New York, Flushing, NY, United States of America
[b] Departments of Biomedical Engineering and Statistics, Northwestern University, Evanston, IL, United States of America
[c] Amsterdam University Medical Centers, location VUmc, Department of Epidemiology and Data Science, Amsterdam Public Health research institute, Amsterdam, the Netherlands
[d] Exercise Science Research Center, Department of Health, Human Performance and Recreation, University of Arkansas, Fayetteville, AR, United States of America
[e] Neurotrack Technologies, 399 Bradford St., Redwood City, CA, United States of America

ARTICLE INFO

Keywords:
Elderly
Physical function
Velocity
Aging
Power

ABSTRACT

Objectives: Lower body power declines with age and is associated with decreased physical function in older adults. However, the majority of the tools available to measure power are expensive and require considerable space and expertise to operate. The purpose of this study was to assess the validity, reliability, and measurement error of a sit-to-stand power test (STSp) to assess lower body power.

Methods: 51 community-dwelling adults, 65 years or older, completed a power test using a pneumatic leg press (LP), the Short Physical Performance Battery (SPPB) that includes a test of balance, usual walking speed, and chair stand tests; Timed Up and Go (TUG) test at both usual and fast paces, and Patient-Reported Outcome Measures (PROMs). A two-week test-retest assessed the reliability in 36 participants. The study hypotheses and analysis were pre-registered prior to data collection and statistical analyses were blinded.

Results: The mean age was 71.3 years, with 63% females, and an average SPPB score of 10.6 (median = 12). STSp peak power was strongly correlated with LP (r = 0.90, 95% CI (0.82, 0.94). As hypothesized, the STSp peak power showed similar or higher correlations with physical function tests relative to LP peak power: SPPB (0.41 vs. 0.29), chair stand test (−0.44 vs. -0.35), TUG test at usual pace (−0.37 vs. −0.29) and fast pace (−0.41 vs. −0.34) and balance (0.33 vs. 0.22), but not for mobility (0.34 vs. 0.38) and function (0.41 vs. 0.48) questionnaire. For discriminant validity, as hypothesized, males showed higher STSp peak power compared to females (Δ = 492 W, p < .001, Cohen's d = 2.0). Test-retest assessment yielded an intraclass correlation coefficient of 0.96 and a standard error of measurement of 70.4 W. No adverse events were reported or observed for both tests.

Conclusion: The STSp showed adequate validity and reliability in measuring lower body power in community-dwelling older adults. The test is quick, relatively inexpensive, safe, and portable and thus should be considered for use in aging research.

Aging, even in the absence of overt disease, leads to a gradual decline in muscle mass and strength(Frontera et al., 2008). This gradual decline can result in the loss of physical independence, increased risk of falls, decreased quality of life, increased health care costs, and lowered life expectancy(Fried and Guralnik, 1997; Bergen et al., 2016; Janssen et al., 2002; Lord et al., 1994). Considering the expected rise in the size of the elderly population(Werner, 2010), preserving physical function is a significant public health concern.

Over the past two decades, mechanical power—the rate of mechanical work, or the dot product of force and velocity—has gained prominence as an important determinant of physical function in the aging population. During the aging process, power declines at a faster rate than strength(Reid et al., 2014; Metter et al., 1997)—the ability to exert force—and shows a stronger association with physical function (Bean et al., 2002; Skelton et al., 1994; Foldvari et al., 2000; Sayers et al., 2005) and falls(Whipple et al., 1987; Skelton et al., 2002) than

* Corresponding author at: Family, Nutrition, and Exercise Sciences, Queens College, The City University of New York, Fitzgerald Building, Room 203C, 65-30 Kissena Blvd., Flushing 11367, NY, United States of America.
E-mail address: athozhuthungalba@qc.cuny.edu (A.T. Balachandran).

https://doi.org/10.1016/j.exger.2020.111202
Received 1 August 2020; Received in revised form 23 November 2020; Accepted 8 December 2020
Available online 19 December 2020
0531-5565/

does strength. This observational work is corroborated by interventional studies. Multiple systematic reviews(Byrne et al., 2016; Reid and Fielding, 2012) and meta-analyses(da Rosa et al., 2019; Tschopp et al., 2011) have now shown that power training is more beneficial for improving physical function than conventional resistance training in older adults. Despite the compelling evidence, unlike grip strength, power assessments are often excluded from multi-center intervention trials to improve function or prevent mobility loss(Stathi et al., 2018; Pahor et al., 2018; Landi et al., 2017).

For a measurement tool to be widely adopted in the aging population, it must not only be valid and reliable, but also safe, inexpensive, portable for home use, easy to use, and time-efficient, such as grip strength assessments. With a few excpetions(Alcazar et al., 2018a; Gray and Paulson, 2014), the majority of the tools available to measure lower body power are expensive and require considerable space and expertise to operate. For example, the Nottingham Power Rig(Bassey and Short, 1990), force plate(Lindemann et al., 2003), 3D motion capture(Ford et al., 2007), and isokinetic dynamometry(Suzuki et al., 2001) measurements are only available in research laboratories due to their high cost and complexity. The pneumatic leg press is most commonly used in trials to measure power(Alcazar et al., 2018b), but it is large, expensive, and immobile. Moreover, most of the aforementioned tools measure power in a seated, artificial manner, which could affect their capacity to predict disability or discriminating elderly with functional limitations (Augustsson et al., 1998a).

The sit-to-stand power test (STSp), using a portable linear transducer, is a promising method to assess lower body power that is relatively inexpensive, safe, and portable. It measures power in a functional manner(Liu et al., 2014), which is specific to upright, weight-bearing, everyday functions such as walking or stair climbing. The few studies that have assessed the construct validity of the STSp had several limitations. First, no study has compared the STSp with a standard method to measure power by juxtaposing their associations with physical functional outcomes(Gray and Paulson, 2014; Glenn et al., 2017b; Sherwood et al., 2019; Glenn et al., 2017a). Considering that power is a surrogate outcome and there is no "gold-standard" test for measuring power, it is prudent to compare the associations of current power measures with standard functional outcomes to establish validity, as opposed to using arbitrary effect size thresholds or significance tests. We chose the pneumatic leg press for comparison since it is valid, reliable(Foldvari et al., 2000; Thomas et al., 1996), and is the most commonly used method (63% of power studies) to measure lower body power(Alcazar et al., 2018b). Second, no study has quantified the test-retest reliability and measurement error of the STSp to measure power, which are necessary for evaluating change in response to behavioral/pharmaceutical intervention(De Vet et al., 2011). Third, previous studies did no pre-register their outcomes or hypotheses. Pre-registration separates hypothesis testing (confirmatory) from hypothesis-generating research (exploratory), thus improving research credibility and reproducibility (Scheel et al., 2020; Kaplan and Irvin, 2015). We have registered both the direction and magnitude of our hypotheses.

The purpose of the study is to investigate the construct validity, reliability, and measurement error of the STSp. As stated in the preregistration plan prior to data collection (https://osf.io/cd4xq), our hypotheses were:

a. STSp peak power will show a positive, moderate correlation (0.5 to 0.8) with peak power measured using the pneumatic leg press.
b. STSp peak power will exhibit a similar or stronger positive correlation (at least 0.05 higher) with physical function measures (performance and patient-reported) in comparison to the pneumatic leg press. For performance measures, we will correlate power measures with Short Physical Performance Battery (SPPB), chair stand (negative correlation), balance, and the Times up and Go test (negative correlation).
c. To assess discriminant validity, STSp peak power will be lower in females than males
d. STSp peak power will have test-retest reliability greater than 0.90.

1. Materials and methods

1.1. Participants

Participants were recruited from the local New York community using flyers, posters, and advertisements in newsletters. The criteria for inclusion were that volunteers should be older than 65 years of age, live independently in the community, and be able to communicate in English. Exclusion criteria were severe knee arthritis (either osteoarthritis or rheumatoid arthritis) that could be exacerbated by exercise, and serious neurological disorders such as Parkinson's disease. Temporary exclusion criteria were major surgery or fracture of the hip or knee, hip/knee replacement or hospitalization in the last 6 months, heart attack or heart disease, major heart surgery, valvular disease, or stroke in the past 6 months. The protocol was approved by the University's Institutional Review Board, and all participants signed informed consent before participation.

Test instructions and procedures were standardized, and the research staff was trained and certified. During the first visit, participants completed the informed consent, reported baseline characteristics, weight and height were measured, and were tested on the following measures in the given sequence:

1.2. Power measurements

1.2.1. Pneumatic leg press

Pneumatic leg press is valid, reliable (Foldvari et al., 2000; Thomas et al., 1996), and is the most commonly used method (63% of power studies) to measure lower body power(Alcazar et al., 2018b). The pneumatic equipment utilizes cylinders pressurized with air to provide resistance rather than weight plates as used in traditional machines. After the tester demonstrated the proper technique, the participant performed 3–5 warm-up repetitions with 50% of their body weight and 1–2 repetitions with their full body weight using a pneumatic leg press (Keiser A300, Keiser Sports Health Equipment, Fresno, CA). The machine was adjusted such that the sitting knee angle was in 90° flexion. If the participant reported pain or was unable to maintain the position due to anatomical restrictions, we moved the seat to the next closest setting. Following the warm-up, resistance was set to their body weight, and participants were instructed, "When you are ready, push as fast as possible," and to perform the lowering phase in a slow, controlled fashion. The software calculates work and power during the concentric phase of each repetition by sampling the system pressure (from which force is calculated) and position at 400 Hz. The highest peak power across three repetitions with 1 min of rest between stands was used for the final analysis.

Peak power was used as the primary outcome since a majority of the studies using the pneumatic leg press have used peak power as the power outcome(Bean et al., 2002; Reid et al., 2008; Marsh et al., 2009). In addition, studies using the sit to-stand test to measure power showed similar correlations for peak power and mean power when compared to functional outcomes. For example, mean power vs peak power using the STSp: TUG test (−0.46 vs −0.46), chair stand (0.63 vs. −0.60), 6 min walk (−0.39 vs. −0.39) (Glenn et al., 2017a). Finally, although the linear transducer provides both peak power and mean power, the pneumatic leg press version (A300) only provides peak power values.

1.2.2. Sit-to-stand power test

A chair and a linear transducer (Tendo Weightlifting Analyzer, Trencin, Slovak Republic) were used to assess peak power during a sit-to-stand test (Gray and Paulson, 2014; Glenn et al., 2017b). The sit to stand test involves standing up from the chair one time. After the tester demonstrated the proper technique, participants performed 3–5 warm-up sit to stands at normal speed before performing the power tests. Subsequently, a belt was secured around the participant's waist (above the Iliac crest). The Kevlar string from the unit was attached to the belt such that the string was perpendicular to the floor when the participant stood up. The participants were instructed to sit in the middle or the edge of the chair to minimize forward trunk lean, and the distance from their feet to the chair was recorded to reduce deviations during subsequent re-testing. The participant began seated with their arms folded across their chest and stood up as quickly and safely as they could before returning to the seated position. The standard instruction before each sit to stand to the participants was, "When you are ready, get up as fast as you can." Power was calculated by the software and displayed from the vertical velocity (m/s) and the mass moved (kg) for the standing portion of the test. The highest peak power across three sit to stand with 1 min of rest between stands was used for analysis (Glenn et al., 2017b; Balachandran et al., 2017). The chair height was 45 cm.

1.3. Physical function measures

1.3.1. SPPB

The Short Physical Performance Battery (SPPB) is widely used in multi-center clinical trials to measure physical function in older adults (Pahor et al., 2018; Pahor et al., 2014). SPPB is reliable and valid for predicting institutionalization, mortality, and disability (Guralnik et al., 1994; Guralnik et al., 2000). The battery involves three tests:

a. Walk speed: A 4-meter walk performed at the usual pace. The faster time out of two trials was recorded.
b. Balance: Three standing balance tests (narrow stance, semi-tandem, and tandem) for 10 s each. The total time per test was recorded.
c. Chair stands: One trial consisting of five chair stand tests performed as quickly as possible. The total time was recorded.

Based on the completion time, each of the three tests is scored between 0 and 4 and subsequently summed to a maximum score of 12 for the total SPPB score, with higher scores indicating better physical performance. The walk speed and chair stand outcomes used for analysis were derived from the SPPB.

1.3.2. Timed up and go

The timed up and go (TUG) measures dynamic balance and agility in older adults (Rikli and Jones, 1999; Podsiadlo and Richardson, 1991). The test involves standing from a chair, walking around a cone 3 m away, and sitting back down. We performed the test at both usual pace and fast pace. There were two trials per pace with 1 min rest, and the faster time was recorded.

1.3.3. Patient-reported outcome measures

For Patient-Reported Outcome Measures (PROMs), we used the Patient Reported Outcomes Measurement Information System (PROMIS) physical function and mobility questionnaire developed by the National Institutes of Health (NIH) (Fries et al., 2014; Fries et al., 2011). PROMIS uses item response theory and computerized adaptive testing to maximize efficiency, and has been shown to be reliable and valid in a large sample of the general population. Participants used the PROMIS iPad App to complete the questionnaires without any help from the study staff.

1.4. Test-retest reliability

To assess test-retest reliability, we assessed 36 out of the 51 participants, who were willing to return to the lab for a second visit. Specifically, participants repeated the pneumatic leg press and chair stand power test on a different occasion within 2 weeks by three assessors. The test instructions, administration, and environment were the same for both tests.

1.5. Sample size

The sample size was pre-registered at 50 participants. Currently, there is a lack of consensus regarding the optimal sample size for studies on the measurement properties for performance measures. According to the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines, a minimum of 50 participants is recommended for studies on reliability and hypotheses testing for construct validity (Mokkink et al., 2010). However, COSMIN was developed for assessing the measurement properties of patient-reported outcome measures (PROMs). For reliability, a sample size of 17 was calculated for test-retest reliability based on ρ_0, the minimally acceptable value (0.7); ρ_1, the hypothesized value of the intraclass correlation coefficient (ICC) (0.9); n the number of observations (2); $\alpha = 0.05$; and $\beta = 0.20$ (Donner and Eliasziw, 1987; Walter et al., 1998).

1.6. Statistical analysis

Continuous variables were expressed as mean (SD) and categorical variables were presented as frequencies and percentages. Data were imported into R (version 4.0.0) for analysis (Team RC, 2013). In accordance with the pre-registration, we assessed the construct validity of the STSp: We calculated Pearson's correlation coefficients, along with their 95% confidence intervals (CI) using the bias-corrected and accelerated bootstrap with 500 replicates. Visual inspection of the data, along with residual and model checking were used to ensure our data were not curvilinear, as has been previously reported for power and physical performance relationships (Bean et al., 2002; Alcazar et al., 2017).

Test-retest reliability was assessed using ICC, calculated using a two-way random-effects model of absolute agreement for single measure (ICC(2,1)). Bland-Altman plots were generated to compare the limits of agreement (LoA) and bias. Smallest Detectable Change (SDC) was calculated as $SDC_{95} = SEM \times 1.96\sqrt{2}$, where the standard error of measurement (SEM) is $\sqrt{\sigma_t^2 + \sigma_{res}^2}$, σ_t^2 is the variance due to systematic differences between time points (test and re-test), and σ_{res}^2 is the residual variance in the random-effects model (de Vet et al., 2006).

Finally, we assessed sex differences within each of the methods using permutation tests. Here, permutation testing was used as a nonparametric alternative to independent samples t-tests. This involves creating a null distribution by randomly shuffling group labels and calculating a test statistic (mean difference between groups) with each shuffle. After 100,000 permutations, we compared our observed mean difference to the permutation (null) distribution, from which we could calculate a z-score and p-value. Again, 95% CIs of these differences were calculated via the bootstrap. Exploratory analyses were conducted using Spearman's correlations and using relative peak power (W/kg). Statistical analyses were blinded by removing data labels (cell scrambling method) for LP and STSp (MacCoun and Perlmutter, 2015).

2. Results

From July 2019 to December 2019, we screened 70 participants and recruited and tested 51 participants. The demographics and characteristics of the participants who completed the validity and reliability section of the study are shown in Table 1. The mean age of the participants was 71.3 years (5.7), 63% females, mean Body Mass Index (BMI) 26.6 (5.4), and the majority (69.9%) were college graduates. The participants were highly physically functioning as shown by mean score of 10.6 (2.6) on the SPPB.

Table 1
Participant characteristics.

	Validity $n = 51$	Reliability $n = 36$
Age, mean, y	71.3 (5.7)	70.4 (5.4)
Gender		
Male	19 (37.3%)	17 (52.8%)
Female	32 (62.7%)	19 (47.2%)
Peak power		
Leg Press (W)	790 (317)	840 (331)
Sit to Stand power test (W)	795 (348)	854 (308)
Leg press (W/kg)	11.3 (4.2)	11.8 (4.5)
Sit to Stand power test (W/kg)	11.6 (4.1)	12.1 (4.5)
BMI, mean (SD)	26.6 (5.4)	27.4 (5.4)
Physical function		
SPPB score, s	10.6 (2.1)	10.6 (2.1)
Chair Stand, s	10.4 (2.8)	10.3 (2.5)
TUG fast, s	6.6 (2.5)	6.4 (1.6)
TUG slow, s	8.5 (2.3)	8.3 (2.2)
PROMIS Mobility t-score	51.6 (7.5)	52.3 (1.6)
PROMIS Function t-score	52.8 (7.1)	53.8 (7.4)
Race/Ethnicity[a]		
White	35 (68.6%)	27 (75.0%)
African American/Black	0 (0.0%)	0 (0.0%)
Asian	12 (23.5%)	6 (16.7%)
Other	4 (8%)	3 (8.4%)
Income (<$75,000/year)[a]	10/31 (32.2%)	6/22(27.0%)
College education[a]	42/49 (69.9%)	32/36 (88.0%)
Conditions, No./total (%)		
Hypertension[a]	20/48 (39.2%)	13/36 (36.0%)
Heart Condition[a]	22/48 (43.1%)	18/36 (50.0%)
Diabetes[a]	4/48 (7.8)	4/36 (11.0%)

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); SPPB, Short Performance Physical Battery. TUG; Timed Up and Go; PROMIS.
[a] Self-reported.

2.1. Construct validity

To determine construct validity, we specified hypotheses which included the direction (negative or positive) and (relative) magnitude of correlations. Table 2 and (Fig. S1 in Supplmementary File) reports Pearson's correlations for the STSp and LP test compared to physical function measures.

Physical function was assessed via:

1. Physical performance measures (SPPB, 4 m walk, Balance, TUG)
2. Patient-Reported Outcome Measures (PROMs), via two questionnaires (mobility and physical function questionnaire).

As shown in Fig. 1 and Table 2, STSp peak power showed a high correlation of 0.90 with LP power. For physical performance outcomes, STSp showed similar or higher correlations compared to the LP test as hypothesized: SPPB ($r = 0.41$ (STSp) vs. 0.29 (LP)), Chair stands (−0.44 vs. -0.35), TUG normal (−0.37 vs. -0.29), TUG Fast (−0.41 vs. -0.34), and balance (0.33 vs. 0.22). However, for PROMs, in contrast to our hypothesis, LP showed a greater correlation with mobility questionnaire (0.34 vs. 0.38) and physical function questionnaire (0.41 vs. 0.48). Exploratory analyses showed Spearman's correlations and peak power relative to body weight (W/kg) to be consistent with our primary analysis as shown in Table S1 and Table S2 in Supplmementary File. For discriminant validity, as expected, males showed higher STSp peak power compared to females ($\Delta = 492$ W, $p < .001$, Cohen's $d = 2.0$) (Table S3 in Supplementary file).

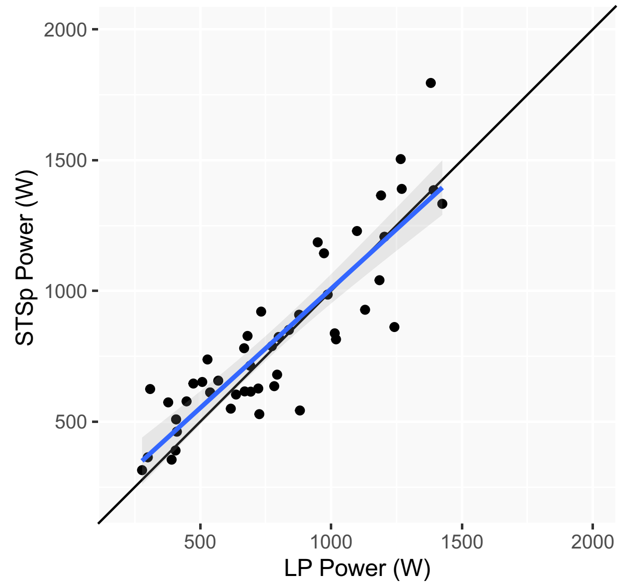

Fig. 1. Scatter plot of STSp and LP peak power.

2.2. Test-rest reliability and measurement error

Test–retest reliability was examined by calculating ICC, SEM, and SDC. As shown in Table 3, the test-retest reliability of the STSp peak power measured on two occasions, within 2 weeks (minimum of 1 week and maximum of 2 weeks), showed an ICC of 0.96 ($CI_{95\%} = 0.93–0.97$). Standard error of measurement (SEM) was 70.4 W and SDC was 192.8 W for STSp. The Bland-Altman plot (Fig. 2) for STSp showed a mean difference (i.e., a systematic change between the occasions) of 6.57 W with LoA of −187.9 to 201.0 W.

2.3. Floor and ceiling effects

Two participants could not perform STSp without the aid of their arms (4%), while two other participants could not perform the LP test (4%), suggesting a floor effect. There were no ceiling effects for either the STSp or the LP as none of the participants reached the maximum or close to the maximum peak power for either the LP or STSp.

Table 2
Pearson's correlation coefficients (and 95% confidence intervals) of the LP and STSp with physical performance and Patient-Reported Outcome Measures.

		Physical performance					Patient-reported outcome measures (PROMIS)	
	$STSp_{Power}$	SPPB	Chair stands	$TUG_{normal,}$	$TUG_{fast,}$	Balance	Mobility	Function
LP_{Power}	0.90 (0.82, 0.94)	**0.29** (0.07, 0.51)	**−0.35** (−0.59, −0.10)	**−0.29** (−0.53, 0)	**−0.34** (−0.57, −0.07)	**0.22** (0, 0.47)	0.38 (0.14, 0.57)	0.48 (0.21, 0.68)
$STSp_{Power}$	–	**0.41** (0.17, 0.59)	**−0.44** (−0.62, −0.12)	**−0.37** (−0.57, −0.05)	**−0.41** (−0.56, −0.14)	**0.33** (0.09, 0.55)	0.34 (0.06, 0.55)	0.41 (0.18, 0.60)

Abbreviations: STSp, Sit to stand power test; LP, Leg press; SPPB, Short Performance Physical Battery. TUG; Timed Up and Go; PROMIS; Patient Reported Outcomes Measurement Information System. Hypothesis accepted in bold.

Table 3
Test re-test reliability and measurement errors.

	$ICC_{agreement}$ (95% CI)	$SEM_{agreement}$ W	SDC W	LoA	Day 1 W	Day 2 W	Day1 – Day 2 W
LP_{Power}	0.99 (0.98, 0.99)	35.7	87.9 W	(−97.5, 100.3)	841 ± 314	840 ± 331	1 ± 51
$STSp_{Power}$	0.96 (0.93, 0.97)	70.4	192.8 W	(−187.9, 201.1)	854 ± 327	848 ± 306	7 ± 101

Abbreviations: STSp, Sit to stand power test; LP, Leg press; ICC, Intraclass correlation coefficient; SEM, Standard Error of Measurement; SDC, Smallest Detectable Change; LOA, Limits of Agreement. The LoA was calculated as the mean difference between the test and retest power mean ± 1.96*SD. SDC based on 95% CI. Day 1, Day 2, and their difference are reported as mean ± SD.

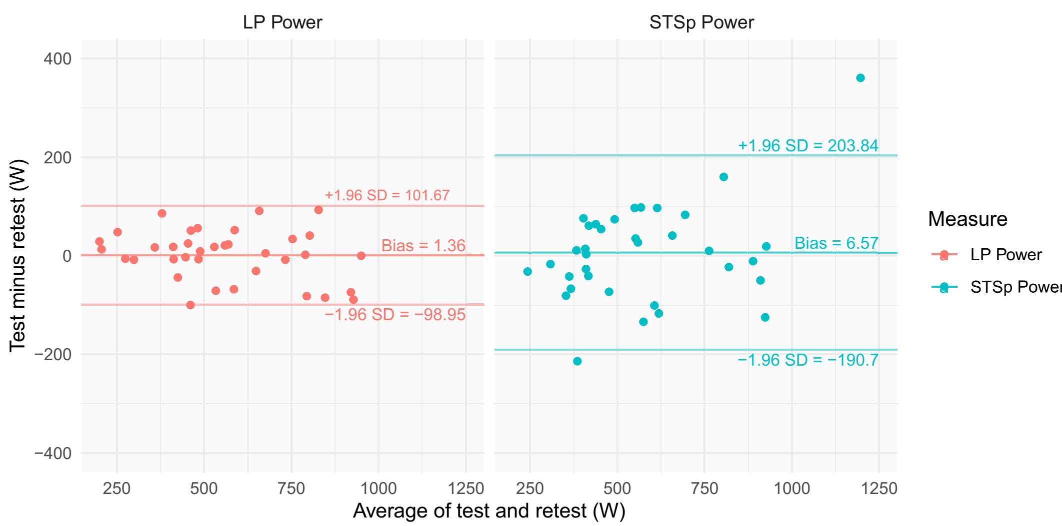

Fig. 2. Bland Altman plot.

2.4. *Missing data*

We omitted two participants from reliability testing since they did not wear footwear for the re-test session. The missing data are reported in Table S4 in Supplmementary File.

2.5. *Adverse events/time taken*

There were no adverse effects reported for either the LP or the STSp. The time taken was 10–15 min for each test, including warm-up and 3 repetitions.

3. Discussion

Power has gained prominence as an important determinant of physical function in the aging population. Our current study expanded the evidence on a promising test to measure lower body power. In this study, we determined construct validity, test-retest reliability, and measurement error of the STSp. As hypothesized, the study showed adequate validity and reliability in assessing lower body power in community-living older adults.

The lower body peak power produced by the LP and STSp had a correlation that was greater than hypothesized (0.90). Despite the strong correlation of 0.90 (Table 2) the STSp showed slightly higher correlations (ranging from 0.07 to 0.12) with performance tests compared to the LP test. We believe that the higher correlations could partly be explained by the ground-based or functional nature of the STSp (Liu et al., 2014; Augustsson et al., 1998b; Balachandran et al., 2016). Since the participant must stand up quickly without support, the STSp involves a dynamic balance and sensorimotor component (Lord et al., 2002). Thus, unlike the LP test, the STSp is not a measure of power in isolation. However, greater correlations were not observed in PROMs, as the LP showed negligibly to slightly greater correlations for both measures of PROMs (mobility and physical function questionnaire). Exploratory analyses using peak power relative to body weight (W/kg) showed slightly higher correlations compared to unnormalized power and were consistent with our primary results. As hypothesized, and similar to the LP, the STSp discriminated between genders (Cohen's d = 2.0, Δ = 455 W), supporting the discriminant validity of the test.

Previous studies that compared STSp to reference tools showed similar correlations. For example, STSp and relative power (W/kg) measured using 2D motion capture during the chair stand test showed a correlation of 0.76; however, the study used the average of the last three trials out of ten trials (Gray and Paulson, 2014). Similarly, STSp had a correlation of 0.65 with the Nottingham Power Rig in older women (Lindemann et al., 2015). However, unlike in our study, the participants used their hands to stand up and used a different linear encoder (MuscleLab Powermodel MLPRO, Ergotest Technology, Langesund, Norway). Another study showed a moderate correlation of 0.70 between STS power and leg press power (Alcazar et al., 2018a). However, unlike our study, the study estimated power using the time taken for five chair stands and the leg press using a linear transducer (T-Force System, Ergotech, Spain) for validity comparison.

For physical performance tests, we observed correlations of 0.44 and 0.41 for chair stands and TUG, respectively, with STSp, while a higher correlation for chair stands (0.58) and TUG (0.48) were observed in a similar study using the same linear encoder (Glenn et al., 2017a). However, the population in that study was slightly older (78 years) and had lower mean peak power (585 W) than our sample. In another study, a correlation of 0.5 with TUG test was observed, but the population was middle-aged, severely obese older adults (>45 BMI) and used a wearable inertial sensor (PUSHTM) which uses an accelerometer (Orange et al., 2019). LP power in low functioning older adults showed a correlation of 0.31 and 0.42 compared to five chair stands and SPPB[60]. Another study showed greater LP mean power correlation with SPPB score ($r = 0.58$ vs. 0.39) than LP peak power (Alcazar et al., 2017). Unfortunately, unlike the upgraded version (A420), we were limited to peak power values with our pneumatic leg press (A300) version. Another study also showed a correlation of 0.45 between equation-based STS power and 3 m walk speed, which is similar to the correlation obtained for STSp power and 4 m walk in the current study as reported in Fig. S1 in Supplementary File (Alcazar et al., 2018a). A recent study showed excellent agreement between STSp using a linear transducer (GymAware™) and Dartfish 2D videography analysis (ICC = 0.98), but they did not report any functional outcomes to compare against (Sherwood et al., 2019). Thus, small differences in correlations between studies could be primarily explained by a combination of factors, including functional status, age, and the methods used in each study. The sample sizes of the above studies ranged from 20 to 138 so differences could be partly due to sample sizes too. In general, our study results are consistent with other studies that used STSp or LP to measure power (Gray and Paulson, 2014; Glenn et al., 2017b; Glenn et al., 2017a; Bean et al., 2007).

Reliability and measurement error are important measurement properties (Streiner et al., 2015), especially when the test involves set-up and proper body positioning. One of the STSp studies reported Cronbach alpha (Vincenzo et al., 2018), but none of the STSp studies assessed test-retest reliability or measurement error. For test-retest reliability, often 0.70 is recommended as a minimum standard for group comparisons and research purposes (Streiner et al., 2015) and greater than 0.90 for individual and clinical decision making. The test-retest reliability of the STSp was excellent (ICC >0.90). SEM and SDC for STSp were 70 W and 193 W, respectively. However, SEM and SDC for LP were half the value of the STSp test. The SEM for 10 step stair climb power is 8.6% in community-dwelling older adults (Ni et al., 2017) and 12.3–12.8% for LP (70% 1RM) (Reid et al., 2015; Chal et al., 2013) in low functioning older adults which is greater than the 8.8% for the STSp reported here. SPPB reported an SEM of 16.4–17% (Mangione et al., 2010; Perera et al., 2006) and 11% for TUG (Mangione et al., 2010; Perera et al., 2006) in low functioning older adults. We would like to note that the percentage values for measurement error reported could change with the sample mean and therefore should be interpreted cautiously. Based on the SDC, any intervention showing an improvement in peak power greater than 193 W in community-living older adults could be considered as 'real' change with a type I error rate of 5% (i.e., assuming any change can be attributed to measurement error and the biological variability occurring over 2 weeks' time). The minimal important change (MIC) for lower body power for the STSp test in community-dwelling older adults is currently unknown and should be established using clinically relevant endpoints.

For a measurement tool to be widely accepted, factors such as feasibility, safety, and cost are just as, if not more, important than the measurement properties. Several tools have been used to assess lower body power in older adults (Table 4): Pneumatic leg press is the most widely used in trials to measure lower body power (Foldvari et al., 2000; Thomas et al., 1996). However, they are expensive, require a compressor, and are not portable. Likewise, isokinetic dynamometry (Suzuki et al., 2001; Leyva et al., 2016) is also widely used, but it is costly, not portable and measures power in a seated, non-functional position. The Nottingham power rig (Bassey and Short, 1990; Bassey et al., 1992) has been used in a few trials, but like other machines, it is expensive, not portable, and measures power in a non-functional manner. Alternatively, jumping tests that use force plates have shown to have sufficient validity and reliability, but there remain safety and feasibility concerns, especially in low functioning older adults (Siglinsky et al., 2015; Buehring et al., 2015). 2D motion analysis (Ford et al., 2007) and ground reaction forces from force plates (Lindemann et al., 2003) have also been used to measure power, but these take considerable time and expertise for analysis. Loaded and unloaded stair climbing has been used to measure power but is limited by accessibility issues (Bean et al., 2007; Ni et al., 2017; Gagliano-Jucá et al., 2020). Chair stand power estimated using time taken is a promising power test that does not require equipment, but the test-retest reliability is still unknown (Alcazar et al., 2018a; Alcazar et al., 2020). Alternatively, STSp is quick, inexpensive, portable, requires no expertise in testing and data acquisition, and importantly, is feasible in a home or clinical environment.

3.1. Limitations and strength

Our study had several limitations: First, the majority of the participants were high functioning and these results cannot be extrapolated to low-functioning older adults; the validity of an instrument is highly dependent on the population and contextual factors (Streiner et al., 2015). Second, although we used correlations to assess validity between instruments, we are unsure how these quantitative differences would impact the interpretation of the test. Third, testers were not blinded to the general hypothesis or the scores. Unlike in a randomized controlled trial, where the fastest time or the heaviest weight is typically better, it can be quite complex to influence a correlation comparison. Having a tester each for each measure—LP, STSp, and functional outcomes—would have been ideal, but unfortunately was not feasible. Finally, we used a sample size of 50, which is considered adequate based on COSMIN (Mokkink et al., 2010), but a larger sample size could have improved the precision as evident by the wide CI's. Some of the delimitations of the current study are: The study was not designed to evaluate responsiveness or the ability to evaluate longitudinal change. If the tool is used for evaluative purposes, the ability to detect change is crucial. Finally, minimal clinically important difference (MCID) or meaningful change of power for STSp is not known (Jaeschke et al., 1989). Without the knowledge of MIC/MCID, the interpretability or the clinical utility of the scores is uncertain.

Limitations of the STSp: First, the knee angles for the STSp could not be controlled since the chair height (45 cm) was fixed. This is a limitation compared to leg press where knee angles can be adjusted. Second, the string from the Tendo unit tethered to the waist was kept perpendicular during standing. Any inclination was kept to a minimum. However, considering the different positions assumed during sitting and standing, slight changes in the string angle is unavoidable. According to the manufacturer, angle deviations relative to greater than 20° can influence power readings. Third, the shank and feet mass during the STSp

Table 4
Comparisons of methods to measure lower body power.

Lower body power tools	Cost $	Safety	Time (min)	Expertise	Portable	Functional
Pneumatic leg press	10–15 K	High	5–10	Low	No	No
Notthingham power rig	10–15 K	High	5–10	Low	No	No
Isokinetic leg extension	40–45 K	High	15–25	Moderate	No	No
STSp w/ linear transducer	1–2 K	High	5–10	Low	Yes	Yes
STS w/equation	NA	High	0–5	Low	Yes	Yes
STSp w/ motion capture	10–15 K	High	15–25	High	No	Yes
Jumping w/ force plate	10–15 K	Low	15–25	High	Yes	Unclear
Star climb w/equation	NA	Moderate	5–10	Low	No	Yes

Abbreviations: STSp, Sit to stand power test.

are not displaced, but the linear transducer includes the whole body weight in calculating power during the STSp. Considering the favorable results of the study, along with those from previous studies, we contend that these limitations of the STSp should not have drastically affected the validity and reliability.

The major strength of the study is the pre-registration of the research questions, outcomes, and the analysis plan before data collection. Pre-registration attenuates *P*-hacking and selective reporting, and has been shown to result in smaller effect sizes being reported compared to non-registered studies (Scheel et al., 2020; Open Science Collaboration, 2015). In addition, we conducted blinded statistical analyses (MacCoun and Perlmutter, 2015), further improving the rigor of the study. Finally, unlike other studies, we examined the relationship to physical function or meaningful outcomes by directly comparing our measures to a reference tool that is widely used. This was done in lieu of 'establishing' validity based on statistical significance or arbitrary effect size thresholds.

3.2. Conclusions

Power is an important determinant of physical function in the aging population. The sit to stand power test is a promising method to assess lower-body power that is quick, relatively inexpensive, safe, portable, feasible, and functional. In the current study, STSp showed adequate validity and test-retest reliability in measuring lower body power in community-dwelling older adults compared to pneumatic leg press and should be considered for future use.

CRediT authorship contribution statement

Anoop T. Balachandran: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Validation. **Andrew D. Vigotsky:** Formal analysis, Writing – review & editing, Validation. **Norberto Quiles:** Conceptualization, Methodology, Writing – review & editing, Validation. **Lidwine B. Mokkink:** Formal analysis, Writing – review & editing, Validation. **Mark A. Belio:** Formal analysis, Writing – review & editing, Validation. **Jordan McKenzie Glenn:** Conceptualization, Methodology, Writing – review & editing, Validation.

Acknowledgements

We thank all our participants for participating in the study. We would also like to thank Commonpoint Queens' Central Queens, Body by Berle, and Pomonok Community center for their help in participant recruitment. We would also like to thank David L. Streiner for his valuable suggestions and guidance.

Declaration of competing interest

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.exger.2020.111202.

References

Alcazar, J., Rodriguez-Lopez, C., Ara, I., et al., 2017. The force-velocity relationship in older people: reliability and validity of a systematic procedure. Int. J. Sports Med. 38 (14), 1097–1104.
Alcazar, J., Losa-Reyna, J., Rodriguez-Lopez, C., et al., 2018a. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Exp. Gerontol. 112, 38–43.
Alcazar, J., Guadalupe-Grau, A., García-García, F.J., Ara, I., Alegre, L.M., 2018b. Skeletal muscle power measurement in older people: a systematic review of testing protocols and adverse events. J. Gerontol. A Biol. Sci. Med. Sci. 73 (7), 914–924.
Alcazar, J., Kamper, R.S., Aagaard, P., et al., 2020. Relation between leg extension power and 30-s sit-to-stand muscle power in older adults: validation and translation to functional performance. Sci. Rep. 10 (1), 1–8.
Augustsson, J., Esko, A., Thomee, R., Svantesson, U., 1998a. Weight training of the thigh muscles using closed vs. open kinetic chain exercises: a comparison of performance enhancement. J Orthop Sports Phys Ther 27 (1), 3–8.
Augustsson, J., Esko, A., Thomee, R., Svantesson, U., 1998b. Weight training of the thigh muscles using closed vs. open kinetic chain exercises: a comparison of performance enhancement. J Orthop Sports Phys Ther 27 (1), 3–8.
Balachandran, A., Martins, M.M., De Faveri, F.G., Alan, O., Cetinkaya, F., Signorile, J.F., 2016. Functional strength training: seated machine vs standing cable training to improve physical function in elderly. Exp. Gerontol. 82, 131–138.
Balachandran, A.T., Gandia, K., Jacobs, K.A., Streiner, D.L., Eltoukhy, M., Signorile, J.F., 2017. Power training using pneumatic machines vs. plate-loaded machines to improve muscle power in older adults. Exp. Gerontol. 98, 134–142.
Bassey, E., Short, A., 1990. A new method for measuring power output in a single leg extension: feasibility, reliability and validity. Eur. J. Appl. Physiol. Occup. Physiol. 60 (5), 385–390.
Bassey, E.J., Fiatarone, M.A., O'neill, E.F., Kelly, M., Evans, W.J., Lipsitz, L.A., 1992. Leg extensor power and functional performance in very old men and women. Clin. Sci. 82 (3), 321–327.
Bean, J.F., Kiely, D.K., LaRose, S., Alian, J., Frontera, W.R., 2007. Is stair climb power a clinically relevant measure of leg power impairments in at-risk older adults? Arch. Phys. Med. Rehabil. 88 (5), 604–609.
Bean, Jonathan F., Kiely, D.K., Herman, S., et al., 2002. The relationship between leg power and physical performance in mobility limited older people. J. Am. Geriatr. Soc. 50 (3), 461–467.
Bergen, G., Stevens, M.R., Burns, E.R., 2016. Falls and fall injuries among adults aged greater than or equal to 65 years–United States, 2014. Morb Mortal Weekly Rep 65 (37), 993.
Buehring, B., Krueger, D., Fidler, E., Gangnon, R., Heiderscheit, B., Binkley, N., 2015. Reproducibility of jumping mechanography and traditional measures of physical and muscle function in older adults. Osteoporosis Int 26 (2), 819–825.
Byrne, C., Faure, C., Keene, D.J., Lamb, S.E., 2016. Ageing, muscle power and physical function: a systematic review and implications for pragmatic training interventions. Sports Med. 46 (9), 1311–1332.
Chal, Ã.A., Cloutier, G.J., Hau, C., Phillips, E.M., Dallal, G.E., Fielding, R.A., 2013. Efficacy of whey protein supplementation on resistance exercise induced changes in lean mass, muscle strength, and physical function in mobility-limited older adults. J. Gerontol. A Biol. Sci. Med. Sci. 68 (6), 682–690.
De Vet, H.C., Terwee, C.B., Mokkink, L.B., Knol, D.L., 2011. Measurement in Medicine: A Practical Guide. University Press, Cambridge.
Donner, A., Eliasziw, M., 1987. Sample size requirements for reliability studies. Stat. Med. 6 (4), 441–448.
Foldvari, M., Clark, M., Laviolette, L.C., et al., 2000. Association of muscle power with functional status in community-dwelling elderly women. J. Gerontol. A Biol. Sci. Med. Sci. 55 (4), M192–M199.
Ford, K.R., Myer, G.D., Hewett, T.E., 2007. Reliability of landing 3D motion analysis: implications for longitudinal analyses. Med. Sci. Sports Exerc. 39 (11), 2021–2028.
Fried, L.P., Guralnik, J.M., 1997. Disability in older adults: evidence regarding significance, etiology, and risk. J. Am. Geriatr. Soc. 45 (1), 92–100.
Fries, J.F., Krishnan, E., Rose, M., Lingala, B., Bruce, B., 2011. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. Arthritis Res. Ther. 13 (5), 1–8.
Fries, J.F., Witter, J., Rose, M., Cella, D., Khanna, D., Morgan-DeWitt, E., 2014. Item response theory, computerized adaptive testing, and PROMIS: assessment of physical function. J. Rheumatol. 41 (1), 153–158.
Frontera, W.R., Reid, K.F., Phillips, E.M., et al., 2008. Muscle fiber size and function in elderly humans: a longitudinal study. J Appl Physiol (1985) 105 (2), 637–642.
Gagliano-Jucá, T., Li, Z., Pencina, K.M., et al., 2020. The stair climb power test as an efficacy outcome in randomized trials of function promoting therapies in older men. J. Gerontol. A Biol. Sci. Med. Sci. 75 (6), 1167–1175.
Glenn, J.M., Gray, M., Binns, A., 2017a. Relationship of sit-to-stand lower-body power with functional fitness measures among older adults with and without sarcopenia. J Geriatr Phys Ther 40 (1), 42–50.
Glenn, J.M., Gray, M., Vincenzo, J., Paulson, S., Powers, M., 2017b. An evaluation of functional sit-to-stand power in cohorts of healthy adults aged 18–97 years. J. Aging Phys. Act. 25 (2), 305–310.
Gray, M., Paulson, S., 2014. Developing a measure of muscular power during a functional task for older adults. BMC Geriatr. 14, 145.
Guralnik, J.M., Simonsick, E.M., Ferrucci, L., et al., 1994. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J. Gerontol. 49 (2), M85–M94.
Guralnik, J.M., Ferrucci, L., Pieper, C.F., 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 (4), M221–M231.
Jaeschke, R., Singer, J., Guyatt, G.H., 1989. Measurement of health status. Ascertaining the minimal clinically important difference. Control. Clin. Trials 10 (4), 407–415.
Janssen, I., Heymsfield, S.B., Ross, R., 2002. Low relative skeletal muscle mass (sarcopenia) in older persons is associated with functional impairment and physical disability. J. Am. Geriatr. Soc. 50 (5), 889–896.

Kaplan, R.M., Irvin, V.L., 2015. Likelihood of null effects of large NHLBI clinical trials has increased over time. PLoS One 10 (8), e0132382.
Landi, F., Cesari, M., Calvani, R., et al., 2017. The sarcopenia and physical frailty in older people: multi-component treatment strategies (SPRINTT) randomized controlled trial: design and methodt. Aging Clin. Exp. Res. 29 (1), 89–100.
Leyva, A., Balachandran, A., Signorile, J.F., 2016. Lower-body torque and power declines across six decades in three hundred fifty-seven men and women: a cross-sectional study with normative values. J Strength Cond Res 30 (1), 141–158.
Lindemann, U., Claus, H., Stuber, M., et al., 2003. Measuring power during the sit-to-stand transfer. Eur. J. Appl. Physiol. 89 (5), 466–470.
Lindemann, U., Farahmand, P., Klenk, J., Blatzonis, K., Becker, C., 2015. Validity of linear encoder measurement of sit-to-stand performance power in older people. Physiotherapy 101 (3), 298–302.
Liu, C., Shiroy, D.M., Jones, L.Y., Clark, D.O., 2014. Systematic review of functional training on muscle strength, physical functioning, and activities of daily living in older adults. Eur. Rev. Aging Phys. Act. 11 (2), 95–106.
Lord, S.R., Ward, J.A., Williams, P., Anstey, K.J., 1994. Physiological factors associated with falls in older community-dwelling women. J. Am. Geriatr. Soc. 42 (10), 1110–1117.
Lord, S.R., Murray, S.M., Chapman, K., Munro, B., Tiedemann, A., 2002. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. J. Gerontol. A Biol. Sci. Med. Sci. 57 (8), M539–M543.
MacCoun, R., Perlmutter, S., 2015. Blind analysis: Hide results to seek the truth. Nature 526 (7572), 187–189.
Mangione, K.K., Craik, R.L., McCormick, A.A., et al., 2010. Detectable changes in physical performance measures in elderly african americans. Phys. Ther. 90 (6), 921–927.
Marsh, A.P., Miller, M.E., Rejeski, W.J., Hutton, S.L., Kritchevsky, S.B., 2009. Lower extremity muscle function after strength or power training in older adults. J. Aging Phys. Act. 17 (4), 416–443.
Metter, E.J., Conwit, R., Tobin, J., Fozard, J.L., 1997. Age-associated loss of power and strength in the upper extremities in women and men. J. Gerontol. A Biol. Sci. Med. Sci. 52 (5), B267–B276.
Mokkink, L.B., Terwee, C.B., Patrick, D.L., et al., 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international delphi study. Qual. Life Res. 19 (4), 539–549.
Ni, M., Brown, L.G., Lawler, D., Bean, J.F., 2017. Reliability, validity, and minimal detectable change of four-step stair climb power test in community-dwelling older adults. Phys. Ther. 97 (7), 767–773.
Open Science Collaboration, 2015. PSYCHOLOGY. Estimating the reproducibility of psychological science. Science 349 (6251), aac4716.
Orange, S.T., Marshall, P., Madden, L.A., Vince, R.V., 2019. Can sit-to-stand muscle power explain the ability to perform functional tasks in adults with severe obesity? J. Sports Sci. 37 (11), 1227–1234.
Pahor, M., Guralnik, J.M., Ambrosius, W.T., et al., 2014. Effect of structured physical activity on prevention of major mobility disability in older adults: the LIFE study randomized clinical trial. JAMA 311 (23), 2387–2396.
Pahor, M., Anton, S.D., Beavers, D.P., et al., 2018. Effect of losartan and fish oil on plasma IL-6 and mobility in older persons. The ENRGISE pilot randomized clinical trial. J. Gerontol. A Biol. Sci. Med. Sci. 73 (12), 1–8.
Perera, S., Mody, S.H., Woodman, R.C., Studenski, S.A., 2006. Meaningful change and responsiveness in common physical performance measures in older adults. J. Am. Geriatr. Soc. 54 (5), 743–749.
Podsiadlo, D., Richardson, S., 1991. The timed "up & go": a test of basic functional mobility for frail elderly persons. J. Am. Geriatr. Soc. 39 (2), 142–148.
Reid, K.F., Fielding, R.A., 2012. Skeletal muscle power: a critical determinant of physical functioning in older adults. Exerc. Sport Sci. Rev. 40 (1), 4–12.
Reid, K.F., Callahan, D.M., Carabello, R.J., Phillips, E.M., Frontera, W.R., Fielding, R.A., 2008. Lower extremity power training in elderly subjects with mobility limitations: a randomized controlled trial. Aging Clin. Exp. Res. 20 (4), 337–343.
Reid, K.F., 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 (1), 29–39.
Reid, K.F., Martin, K.I., Doros, G., et al., 2015. Comparative effects of light or heavy resistance power training for improving lower extremity power and physical performance in mobility-limited older adults. J. Gerontol. A Biol. Sci. Med. Sci. 70 (3), 374–380.
Rikli, R., Jones, C., 1999. Development and validation of a functional fitness test for community- residing older adults. J. Aging Phys. Act. 7 (2), 129–161.
da Rosa, Orssatto, Bet, Lucas, de la Rocha, Freitas, Cintia, Shield, A.J., Pinto, R.S., Trajano, G.S., 2019. Effects of resistance training concentric velocity on older adults' functional capacity: a systematic review and meta-analysis of randomised trials. Exp. Gerontol. 127, 110731.
Sayers, S.P., Guralnik, J.M., Thombs, L.A., Fielding, R.A., 2005. Effect of leg muscle contraction velocity on functional performance in older men and women. J. Am. Geriatr. Soc. 53 (3), 467–471.
Scheel, A.M., Schijen, M., Lakens, D., 2020. An excess of positive results: comparing the standard psychology literature with registered reports. Psyarxiv. https://doi.org/10.31234/osf.io/p6e9c.
Sherwood, J.J., Inouye, C., Webb, S.L., Jenny, O., 2019. Reliability and validity of the sit-to-stand as a muscular power measure in older adults. J. Aging Phys. Act. 28 (3), 455–466.
Siglinsky, E., Krueger, D., Ward, R.E., et al., 2015. Effect of age and sex on jumping mechanography and other measures of muscle mass and function. J. Musculoskelet. Neuronal Interact. 15 (4), 301–308.
Skelton, D.A., Greig, C.A., Davies, J.M., Young, A., 1994. Strength, power and related functional ability of healthy people aged 65-89 years. Age Ageing 23 (5), 371–377.
Skelton, D.A., Kennedy, J., Rutherford, O.M., 2002. Explosive power and asymmetry in leg muscle function in frequent fallers and non fallers aged over 65. Age Ageing 31 (2), 119–125.
Stathi, A., Withall, J., Greaves, C.J., et al., 2018. A community-based physical activity intervention to prevent mobility-related disability for retired older people (REtirement in ACTion (REACT)): study protocol for a randomised controlled trial. Trials 19 (1), 228.
Streiner, D.L., Norman, G., Cairney, J., 2015. Health Measurement Scales. A Practical Guide to their Development and Use, 5th ed. Oxford University Press.
Suzuki, T., Bean, J.F., Fielding, R.A., 2001. Muscle power of the ankle flexors predicts functional performance in community dwelling older women. J. Am. Geriatr. Soc. 49 (9), 1161–1167.
Team RC, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.
Thomas, M., Fiatarone, M.A., Fielding, R.A., 1996. Leg power in young women: relationship to body composition, strength, and function. Med. Sci. Sports Exerc. 28 (10), 1321–1326.
Tschopp, M., Sattelmayer, M.K., Hilfiker, R., 2011. Is power training or conventional resistance training better for function in elderly persons? A meta-analysis. Age Ageing 40 (5), 549–556.
de Vet, H.C., Terwee, C.B., Knol, D.L., Bouter, L.M., 2006. When to use agreement versus reliability measures. J. Clin. Epidemiol. 59 (10), 1033–1039.
Vincenzo, J.L., Gray, M., Glenn, J.M., 2018. Validity of a novel, clinically relevant measure to differentiate functional power and movement velocity and discriminate fall history among older adults: a pilot investigation. Innov. Aging 2 (3), igy028.
Walter, S., Eliasziw, M., Donner, A., 1998. Sample size and optimal designs for reliability studies. Stat. Med. 17 (1), 101–110.
Werner, C., 2010. The Older Population. http://www.census.gov/prod/cen2010/briefs/c2010br-09.pdf.
Whipple, R., Wolfson, L., Amerman, P., 1987. The relationship of knee and ankle weakness to falls in nursing home residents: an isokinetic study. J. Am. Geriatr. Soc. 35 (1), 13–20.