Original Paper

Remote Asynchronous Smartphone 5 Times Sit-to-Stand Testing in Older Adults: Pragmatic Validity Trial

Ethan CJ Berry[1], BSc, MSc; Lawrence D Hayes[2], BSc, MSc, PhD; Ashley Warner[3], BSc, MSc, PhD; Nicholas F Sculthorpe[3], BSc, MSc, PhD

[1]School of Medicine, Medical Sciences & Nutrition, University of Aberdeen, Aberdeen, Scotland, United Kingdom

[2]Lancaster Medical School, Faculty of Health & Medicine, Lancaster University, Lancaster, England, United Kingdom

[3]Department of Sports and Sciences, Faculty of Computing, Sciences and Engineering, University of the West of Scotland, Glasgow, Scotland, United Kingdom

Corresponding Author:
Lawrence D Hayes, BSc, MSc, PhD
Lancaster Medical School
Faculty of Health & Medicine
Lancaster University
Sir John Fisher Drive
Lancaster, England, LA1 4YW
United Kingdom
Phone: 44 7743 061291
Email: l.hayes4@lancaster.ac.uk

Abstract

Background: Age-related declines in muscle function threaten independence and increase the risk of falls in community-dwelling older adults. Muscle-strengthening exercises delivered via mobile health (mHealth) apps could help address access barriers, but scalable programs require valid remote assessments of lower-limb function. The 5 times sit-to-stand (5TSTS) test is a well-established indicator of lower-limb strength and functional capacity, yet its validity in asynchronous, smartphone video formats remains unclear.

Objective: This study aimed to examine validity, intermethod relative agreement, intermethod measurement variability, and convergent validity of a single smartphone-recorded, asynchronously scored 5TSTS.

Methods: A total of 32 older adults aged >60 years attended a single laboratory visit for isokinetic dynamometry of dominant-limb knee extensors and standardized 5TSTS testing and then completed a home-based 5TSTS test after a washout period of at least 3 days, recording the test on their own smartphone. Videos were timed offline by the lead author (EB) using video analysis software. Validity was assessed using a Bland-Altman analysis comparing laboratory and remote 5TSTS tests, while intermethod relative agreement was assessed using the intraclass correlation coefficient (ICC). Intermethod measurement variability was quantified using SEM. Pearson correlations were used to examine associations between body weight–normalized peak torque and both 5TSTS measures to test convergent validity.

Results: A total of 29 participants provided complete laboratory and remote data. Bland-Altman analysis identified that remote 5TSTS times were, on average, 0.3 seconds faster than laboratory values, with 95% limits of agreement of −2.24 to 2.84 seconds, which is promising. Intermethod relative agreement and measurement variability metrics indicated good agreement between laboratory and remote 5TSTS (ICC=0.89; SEM=0.9 seconds). Both laboratory and remote 5TSTS showed moderate negative correlations with normalized knee extensor torque, indicating that greater strength was associated with faster performance in both conditions.

Conclusions: A single asynchronous smartphone video 5TSTS demonstrates promising pragmatic validity, relative agreement with laboratory assessment, and moderate convergent validity with isokinetic knee extensor strength. This could potentially support its use as a future pragmatic remote tool for monitoring progress within mHealth muscle-strengthening programs for community-dwelling older adults at the group level; however, research on test-retest reliability and responsiveness is required before it can be implemented.

Trial Registration: OSF Registries osf.io/xqt52; https://osf.io/xqt52

XSL•FO
RenderX

(JMIR Aging 2026;9:e108345) doi: 10.2196/108345

KEYWORDS

mobile health; mHealth; muscle strengthening; older adults; 5 times sit-to-stand test; remote assessment; aged

Introduction

Background

The 5 times sit-to-stand (5TSTS) test is a performance-based and clinical test that is well standardized and devised by Csuka and McCarty [1]. The main purpose of using the test is to assess lower-limb strength, balance, and postural control in older adults, and it has high reliability [2]. The 5TSTS test is widely used due to its simplicity; minimal equipment requirements; and strong evidence supporting its validity, reliability, and safety [3-5]. The 5TSTS test measures the power of hip and knee extensors and provides information on balance and postural control relevant to fall prevention [3,6]. Its clinical utility lies in its ability to objectively track changes in strength and mobility [7]. Previous research has used this test in older adults in both community and laboratory or clinical settings due to its quick and easy implementation and low time or space requirements [8-11]. The test has been previously validated for predicting adverse health outcomes in community-dwelling older adults, with 1 study finding times greater than 10 seconds were associated with increased risk of developing disability [12]. van Cappellen-van Maldegem et al [13] highlighted the 5TSTS test's ability to act as a screening tool for frailty in clinical populations (intraclass correlation coefficient [ICC]=0.70; SEM=2.2; minimally important clinical difference (MID)=3.8). Trivedi et al [14] demonstrated the 5TSTS test as responsive to exercise-based interventions, highlighting its utility for tracking improvements. Moreover, previous work has highlighted a minimal detectable change at the 95% CI (MDC95) of 3.12 seconds in older adults, indicating the magnitude of change required to exceed measurement error [15].

Steffens et al [5] demonstrated that remote delivery of the 5TSTS test via Zoom was feasible, reliable, and safe in 37 patients with cancer preparing for surgery. In older adults, van Schooten et al [16] similarly reported good agreement between face-to-face and remote (Zoom-based) 5TSTS assessments, highlighting their suitability for large-scale remote muscle strength testing. Within mobile health (mHealth), the PowerFrail app developed by Alcazar et al [17] enables remote sit-to-stand assessment to estimate relative muscle power and strength and assess frailty. The mobile app estimates relative muscle power using a validated equation and the user's anthropometric data. The app can work using any sit-to-stand test, including both the 5TSTS and 30-second sit-to-stand tests. It can take the result either from direct entry or video recording, compare it with normative values to identify low power, and flag which features may need improvement. It also assesses frailty using the validated Frailty Trait Scale—Short Form, which produces a frailty score from 0 to 50 and provides tailored physical activity recommendations. The app saves data and generates a report, making it suitable for clinical, research, and public health settings. Other studies have used smartphones attached to the back or chest to capture accelerometry data during the 5TSTS test [18,19]. Moreover, Marques et al [20] used a smartphone app to quantify the validity and reliability of sit-to-stand performance among older adults living in care. Taken together, existing approaches rely either on synchronous videoconferencing with live supervision or on additional hardware or attachment requirements, and tools such as PowerFrail have been designed primarily for classification of low relative muscle power and frailty, rather than for monitoring strength adaptations within exercise interventions in community-dwelling older adults. Moreover, research using smartphone apps to quantify the validity and reliability of sit-to-stand performance has focused on older adults living in care, limiting generalizability to community settings.

There remains limited evidence on the validity and agreement between laboratory 5TSTS assessment and asynchronous smartphone video 5TSTS measurement in community-dwelling older adults. Establishing this is an essential first step toward simple, unsupervised, app-enabled monitoring of functional performance in remote smartphone-delivered muscle-strengthening programs. This will help move toward future automated 5TSTS assessment. We were recently involved in co-designing an mHealth tool to help older adults perform muscle-strengthening exercises at home [21]. The app includes an integrated remote strength and functional assessment using smartphone video to record the 5TSTS test, which is manually but asynchronously scored by a researcher. This provides a practical, at-home indicator of lower-limb strength and function. While the app is under development, it is prudent to assess the validity of smartphone camera–based remote, asynchronous 5TSTS assessment.

Aims and Objectives

The overarching aim of this study was to pragmatically assess the validity and relative agreement between remote 5TSTS performance measured via smartphone camera and laboratory-based assessment using Bland-Altman analysis. By "pragmatically," we mean a single 5TSTS recording to reduce user burden. The following objectives were addressed: evaluation of the validity of remote asynchronous 5TSTS assessment from a single trial and evaluation of the correlation between remote and laboratory-based assessments of 5TSTS and knee extensor strength assessed via isokinetic dynamometry.

Methods

Study Preregistration

The study was preregistered on OSF [22], which included full details on study design, approach, and data analysis plan prior to the commencement of recruitment and data collection. The raw data can be found in the associated project linked to the OSF registration. The study was reported in line with the Guidelines for Reporting Reliability and Agreement Studies [23]. This 15-item checklist covers essential reporting

XSL•FO
RenderX

requirements for the Introduction, Methods, Results, and Discussion sections.

Ethical Considerations

Ethics approval was obtained from the University of the West of Scotland's Health and Life Sciences Academic Integrity and Ethics Committee (23548). The study was conducted in accordance with the Declaration of Helsinki (2024 version) issued by the World Medical Association.

Study Design

This study was a cross-sectional study evaluating the validity, intermethod relative agreement, intermethod measurement validity, and convergent validity of laboratory and remote 5TSTS tests. The primary outcome was validity, quantified using Bland-Altman analysis. The secondary outcomes were intermethod relative agreement (by ICC) and intermethod measurement variability (SEM). The study's tertiary outcome was convergent validity, assessed through correlations between remote and laboratory-based 5TSTS tests and knee extensor strength measured using isokinetic dynamometry.

Participants and Recruitment Strategy

Participants were recruited through local organizations (eg, South and North Lanarkshire councils, the University of the Third Age, Men's Sheds, and Age Scotland), which circulated study information to eligible older adults, and via the social media of the lead author (EB) and posters with a QR code linking to a Microsoft Forms sign-up page. Interested individuals could leave contact details, and some participants also recruited peers by word of mouth through community groups and sports clubs.

Inclusion criteria were as follows: community-dwelling adults aged ≥60 years, with access to a mobile device, and no serious musculoskeletal injury preventing muscle-strengthening exercise. Potential participants received a detailed information sheet and, if willing to take part, provided informed consent, provided demographic data, and completed a Physical Activity Readiness Questionnaire; any contraindications or relevant injuries led to exclusion, and new health issues during the study required medical clearance.

All participants attended a single laboratory at the UWS Lanarkshire Campus Physiotherapy Laboratory (June-August 2025) for measurements and procedures. After a washout period of at least 3 days, they completed the remote part of data collection at home, using their smartphone to video-record themselves performing a functional capacity test. This conservative period was selected to mitigate the learning effects of the laboratory effort. As reported in the literature, older adults do not demonstrate a learning effect on the 5TSTS test with repeated attempts [24]. Participants received instructions in the laboratory to minimize their burden.

Procedures

Preparation and Warm-Up—Laboratory Visit

Prior to all the following procedures, participants were asked not to consume caffeine or alcohol and not to take part in vigorous physical exertion 24 hours prior to the testing for both the laboratory and remote arms of data collection. They complied according to self-report. Participants were advised to wear appropriate clothing for the testing (ie, loose-fitting clothing and shorts). Upon entry to the laboratory, participants were greeted by the lead author (EB), given an outline of what would be expected of them, and asked to read and complete all necessary paperwork if not already completed. Participants then provided demographic data, including age, height, body mass, BMI, blood pressure, sex, ethnicity, nationality, occupation, and education. Height, body mass, and blood pressure were measured in the laboratory as a means of profiling the participants.

A standardized warm-up was completed by each participant. The warm-up consisted of 5 minutes on a cycle ergometer (Wattbike Pro; Wattbike) with the fan resistance progressively increased over time to achieve 100 W and 60 r/min, followed by dynamic stretches of the major muscle groups performed on a matted area. A rest period of at least >5 minutes was given to participants between each of the laboratory assessments to mitigate adverse effects on performance in the tests. Isokinetic dynamometry was performed first, as it was anticipated to be the most physically demanding procedure for participants.

Isokinetic Dynamometry

Dominant-limb knee extensor strength was assessed using isokinetic dynamometry (HUMAC NORM; CSMi). The isokinetic dynamometer (IKD) was calibrated and set up in a standardized position [25-27], and participant positioning was carefully aligned to isolate the quadriceps during testing. Participants sat with the hip flexed at 90°, the knee aligned to the IKD axis, and the lower leg and trunk secured with straps to minimize compensatory movement. The test was performed at an angular velocity of 60°/s, with 3 submaximal familiarization repetitions (20%, 40%, and 60% effort), followed by a 1-minute rest, and then 1 set of 5 maximal knee extensions on the dominant leg with at least 30-second rest as needed between repetitions. Participants were instructed to extend the knee as hard and fast as possible through full range, with verbal encouragement provided. Body weight–normalized torque (N m/kg) was recorded, with the best of the 5 trials used for analysis, and testing was stopped if pain, instability, or significant fatigue occurred.

5TSTS Test

Laboratory

Participants were given a further 5-minute warm-up after their rest time upon completion of the dynamometry testing. The warm-up included 3 main phases. First, participants engaged in a general warm-up, which included marching on the spot, standing leg swings, and ankle rotations. Next, participants were advised to complete seated knee extensions and body weight squats. Finally, participants practiced the sit-to-stand test using slow, controlled movements.

A standard chair (46 cm high) was used for the test. It was placed on a flat, stable surface and against a wall to avoid sliding during the test. Participants were given the opportunity to complete further practice repetitions. Participants were then instructed to sit with their feet flat on the floor, back against the

XSL•FO
RenderX

chair, arms crossed over their chests, and knees at 90° flexion. On researcher command, the participants rose to full standing position and then sat back down on the chair. This movement was completed 5 times as quickly as possible without deviation from the technique. Two full trials were completed with a 1-minute rest period between trials, and the mean time of both trials was used to score the test; time was recorded using a stopwatch. If participants fell or lost balance during the test, the test was voided and restarted after a rest period [5,28].

Remote

Participants were asked to replicate the abovementioned protocol as closely as possible in their own homes, completing a single trial. Verbal instructions were reinforced to participants to ensure they were clear on how to safely and effectively complete the test. Instruction was given to participants to ensure they found a suitable, spacious, and hazard-free area in the home to complete the test and to use a chair of the same height as the one used in the laboratory, positioned against a wall to prevent it from sliding during the test. Although the chair height was standardized in the laboratory, this could not be controlled in participants' homes. Participants were therefore instructed to use a stable chair of similar height, where possible, aiming for 90° knee flexion, as was the case in the laboratory, reflecting the pragmatic conditions under which the app would be used in real-world settings. This approach prioritized feasibility and ecological validity while reducing participant burden, for example, by eliminating the need to measure chair height. Similar approaches have been used in previous protocols [5]. Participants were also encouraged to have someone present while performing the test in case of an injury. Participants were asked to start recording on their mobile devices prior to completing the test to ensure the whole test was recorded. Participants were asked to ensure that their phone cameras were set to 30 frames per second, as is typically the smartphone default, to send this file to the research team. Video files were scored asynchronously, remotely, and manually by the lead author (EB) using the DaVinci Resolve video analysis software (Blackmagic Design Pty Ltd) [29]. EB was not blinded to the laboratory-based 5TSTS scores because he conducted the in-person testing. The score by EB was taken as the ground truth for the remote 5TSTS value for comparison with laboratory testing, but to test EB's scoring, a second rater (LH) also scored all video files. LH was blinded to the scores by EB and to the laboratory scores. The inter-rater agreement for video files was 0.94 (95% CI 0.88-0.97), indicating excellent inter-rater relative agreement for the remote manual scoring of 5TSTS.

Statistical Analysis

Data were recorded in the laboratory using a Microsoft 365 Excel spreadsheet on a secure password-protected device. Data on knee extensor strength were automatically saved on the IKD. All data analysis and visualization were completed in RStudio (version 2026.01.0+392; Posit PBC) [30]. Descriptive statistics were produced, including mean (SD), to characterize the included sample. To test validity, the agreement between remote and laboratory-based 5TSTS was assessed using Bland-Altman analysis. This included the mean bias between both measures and the limits of agreement around the mean bias. The maximum clinically acceptable disagreement between laboratory and remote assessments was not prespecified because established thresholds for this specific study were scarce. Therefore, agreement between measures was assessed in conjunction with previously reported 5TSTS literature. Intermethod relative agreement was assessed via 2-way random-effects, absolute-agreement ICC(2,1). A value <0.50 was interpreted as poor, 0.50-0.75 as moderate, 0.75-0.90 as good, and ≥0.90 as excellent. Moreover, intermethod SEM was reported to express intermethod measurement variability. Two-way random-effects, absolute-agreement ICC(2,1) was used to assess relative inter-rater agreement after a second rater (LH) carried out a blind scoring of the remote 5TSTS video files. Pearson correlation (r) quantified the strength and direction of linear associations between body weight–normalized IKD torque and each 5TSTS measure (IKD vs remote 5TSTS and IKD vs laboratory 5TSTS) to express convergent validity. Where data met assumptions of linearity and normality (assessed using Shapiro-Wilk tests), Pearson r was used; otherwise, the Spearman rank correlation was applied. Correlations were interpreted as r<0.3 (weak), 0.3-0.5 (moderate), and >0.5 (strong), with statistical significance set at P<.05.

An a priori sample size calculation was performed in RStudio (version 2026.01.0+392) for the primary outcome of agreement between the remote and laboratory 5TSTS tests using Bland-Altman analysis. Sample size was determined based on the precision of the estimated 95% limits of agreement. Assuming an SD of paired differences of 3 seconds [5], a sample of 26 participants was required to estimate each limit of agreement, with an approximate 95% CI half-width of ≤2 seconds. This precision-based approach followed recommendations for the design of method comparison studies [31].

Results

Overview

A total of 40 participants were recruited to take part in the study. Of these, 8 (20%) initially agreed but declined to participate due to practicality and scheduling issues that prevented them from attending both the laboratory and remote arms of the trial. A total of 32 (80%) participants attended at least the laboratory-based elements of data collection. However, 1 (2.5%) participant had medical issues in the time between the laboratory visit and remote data collection, which were unrelated to any aspects of the data collection, and therefore could not participate in the remote aspect of the study. A further 2 (5%) participants dropped out after attending the laboratory due to practicality and scheduling issues that prevented them from being able to participate in the remote data collection. Thus, 29 participants were included in the final analysis. Participant characteristics and physical performance are reported in Tables 1 and 2, respectively.

Of the 2 participants who did not continue participation between the laboratory and remote arms of data collection, 1 had technical issues that prevented them from recording the 5TSTS. No further technical issues were reported by participants, and all participants were able to record and export their 5TSTS video

XSL•FO
RenderX

for analysis. Of the 29 participants included, 17 (59%) were Android (Google) device users, and 12 (41%) were iOS (Apple) device users.

Table 1. Demographic characteristics of included participants (N=29).

Characteristics	Participants
Sex, n (%)	
Female	17 (59)
Male	12 (41)
Age (y), mean (SD)	66.9 (6.0)
Body mass (kg), mean (SD)	81.6 (17.9)
Stature (cm), mean (SD)	165.1 (8.5)
BMI (kg/m^2), mean (SD)	29.9 (5.6)
Blood pressure, mean (SD)	
Systolic	147 (17)
Diastolic	81 (11)

Table 2. Physical performance measures and results of laboratory-based and remote 5TSTS assessments and normalized peak torque.

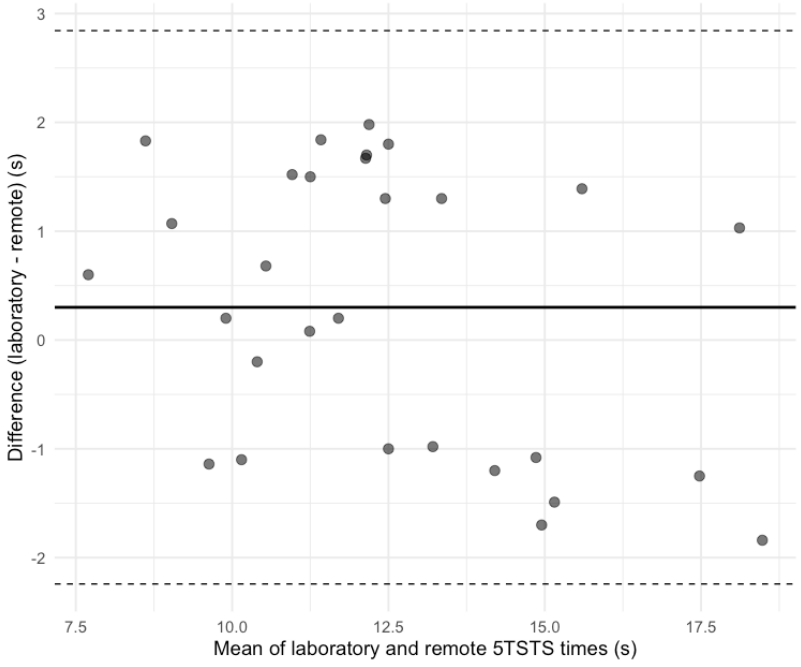
Variables	Values
Laboratory 5TSTS[a] (s), mean (SD)	12.6 (2.6)
Remote 5TSTS (s), mean (SD)	12.3 (3.0)
Normalized peak torque (N m/kg), mean (SD)	1.4 (0.4)

[a]5TSTS: 5 times sit-to-stand.

Validity

Bland-Altman analysis showed that the mean difference (bias) between laboratory-based and remote 5TSTS was 0.3 (95% CI −0.19 to 0.76) seconds, indicating that remote assessment times were, on average, slightly shorter than laboratory-based assessment times. The 95% limits of agreement ranged from −2.24 (95% CI −3.08 to −1.38) seconds to 2.84 (95% CI 1.99 to 3.67) seconds. Figure 1 highlights that there appears to be no obvious visual evidence of proportional bias at low or high values (ie, the spread around the mean bias is consistent across measured values).

Figure 1. Bland-Altman plot for laboratory-based vs remote 5 times sit-to-stand (5TSTS) assessment.

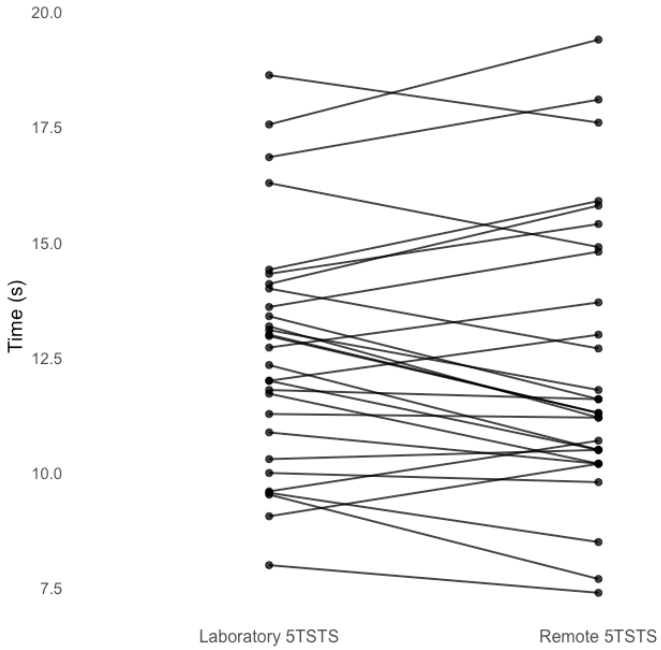


XSL•FO
RenderX

Intermethod Relative Agreement and Intermethod Measurement Variability

The ICC was 0.89 (95% CI 0.81-0.94), indicating good intermethod relative agreement between the remote 5TSTS measurement and the laboratory reference. The SEM was 0.9 seconds, corresponding to an intermethod measurement variability of less than 1 second. These metrics collectively indicate strong relative and absolute agreement between remote and laboratory-based 5TSTS measurements, while also reflecting measurement variability (Figure 2).

Figure 2. Individual-level comparison of laboratory-based and remote 5 times sit-to-stand (5TSTS) performance.



Convergent Validity Correlative Analysis

Pearson correlations revealed a moderate negative relationship between normalized peak torque and laboratory 5TSTS time and between normalized peak torque and remote 5TSTS time, indicating that higher knee extensor strength was associated with faster 5TSTS performance (Table 3). These correlation coefficients were similar in magnitude.

Table 3. Pearson correlation matrix between laboratory and remote 5TSTS[a] (seconds) and normalized peak torque (N m/kg).

	Laboratory 5TSTS	Remote 5TSTS	Normalized peak torque
Laboratory 5TSTS			
r	—[b]	—	—
P value	—	—	—
Remote 5TSTS			
r	0.907	—	—
P value	<.001	—	—
Normalized peak torque			
r	−0.519	−0.518	—
P value	0.004	0.004	—

[a]5TSTS: 5 times sit-to-stand.

[b]Not applicable.

Discussion

Main Findings

The objectives of this study were to evaluate the validity of remote asynchronous 5TSTS assessment from a single trial. An additional objective was to assess the convergent validity of remote and laboratory-based 5TSTS with knee extensor strength. The current study highlighted that both remote and laboratory-based 5TSTS assessments showed promising pragmatic agreement, indicated by a small mean bias and promising limits of agreement of <3 seconds. We interpret these findings to potentially support remote 5TSTS as a valid means for evaluating lower-limb muscle strength and functional mobility in community-dwelling older adults at the group level. However, further research on reliability and responsiveness is

XSL•FO
RenderX

required. We found good intermethod relative agreement between the remote smartphone assessment and the laboratory assessment, as indicated by ICC, and good intermethod measurement validity, as indicated by SEM. Both laboratory and remote 5TSTS showed moderate negative correlations with knee extensor torque.

Validity Analysis

The 5TSTS test is a well-established test of lower-limb muscle strength and functional mobility in community-dwelling older adults [7]. In practice, the findings from the current study could indicate that the mean difference from our Bland-Altman analysis is not clinically important [4], meaning the trivial mean bias between the laboratory and remote 5TSTS observed in this study could potentially support the use of remote smartphone 5TSTS for assessing lower-limb functional mobility in older adults at the group level when further test-retest evaluation has been conducted. The 95% limits of agreement indicated that remote 5TSTS times differed from laboratory-based assessment by −2.24 seconds (lower) and 2.84 seconds (upper). Previous research in older populations has highlighted similar results. For instance, van Schooten et al [16] assessed a variety of physical function tests in laboratory and remote setting (via Zoom) in 37 community-dwelling older adults, reporting lower and upper limits of agreement of more than 2 seconds for the 5TSTS test. Fyfe et al [32] included 19 community-dwelling older adults and assessed the reliability and feasibility of remote physical function testing. They conducted 3 separate tests across 3 videoconferencing calls and highlighted lower and upper limits of agreement of 2.59 and 4.41 seconds between tests 1 and 2, and 3.14 and 3.13 seconds between tests 2 and 3, respectively. These previous findings highlight that agreement between 5TSTS tests in older adults can vary due to several factors, including test familiarity and physical demands [32]. As such, the mean bias and limits of agreement comparing remote and laboratory testing reported within the present study are encouraging.

It must be noted that our sample included a substantial proportion of participants who exceeded the 12-second clinical cut-off and the 15-second threshold associated with an increased risk of falls [12]. Although a more homogeneous sample may have resulted in narrower limits of agreement, it would also have introduced spectrum bias and reduced the clinical relevance of the findings. In practice, the purpose of the test is to identify functional status without prior knowledge of an individual's risk profile; therefore, the heterogeneity of the sample likely widened the limits of agreement but also improved the ecological validity and representativeness of the study.

It should also be noted that, in this protocol, the laboratory-based 5TSTS time was calculated as the mean of 2 attempts, whereas only 1 attempt was scored for the remote 5TSTS assessment to minimize the technical burden on participants associated with recording multiple attempts. Therefore, the reduced measurement variability resulting from averaging 2 attempts in the laboratory may have contributed to the observed bias and limits of agreement.

The findings of the current study are consistent with those of previous studies that used videoconferencing software to remotely assess 5TSTS performance [16,32]. However, the current study helps advance the previous literature by introducing an asynchronous smartphone-based alternative to live videoconferencing assessment, therefore not requiring a researcher at that moment and reducing the resources required for remote assessment. As mentioned, previous apps have used remote 5TSTS to estimate relative muscle power, assess frailty, and store results, showing that smartphone-based 5TSTS testing is clinically feasible. Other studies have attached smartphones to the trunk to collect accelerometer data during 5TSTS, but these methods are technically demanding and less scalable.

Intermethod Relative Agreement and Intermethod Measurement Variability

The intermethod relative agreement data highlighted that the remote measurement of the 5TSTS test showed a moderate to strong agreement with the laboratory assessment. The ICC of 0.89 indicates good intermethod relative agreement [7]. This finding sits at the upper end of the threshold of good intermethod agreement and is close to excellent, according to widely used criteria [33]. The results of the current study indicate an SEM of approximately 0.9 seconds; however, this reflects intermethod measurement variability between laboratory-based and remote 5TSTS assessments and should not be interpreted as remote test-retest error, which requires further investigation. The intermethod relative agreement and measurement variability results in the current study suggest that remote 5TSTS could potentially be suitable for several applications. Measurement theory emphasizes that a specific measure does not need to be perfect to be useful but must be sufficiently stable for the decisions being made from it [34]. In this context, the good relative intermethod relative agreement findings potentially support the utility of remote 5TSTS for evaluating group-level responses to muscle-strengthening interventions in older adults [35]. However, it is important to note that research on reliability and responsiveness is required before this can be applied. This is relevant for mHealth apps and interventions, where repeated remote monitoring may be used to follow progress across a 12- to 16-week program or to compare average changes between groups [36]. As the present study did not examine repeated remote assessments, the measurement error and MDC95 are still to be established. Further research using repeated remote assessments is required to establish test-retest reliability and ability to detect meaningful individual changes.

Convergent Validity

The correlations between knee extensor strength and 5TSTS performance were moderate and negative for both the laboratory and remote 5TSTS assessments (laboratory: r=−0.519; remote: r=−0.518). These results indicate that participants with greater quadriceps torque tended to complete both 5TSTS tests more quickly [37]. This magnitude of association is consistent with the concept that the 5TSTS test is a functional proxy for lower-limb muscle strength in older adults rather than a "pure" strength measure. The performance of the 5TSTS test in older adults reflects force production but also other factors such as balance, coordination, movement strategy, and familiarity or confidence [38]. Previous work by Yee et al [10] in community-dwelling older adults confirmed that the

XSL•FO
RenderX

performance of the 5TSTS test is a functional maneuver that relies in part on the strength of the knee extensors and also on the stability of the trunk, sensation, and balance. Ruiz-Cárdenas et al [39] validated a smartphone sit-to-stand app as a diagnostic tool for sarcopenia in older adults. This work considered the velocity of hip extension as a key factor in the performance of the 5TSTS, indicating that while the ability of the knee extensor muscles to produce required force is essential, their ability to produce this force quickly (power) is also a key determinant [40]. In relation to the study's current findings, the moderate strength of correlations suggests that knee extensor strength makes an important but not exclusive contribution to 5TSTS performance in community-dwelling older adults. Encouragingly, the remote assessment *r* was analogous to the laboratory assessment *r*.

Strengths and Limitations

The current study exhibits several strengths, including preregistration, which enhances transparency. Additionally, the precision-based sample size calculation undertaken a priori ensures our conclusions are robust. The sample therefore provides adequate precision of the Bland-Altman limits of agreement, which is consistent with recommendations for method comparison studies [31]. In addition, we recruited a sample exceeding the minimum number, which further supports the adequacy of the study for its primary analysis.

An additional important strength is that participants performed the 5TSTS test at home, which may have reduced apprehension and encouraged a more natural testing environment where the participant could proceed at their own pace without external pressures [25]. The laboratory setting may be perceived as more formal and more intimidating [41]. Thus, the remote 5TSTS results may be considered more ecologically valid.

Every participant in the sample was considered an older adult (aged ≥60 years), with an age range from 60 to 80 years. While this could result in differing physical abilities across these ages, widening the limits of agreement, this enhances the validity and applicability of the present study, as opposed to examining agreement in younger adults. Our participants differed in exercise training experience, which could have led to differences in strength and power, balance and postural control, joint pain and stiffness, and confidence [28]. When conditions of the tests were changed (laboratory vs remote), the aforementioned factors may be exacerbated. For example, a trained 60-year-old individual may perform the 5TSTS test almost identically in both laboratory and remote settings; however, a frail 80-year-old individual with marginal balance might be more cautious in one setting, in this case the laboratory (unknown environment), but not at home, where they were comfortable [41]. We consider the rigorous assessment of agreement between laboratory-based and remote 5TSTS assessments a strength of the present study.

The study was not without limitations. First, as the data collection required both laboratory and remote 5TSTS assessments, there was an inherent variability between the laboratory environment and participants' homes. Participants received clear instructions on completing and recording the remote 5TSTS assessment to replicate the laboratory conditions as closely as possible. Because this was a pragmatic evaluation of real-world 5TSTS scores, chair height, chair characteristics, floor stability, number of trials, and timing procedures were not analogous to laboratory testing. However, this was necessary to collect data from participants in real-world settings. Therefore, the remote test should be interpreted as a pragmatic home-based measure, not a standardized laboratory equivalent. Most notably, the laboratory assessment was the mean of 2 trials, and the remote assessment was only 1 trial, which may contribute to between-method variability. That said, the approach reflected real-world conditions in which the app will be deployed. Requiring participants to use a chair of a fixed height at home or to source, measure, and precisely adjust a chair to a specified height would increase participant burden, reduce adherence, and be impractical for older adults.

Importantly, recruitment potential was limited by travel burden, as participants had to travel to the university campus to complete the laboratory-based data collection. This limited recruitment to those who lived near the campus or had the means to access it, meaning the sample may not be entirely representative. However, results include scores ranging between good and below thresholds, suggesting a range of abilities.

It must be noted that laboratory assessment always preceded remote assessment; order and familiarization effects cannot be excluded. The small tendency toward faster remote effects may therefore reflect learning, environmental comfort, or other environmental factors rather than the protocol alone, and future studies should consider counterbalancing the order of assessments.

It is important to note that when interpreting the findings of this study, the mean age of the sample was 66.9 (6.0) years. While age ranged from 60 to 80 years, participants were community dwelling and mainly healthy. Therefore, the findings may not apply to older adults aged >80 years or to those with frailty, chronic conditions, and limited digital experience. As such, this should be a key focus of future research.

Finally, the remote arm of the data collection required manual scoring for the 5TSTS tests. Although manageable for a sample of 29 participants, this approach may impose a substantial researcher burden in studies involving larger sample sizes.

Future Research

Given the current study's promising findings, there are avenues for future research to build on these findings. First, research is needed to explore the automation of remote 5TSTS assessment in a cost-effective manner. One way to do this is to investigate AI for video analysis in remote assessment. This would permit 5TSTS to be automatically scored and categorized upon uploading, reducing researcher burden and resulting in real-time feedback. Furthermore, larger scale mHealth muscle-strengthening interventions are needed to provide longitudinal data.

As mentioned, future research in this area should also seek to address the usability of remote 5TSTS assessment to ensure that this strength assessment type does not lead to digital exclusion. Future trials may benefit from investigating usability and acceptability of smartphone-based 5TSTS assessment in older adults with lower digital literacy. One avenue that has

XSL•FO
RenderX

proven successful when it comes to digital tools in this population is co-design, and well-designed trials should inform adaptations to digital platforms to ensure a smooth, user-designed platform [21]. In turn, the above future direction would help position remote, smartphone-based strength assessment as a core component of a broader digital strategy for the support and promotion of muscle-strengthening exercise in older adults.

Conclusions

The current study demonstrates that remote smartphone-based 5TSTS assessment provides promising pragmatic agreement with laboratory assessment under controlled conditions in community-dwelling older adults, supporting its potential future application for group-level evaluation once future reliability and responsiveness research has been conducted. Our data suggest that remote 5TSTS performance provides comparable estimates of lower-limb functional performance to laboratory assessment and preserves the observed relationship between 5TSTS performance and lower-limb strength. These findings could potentially support remote 5TSTS assessment as a practical and scalable alternative to laboratory-based assessment for monitoring lower-limb function in older adults at the group level in the future. However, as mentioned, future research is required to confirm test-retest reliability and responsiveness before this conclusion can be applied. The study has highlighted good intermethod relative agreement but has not yet established interchangeability for monitoring change at the individual level. As repeated remote assessments were not examined, further research is required to establish test-retest reliability, responsiveness to longitudinal change, and the ability of remote 5TSTS assessment to detect meaningful individual changes over time.

Data Availability

The dataset analyzed during this study is available in the OSF repository [22].

Funding

The authors declared no financial support was received for this work.

Generative AI Use

The authors declare that generative AI was not used in the preparation of this manuscript.

Authors' Contributions

Conceptualization: ECJB (lead), LDH (equal), AW (equal), NFS (equal)
Data curation: ECJB
Formal analysis: ECJB (lead), LDH (supporting), AW (supporting)
Funding acquisition: LDH
Investigation: ECJB
Methodology: ECJB, LDH, AW, NFS
Project administration: ECJB
Resources: ECJB
Supervision: LDH (equal), AW (equal), NFS (equal)
Validation: ECJB
Visualization: ECJB (lead), LDH (supporting), AW (supporting), NFS (supporting)
Writing – original draft: ECJB (lead), LDH (supporting), AW (supporting), NFS (supporting)
Writing – review & editing: ECJB (lead), LDH (supporting)

Conflicts of Interest

In the interest of complete transparency, the lead author (EB) was involved in the creation, design, and development of the Strength Buddy mobile app as part of their wider PhD project alongside the other authors of this manuscript. The app exists as a test prototype and currently has no commercial value. The submitted work was not carried out in the presence of any personal, professional, or financial relationships that could potentially be construed as conflicts of interest.

References

1. Csuka M, McCarty DJ. Simple method for measurement of lower extremity muscle strength. Am J Med. Jan 1985;78(1):77-81. [doi: 10.1016/0002-9343(85)90465-6] [Medline: 3966492]
2. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. 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. Aug 2002;57(8):M539-M543. [doi: 10.1093/gerona/57.8.m539] [Medline: 12145369]
3. De Melo TA, Silva Guimarães F, Lapa E Silva JR. The five times sit-to-stand test: safety, validity and reliability with critical care survivors's at ICU discharge. Arch Physiother. Dec 18, 2022;13(1):2. [FREE Full text] [doi: 10.1186/s40945-022-00156-z] [Medline: 36597159]

XSL•FO
RenderX

4. Albalwi AA, Alharbi AA. Optimal procedure and characteristics in using five times sit to stand test among older adults: a systematic review. Medicine (Baltimore). Jun 30, 2023;102(26):e34160. [FREE Full text] [doi: 10.1097/MD.0000000000034160] [Medline: 37390277]
5. Steffens D, Pocovi NC, Bartyn J, Delbaere K, Hancock MJ, Koh C, et al. Feasibility, reliability, and safety of remote five times sit to stand test in patients with gastrointestinal cancer. Cancers (Basel). Apr 24, 2023;15(9):2434. [FREE Full text] [doi: 10.3390/cancers15092434] [Medline: 37173899]
6. Porto JM, Peres-Ueno MJ, de Matos Brunelli Braghin R, Scudilio GM, de Abreu DC. Diagnostic accuracy of the five times stand-to-sit test for the screening of global muscle weakness in community-dwelling older women. Exp Gerontol. Jan 2023;171:112027. [FREE Full text] [doi: 10.1016/j.exger.2022.112027] [Medline: 36395978]
7. Muñoz-Bermejo L, Adsuar JC, Mendoza-Muñoz M, Barrios-Fernández S, Garcia-Gordillo MA, Pérez-Gómez J, et al. Test-retest reliability of five times sit to stand test (FTSST) in adults: a systematic review and meta-analysis. Biology (Basel). Jun 09, 2021;10(6):510. [FREE Full text] [doi: 10.3390/biology10060510] [Medline: 34207604]
8. Hayes LD, Elliott BT, Yasar Z, Bampouras TM, Sculthorpe NF, Sanal-Hayes NE, et al. High intensity interval training (HIIT) as a potential countermeasure for phenotypic characteristics of sarcopenia: a scoping review. Front Physiol. Aug 24, 2021;12:715044. [FREE Full text] [doi: 10.3389/fphys.2021.715044] [Medline: 34504439]
9. Alcazar J, Losa-Reyna J, Rodriguez-Lopez C, Alfaro-Acha A, Rodriguez-Mañas L, Ara I, et al. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Exp Gerontol. Oct 02, 2018;112:38-43. [doi: 10.1016/j.exger.2018.08.006] [Medline: 30179662]
10. Yee XS, Ng YS, Allen JC, Latib A, Tay EL, Abu Bakar HM, et al. Performance on sit-to-stand tests in relation to measures of functional fitness and sarcopenia diagnosis in community-dwelling older adults. Eur Rev Aging Phys Act. Jan 08, 2021;18(1):1. [FREE Full text] [doi: 10.1186/s11556-020-00255-5] [Medline: 33419399]
11. Søgaard D, Lund MT, Scheuer CM, Dehlbaek MS, Dideriksen SG, Abildskov CV, et al. High-intensity interval training improves insulin sensitivity in older individuals. Acta Physiol (Oxf). Apr 2018;222(4):e13009. [doi: 10.1111/apha.13009] [Medline: 29197155]
12. Makizako H, Shimada H, Doi T, Tsutsumimoto K, Nakakubo S, Hotta R, et al. Predictive cutoff values of the five-times sit-to-stand test and the timed "up and go" test for disability incidence in older people dwelling in the community. Phys Ther. Apr 01, 2017;97(4):417-424. [FREE Full text] [doi: 10.2522/ptj.20150665] [Medline: 28371914]
13. van Cappellen-van Maldegem SJ, Hoedjes M, Seidell JC, van de Poll-Franse LV, Buffart LM, Mols F, et al. Self-performed five times sit-to-stand test at home as (pre-)screening tool for frailty in cancer survivors: reliability and agreement assessment. J Clin Nurs. Apr 2023;32(7-8):1370-1380. [doi: 10.1111/jocn.16299] [Medline: 35332600]
14. Trivedi P, Patel S, Edwards G, Jenkins T, Man WD, Nolan CM. Five-repetition sit-to-stand test: responsiveness and minimal important difference in idiopathic pulmonary fibrosis. Ann Am Thorac Soc. Apr 2024;21(4):577-584. [FREE Full text] [doi: 10.1513/AnnalsATS.202306-561OC] [Medline: 37847730]
15. Puthoff ML, Saskowski D. Reliability and responsiveness of gait speed, five times sit to stand, and hand grip strength for patients in cardiac rehabilitation. Cardiopulm Phys Ther J. Mar 2013;24(1):31-37. [FREE Full text] [Medline: 23754937]
16. van Schooten KS, Steffens D, Engeler A, Letton M, Brown A, Perram A, et al. Remote assessment of physical function in older people: feasibility, safety and agreement with in-person administration. Age Ageing. Aug 29, 2025;54(9):afaf266. [FREE Full text] [doi: 10.1093/ageing/afaf266] [Medline: 41014506]
17. Alcazar J, Lopez-Moranchel I, Losa-Reyna J, Baltasar-Fernandez I, Garcia-Garcia FJ, Alegre LM, et al. App Review: PowerFrail for the assessment of relative muscle power and frailty in older people. BJSM Blog. 2021. URL: https://blogs.bmj.com/bjsm/2021/02/04/app-review-powerfrail-for-the-assessment-of-relative-muscle-power-and-frailty-in-older-people/#_ENREF_2 [accessed 2026-11-13]
18. Cerrito A, Bichsel L, Radlinger L, Schmid S. Reliability and validity of a smartphone-based application for the quantification of the sit-to-stand movement in healthy seniors. Gait Posture. Feb 2015;41(2):409-413. [doi: 10.1016/j.gaitpost.2014.11.001] [Medline: 25467428]
19. Hayek R, Gutman I, Baranes G, Nudelman Y, Springer S. Smartphone-based sit-to-stand analysis for mobility assessment in middle age. Innov Aging. Aug 28, 2024;8(10):igae079. [FREE Full text] [doi: 10.1093/geroni/igae079] [Medline: 39391811]
20. Marques DL, Neiva HP, Pires IM, Zdravevski E, Mihajlov M, Garcia NM, et al. An experimental study on the validity and reliability of a smartphone application to acquire temporal variables during the single sit-to-stand test with older adults. Sensors (Basel). Mar 15, 2021;21(6):2050. [FREE Full text] [doi: 10.3390/s21062050] [Medline: 33803927]
21. Berry EC, Sculthorpe NF, Warner A, Sanal-Hayes NE, Hayes LD. Designing an mHealth app to encourage uptake of muscle-strengthening exercise in older adults: co-design focus group study. JMIR Aging. Mar 12, 2026;9:e87332. [FREE Full text] [doi: 10.2196/87332] [Medline: 41818741]
22. Berry E, Hayes LD, Sculthorpe N, Warner A. Supporting and tracking muscle strengthening exercise in older adults: a validity and feasibility trial. Open Science Framework. 2025. URL: https://osf.io/xqt52/overview [accessed 2026-11-13]
23. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. Jan 2011;64(1):96-106. [doi: 10.1016/j.jclinepi.2010.03.002] [Medline: 21130355]

XSL•FO
RenderX

24. Vilarinho R, Montes AM, Melo C. Age-related influence on reliability and learning effect in the assessment of lower limb strength using sit-to-stand tests: a cross-sectional study. Health Sci Rep. May 20, 2024;7(5):e2064. [FREE Full text] [doi: 10.1002/hsr2.2064] [Medline: 38774592]
25. Hayes LD, Morse CI. The effects of progressive dehydration on strength and power: is there a dose response? Eur J Appl Physiol. Mar 2010;108(4):701-707. [doi: 10.1007/s00421-009-1288-y] [Medline: 19908058]
26. Bloch-Ibenfeldt M, Theil Gates A, Karlog K, Demnitz N, Kjaer M, Boraxbekk CJ. Heavy resistance training at retirement age induces 4-year lasting beneficial effects in muscle strength: a long-term follow-up of an RCT. BMJ Open Sport Exerc Med. Jun 18, 2024;10(2):e001899. [FREE Full text] [doi: 10.1136/bmjsem-2024-001899] [Medline: 38911477]
27. Guadalupe-Grau A, Aznar-Laín S, Mañas A, Castellanos J, Alcázar J, Ara I, et al. Short- and long-term effects of concurrent strength and HIIT training in octogenarians with COPD. J Aging Phys Act. Jan 2017;25(1):105-115. [doi: 10.1123/japa.2015-0307] [Medline: 27402660]
28. Hayes LD, Sanal-Hayes NE, Mclaughlin M, Berry EC, Sculthorpe NF. People with long COVID and ME/CFS exhibit similarly impaired balance and physical capacity: a case-case-control study. Am J Med. Jan 2025;138(1):140-147. [FREE Full text] [doi: 10.1016/j.amjmed.2023.06.028] [Medline: 37490948]
29. DaVinci resolve 21. Blackmagic Design. URL: https://www.blackmagicdesign.com/uk/products/davinciresolve/ [accessed 2026-11-13]
30. RStudio IDE user guide. Posit Software. URL: https://posit.co/download/rstudio-desktop/ [accessed 2026-11-13]
31. Monti CB, Ambrogi F, Sardanelli F. Sample size calculation for data reliability and diagnostic performance: a go-to review. Eur Radiol Exp. Jul 05, 2024;8(1):79. [FREE Full text] [doi: 10.1186/s41747-024-00474-w] [Medline: 38965128]
32. Fyfe JJ, Sousa F, Watson-Mackie K, Jansons P, Scott D, Daly RM. Remote physical function testing in older adults: a mixed methods study exploring test reliability, feasibility, and perceptions of participants and assessors. PLoS One. Sep 19, 2025;20(9):e0332691. [FREE Full text] [doi: 10.1371/journal.pone.0332691] [Medline: 40971978]
33. Benhida M, Zeghari L, El Mouahid K, El Morchidy S, El Madhi Y, Enneya N, et al. Physical performance profiling in Moroccan women's football: a functional assessment from the Rabat-Salé-Kénitra regional league. BMC Sports Sci Med Rehabil. Dec 03, 2025;18(1):13. [FREE Full text] [doi: 10.1186/s13102-025-01453-3] [Medline: 41331491]
34. Dimitrov D, Rumrill P, Fitzgerald S, Hennessey M. Reliability in rehabilitation measurement. Work. Mar 2001;16(2):159-164. [doi: 10.3233/wor-2001-00148]
35. Takimoto K, Takebayashi H, Kondo H, Ikeda K. Association between muscle quality assessed by the 5-repetition sit-to-stand test and falls in community-dwelling older adults in Japan: a cross-sectional study. Geriatrics (Basel). Jun 07, 2025;10(3):78. [FREE Full text] [doi: 10.3390/geriatrics10030078] [Medline: 40558617]
36. Jungreitmayr S, Kranzinger C, Venek V, Ring-Dimitriou S. Effects of an app-based physical exercise program on selected parameters of physical fitness of females in retirement: a randomized controlled trial. Front Physiol. Mar 4, 2022;13:821773. [FREE Full text] [doi: 10.3389/fphys.2022.821773] [Medline: 35317213]
37. Steinbrink GM, Martinez J, Swartz AM, Strath SJ. Sit-to-stand power is a stronger predictor of gait speed than knee extension strength. J Funct Morphol Kinesiol. Jun 13, 2024;9(2):103. [FREE Full text] [doi: 10.3390/jfmk9020103] [Medline: 38921639]
38. Niklasson J, Fagerström C, Backåberg S, Bergman P, Lindberg T. Exploring predictors of the five-time sit-to-stand test based on cross-sectional findings from the Swedish National Study on Aging and Care (SNAC). BMC Geriatr. Feb 04, 2025;25(1):79. [FREE Full text] [doi: 10.1186/s12877-025-05737-8] [Medline: 39905293]
39. Ruiz-Cárdenas JD, Montemurro A, Martínez-García MD, Rodríguez-Juan JJ. Sit-to-stand video analysis-based app for diagnosing sarcopenia and its relationship with health-related risk factors and frailty in community-dwelling older adults: diagnostic accuracy study. J Med Internet Res. Dec 08, 2023;25:e47873. [FREE Full text] [doi: 10.2196/47873] [Medline: 38064268]
40. Mato L, Chankavee N, Amatachaya S, Thaweewannakij T. Knee extensor muscle strength to measure the ability of five times sit to stand independently in patients with incomplete spinal cord injury. Malays J Med Sci. Oct 2022;29(5):74-82. [FREE Full text] [doi: 10.21315/mjms2022.29.5.8] [Medline: 36474537]
41. Nebeker C, Zlatar ZZ. Learning from older adults to promote independent physical activity using mobile health (mHealth). Front Public Health. Aug 12, 2021;9:703910. [FREE Full text] [doi: 10.3389/fpubh.2021.703910] [Medline: 34476229]

Abbreviations

5TSTS: 5 times sit-to-stand
ICC: intraclass correlation coefficient
IKD: isokinetic dynamometer
MDC95: minimal detectable change at the 95% CI
mHealth: mobile health
MID: minimally important clinical difference
OSF: Open Science Framework

XSL•FO
RenderX

Edited by KJ Moon; submitted 30.Jul.2026; peer-reviewed by PR Venna, J Loureiro, B Unver; comments to author 14.Sep.2026; revised version received 10.Nov.2026; accepted 10.Nov.2026; published 23.Nov.2026

Please cite as:
Berry ECJ, Hayes LD, Warner A, Sculthorpe NF
Remote Asynchronous Smartphone 5 Times Sit-to-Stand Testing in Older Adults: Pragmatic Validity Trial
JMIR Aging 2026;9:e108345
URL: https://aging.jmir.org/2026/1/e108345
doi: 10.2196/108345
PMID:

©Ethan CJ Berry, Lawrence D Hayes, Ashley Warner, Nicholas F Sculthorpe. Originally published in JMIR Aging (https://aging.jmir.org), 23.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX