Feasibility, reliability, and validity of the two-minute step test for assessing aerobic exercise capacity and functional endurance in participants with anterior cruciate ligament reconstruction

Caner Karartia, Fatih Özyurta, Ömer Alperen Gürsesa, Hakkı Çağdaş Basatb, İsmail Özsoyc, Sümeyye Erola and Kubra Caylan Gursesd

The aim of this study was to investigate the feasibility, test–retest reliability, and construct validity of the 2-min step test (2MST) in measuring aerobic exercise capacity and functional endurance in participants with anterior cruciate ligament reconstruction (ACLR). Fifty participants with quadriceps tendon autograft, bone-patellar tendon-bone graft, or hamstring autograft were included. Feasibility was assessed by the time and support required to complete the measures and the feedback from participants, including their satisfaction ratings. Test–retest reliability was assessed using the intraclass correlation coefficients ($ICC_{2,1}$), a Bland–Altman plot with 95% limits of agreement (LoA), SEM, SEM%, and minimum detectable change (MDC_{95}). The participants were assessed with the visual analogue scale (VAS), Lysholm Knee Scoring Scale, Knee Injury and Osteoarthritis Outcome Score (KOOS), stair climbing test (SCT), and 6-min walk test to assess construct validity. The test took less than 5 min to describe and perform, and participants required minimal verbal support. The satisfaction rate was notably high. The $ICC_{2,1}$ was 0.98 (0.96–0.99). SEM and MDC_{95} were 2.96 and 8.20 (2.41–14.00), respectively. The SEM% of 2.99 was considered to be at a very good level. The Bland–Altman plot illustrates that the 2MST scores had a 95% LoA ranging from −21.52 to 2.68 steps. The SCT score and VAS rest score were found to be associated with the 2MST performance, explaining 42.5% of the variance. The 2MST is a feasible, reliable, and valid test to measure aerobic exercise capacity and functional endurance in participants with ACLR. *International Journal of Rehabilitation Research* **47: 260–267**

International Journal of Rehabilitation Research 2024, **47**:260–267

Keywords: 2-min step test, anterior cruciate ligament reconstruction, physical functional performance, rehabilitation

[a]Department of Physiotherapy and Rehabilitation, [b]Department of Orthopedics and Traumatology, Kırşehir Ahi Evran University, Kırşehir, [c]Department of Physiotherapy and Rehabilitation, Selçuk University, Konya, and [d]Department of Physiotherapy and Rehabilitation, Hacettepe University, Ankara, Turkey

Correspondence to Caner Karartı, PT, PhD, Department of Physiotherapy and Rehabilitation, Kırşehir Ahi Evran University, Bağbaşı Street, 40100 Kırşehir, Turkey
Tel: +90 553 223 5030; e-mail: fzt.caner.92@gmail.com

Received 5 August 2024 **Accepted** 31 August 2024.

Introduction

Anterior cruciate ligament reconstruction (ACLR) is a commonly used procedure for treating anterior cruciate ligament (ACL) injuries [1]. Strength testing, self-reported knee function, and performance-based functional tests are widely used to assess the functional performance level of participants [2]. There is a considerable variability in the use of the most appropriate performance-based test during the assessment of physical performance [3]. Single- and triple-leg hop tests, crossover hop tests, 6-m timed hop test, isokinetic tests, functional movement screen, 6-min walk test (6MWT), stair climbing test (SCT), etc. are generally used [3,4].

There are studies investigating the potential value of the 2-min step test (2MST) in various populations, including those with low back pain [5], knee osteoarthritis [6], total knee arthroplasty (TKA) [7], obesity [8], and peripheral arterial disease [9]. The 2MST is an easier and more time-efficient tool that does not require a specialized equipment or a large space compared to the 6MWT [10]. The test is designed to assess aerobic exercise capacity and functional endurance. This is achieved by counting the number of times the individual lifts their knees to the specified height within the 2-min period [10]. It offers valuable insight into an individual's capacity to sustain physical activity over an extended period in terms of functional endurance. The 2MST is supported by content validity, which is demonstrated by its capacity to reflect aspects of aerobic capacity and functional endurance. This is based on the assumption that maintaining a consistent pace while stepping in place for a period of 2 min is indicative of the endurance and cardiovascular demands typically encountered in daily activities [10]. It has been demonstrated that stepping in place represents a valid method for assessing the endurance and cardiovascular capabilities that are relevant for evaluating functional health in a range of populations [5–10].

The 2MST has the potential to serve as a valuable tool for the assessment of aerobic capacity, functional endurance, and progress tracking in participants with ACLR.

 DOI: 10.1097/MRR.0000000000000643

The test assesses the participant's capacity to engage in a prolonged, high-intensity, weight-bearing activity, which is indicative of their cardiovascular fitness and lower extremity endurance [5–10]. This is of particular importance for those undergoing ACLR, as the recovery process necessitates the restoration of both strength and endurance [1]. Moreover, the test can be used to track functional improvements and the effectiveness of rehabilitation over time. The test, which encompasses the relevant parameters, can be conducted repeatedly throughout the rehabilitation process to track changes in the individual over time. The act of omitting this test could result in an incomplete assessment, a delayed return to preinjury activity levels, and a potential overemphasis on the strength of the participant's functional capacity. This could lead to the neglect of deficits in endurance and functional performance that are crucial for a successful return to sports or other high-intensity activities. The test could facilitate the tracking of progress, the adjustment of treatment plans, and the determination of readiness for high-demand activities.

A more detailed evaluation, however, may be required when conducting 2MST in individuals with ACLR in comparison to those with knee osteoarthritis and TKA [6,7]. Due to the necessity of a sustained stepping motion over a defined distance, the 2MST may prove challenging for ACL participants experiencing pain, swelling, or range of motion limitations [1]. It is essential to consider the timing of test administration in terms of the reliability and validity of the test, given that these symptoms have the potential to negatively impact stepping performance during testing [1]. The 2MST is an appropriate instrument for individuals with knee osteoarthritis and TKA, as it assesses the cumulative impact of joint degeneration and the individual's level of functional endurance [6,7]. Nevertheless, the principal objective of surgical intervention in ACLR is to eliminate ligamentous instability, rather than to address joint degeneration. Individuals with ACLR frequently encounter challenges such as altered gait mechanics, compensatory movement patterns, muscle imbalances, and proprioceptive deficits [2]. These problems may negatively impact their performance during 2MST in comparison to individuals with osteoarthritis or TKA. The 2MST, which is a variation of the vertical and catalytic tests, may be less sensitive to parameters such as power asymmetries, dynamic stability, or neuromuscular control [10]. Consequently, additional and comprehensive assessments may be required to determine functional improvements and limitations in ACLR participants. The combination of 2MST with other functional and individual-specific tests may be an appropriate approach to optimize the assessment of participants.

Therefore, the aim of the present study was to investigate the feasibility, test–retest reliability, and construct validity of the 2MST in participants with ACLR. The secondary aim of this study was to investigate the relationship between the 2MST performance and clinical variables. It was hypothesized that the 2MST would be a feasible, reliable, and valid test to use in participants with ACLR.

Methods

Fifty participants with quadriceps tendon autograft ($n = 15$), bone-patellar tendon-bone graft ($n = 10$), or hamstring autograft ($n = 25$) and completed a total of 12 weeks of physiotherapy program were included. The diagnosis of an ACL tear was based on an MRI and clinical assessment by an orthopedic surgeon. ACLR was performed by an orthopedic surgeon. A 12-week long, three-phase structured rehabilitation program based on the recommendations of Kotsifaki *et al*. [11] was carried out by an experienced physiotherapist. The techniques for applying the exercises in the program were arranged according to the graft type. Program consisted of cryotherapy, neuromuscular electrical stimulation, kinesio taping, gradual weight-bearing, open kinetic chain exercises, closed kinetic chain exercises, isometric quadriceps exercises including static quadriceps contractions and straight leg raises, leg press, quadriceps eccentric strengthening using eccentric cycle between 20° and 60° of knee flexion, lumbopelvic motor control exercises, and plyometric and agility training [11]. Participants who were still being treated in our clinic were expected to complete treatment. Those who had completed 12th week and were already discharged from our clinic within the last 9 months, however, were also invited to participate in the study. The local ethics committee approved the study protocol (2023/1221). Written and verbal informed consent was obtained from all participants. The study protocol was conducted in accordance with the COSMIN guidelines [12].

Inclusion criteria were as follows: participants over 18 years old, with a BMI between 18.5 and 24.9 kg/m^2, who underwent ACLR following unilateral ACL rupture, with a complete range of motion, and were between 3 and 12 months postsurgery [13]. Exclusions were as follows: reinjury, ACLR revision surgery, current ligament and/or meniscus injury, patellofemoral pain, muscle injury in the previous 6 months, respiratory or cardiovascular diseases [13].

Outcome measures

All tests were carried out by a physiotherapist. The 2MST required participants to march in place for 2 min as quickly as possible while raising their knees to a level halfway between the patella and the iliac crest in a standing position [14]. Prior to the start of the examination, the assessor placed a mark on the wall with the tape at the mid-point between the iliac crest and the mid-point of the patella. To help participants to understand the test,

Fig. 1

2-min step test.

they completed a 30-s trial [7]. During the test, the examiner remained beside the participants to ensure their safety and provided them with verbal instructions [7]. Performance was defined as the number of steps taken on the right side to reach the criterion height in 2 min [14]. The 2MST was administered twice using the same procedure, with a 7-day interval, to assess test–retest reliability (Fig. 1).

The feasibility of the study was evaluated based on the time and support required to complete the measures, as well as the feedback from participants, including their satisfaction ratings on a scale of 0–10 [15].

Construct validity

In accordance with the COSMIN guidelines, construct validity could be defined as the extent to which the

scores obtained from a PROM are consistent with the pre-established hypotheses. These hypotheses may pertain to internal relationships, correlations with scores from other instruments, or differences between relevant groups. The underlying assumption is that the PROM validly measures the construct in question.

In accordance with the COSMIN guidelines, it is recommended that hypotheses be formulated in advance with a specific and clearly defined direction, magnitude, and rationale [16]. The initial hypotheses were developed by the first and last authors, respectively. Following a process of discussion, a consensus was reached regarding the formulation of 11 independent hypotheses. The formulated hypotheses included a clear definition of the anticipated correlation direction and magnitude, as well as an explicit rationale for each hypothesis. Table 1 presents the initial hypotheses. The construct validity rating for the 2MST was based on the total number of confirmed hypotheses, with ratings of high (equivalent to or greater than 75%), moderate (equivalent to or greater than 50% and less than 75%), or poor (less than 50%) assigned accordingly [16].

The participants were evaluated with the following tests and scales in terms of construct validity in the first session. A hypothesis testing approach was employed to further substantiate the construct validity of the test in accordance with the COSMIN guidelines [16]. Visual analog scale (VAS) for the assessment of pain (rest, activity, night) [17]; Lysholm Knee Scoring Scale and Knee injury and Osteoarthritis Outcome Score (KOOS) for the participant-reported symptoms and function [18]; SCT for the assessment of functional strength, balance, and agility [4]; and 6MWT for the aerobic capacity and endurance [19].

Sample size and statistics

The power analysis was conducted using the COSMIN guidelines. Based on the model (two-way mixed-effects model), type (multiple measures), and definition (absolute agreement) of the intraclass correlation coefficients (ICC), the minimum required sample size for adequate power was calculated and 50 participants were included [12].

The IBM SPSS Statistics for Windows software (ver. 22.0; IBM Corp., New York) was used. Descriptive statistics were expressed as mean (SD) for continuous variables and ratios (%) for categorical variables. Test–retest reliability of the 2MST was assessed using the $ICC_{2,1}$ with 95% confidence intervals (CIs) [7,20]. The ICC was defined as a value between 0 and 1, with values below 0.5 indicating poor reliability, values between 0.5 and 0.75 indicating moderate reliability, values between 0.75 and 0.9 indicating good reliability, and values above 0.9 indicating excellent reliability [7,20]. To examine the visual agreement between the first and second 2MST assessments, a Bland–Altman plot was created [7]. The 95% CIs and 95% limits of agreement (LoA) were shown for the differences between the two measurements [7]. SEM and minimum detectable change (MDC) were calculated [7]. The formula for estimating the SEM is $SD \times \sqrt{(1 - ICC)}$. Here, SD refers to the SD obtained from the first assessment [7,20]. The SEM% was reported by calculating the SEM divided by the mean of the 2MST test and retest scores [7]. The SEM% values were interpreted using the procedures of Akkan *et al.* [7]. The MDC_{95} threshold value is calculated by multiplying the SEM by $1.96 \times \sqrt{2}$, where 1.96 corresponds to the 95% CI and $\sqrt{2}$ represents the measurement error associated with two measurements [7,20].

Table 1 **Initial hypotheses for the assessment of construct validity**

Hypothesis	Rationale	Correlation expected
1. A moderate negative correlation is to be expected between the 2MST performance and VAS rest score	Higher levels of pain at rest may be associated with a deterioration in performance during sustained stepping	0.3–0.5
2. A moderate negative correlation is to be expected between the 2MST performance and VAS activity score	Higher levels of pain during activity may be associated with a deterioration in performance during sustained stepping	0.3–0.5
3. A moderate negative correlation is to be expected between the 2MST performance and VAS night score	Higher levels of pain at night may be associated with a deterioration in performance during sustained stepping	0.3–0.5
4. A strong positive correlation is to be expected between the 2MST performance and 6MWT performance	The 2MST and 6MWT are both designed to assess the same underlying construct, namely aerobic exercise capacity	>0.5
5. A strong negative correlation is to be expected between the 2MST performance and SCT performance	The SCT assesses the participant's functional strength, balance, and agility. These are interrelated with functional endurance assessed by the 2MST	>0.5
6. A strong positive correlation is to be expected between the 2MST performance and KOOS pain score	The presence of low levels of pain during functional activities is associated with higher levels of performance in sustained stepping	>0.5
7. A moderate positive correlation is to be expected between the 2MST performance and KOOS symptoms score	The presence of low levels of symptom severity during functional activities is associated with higher levels of performance in sustained stepping	0.3–0.5
8. A moderate positive correlation is to be expected between the 2MST performance and KOOS ADL score	The presence of low levels of difficulty during activities of daily living is associated with higher levels of performance in sustained stepping	0.3–0.5
9. A strong positive correlation is to be expected between the 2MST performance and KOOS Sports score	The presence of low levels of difficulty during sport and recreation function is associated with higher levels of performance in sustained stepping	>0.5
10. A moderate positive correlation is to be expected between the 2MST performance and KOOS QoL score	The presence of higher levels of knee-related quality of life is associated with higher levels of performance in sustained stepping	0.3–0.5
11. A moderate positive correlation is to be expected between the 2MST performance and Lysholm score	The improvement in participant-reported symptoms is associated with higher levels of performance in sustained stepping	0.3–0.5

2MST, 2-min step test; 6MWT, 6-min walk test; ADL, activities of daily living; KOOS, Knee injury and Osteoarthritis Outcome Score; QoL, knee-related quality of life; SCT, Stair climbing test; Sports, Sport and recreation function; VAS, visual analog scale.

Table 2 Descriptive and clinical characteristics of the participants

Variable	Participants with ACLR, mean (SD)
Age (years)	29.9 (10.9)
BMI (kg/m²)	23.3 (2.9)
Female (%)	14 (28.0%)
Duration of symptoms (months)	6.1 (4.5)
Dominant side (right %)	47 (94%)
Injured side (right %)	26 (52%)
$2MST-_1$	93.2 (22.0)
$2MST-_2$	102.6 (24.5)
VAS rest	1.6 (2.0)
VAS activity	2.0 (2.7)
VAS night	1.6 (2.4)
6MWT	536.3 (110.0)
SCT	8.9 (6.0)
KOOS	
Pain	79.2 (18.7)
Symptoms	81.8 (15.5)
Activities of daily living	83.9 (17.2)
Sport and recreation function	63.3 (25.5)
Knee-related quality of life	54.3 (24.5)
Lysholm score	80.7 (19.4)

2MST, 2-min step test; 6MWT, 6-min walk test; ACLR, anterior cruciate ligament reconstruction; KOOS, Knee injury and Osteoarthritis Outcome Score; SCT, stair climbing test; VAS, visual analog scale.

The associations between 2MST and clinical variables were examined using Pearson correlation coefficients [21]. Coefficients greater than 0.5 were considered strong correlations, 0.3–0.5 were considered moderate, and 0.2–0.3 were considered weak [21]. To identify the variables associated with the average of 2MST performance, stepwise multiple linear regression analysis was used. Significantly correlated variables were included in the regression model. Outliers were identified and treated using Cook's distance and centered leverage value. Multicollinearity was evaluated for each independent variable using an R value of less than 0.6. The level of significance was set at $P < 0.05$.

Results

Fifty participants [age: 29.9 (10.9) years; BMI: 23.3 (2.98) kg/m²] with quadriceps tendon autograft ($n = 15$), bone-patellar tendon-bone graft ($n = 10$) or hamstring autograft ($n = 25$) and completed a total of 12 weeks of physiotherapy were included. Descriptive statistics are shown in Table 2.

It was found that the 2MST took less than 5 min to describe and perform and that participants required minimal verbal support during the test. The participant satisfaction rate was found to be 92%.

The $ICC_{2,1}$ of the 2MST was 0.98 (0.96–0.99). SEM and MDC_{95} were 2.96 and 8.20 (2.41–14.00), respectively. The SEM% of 2.99 (≤5% is very good) was considered to be at a very good level (Table 3). The Bland–Altman plot illustrates that the 2MST scores had a 95% LoA ranging from –21.52 to 2.68 steps (Fig. 2).

There was a moderate-to-strong correlation between the average 2MST performance and the VAS rest score ($r = -0.46$, $P = 0.001$), VAS activity score ($r = -0.46$, $P = 0.001$), VAS night score ($r = -0.37$, $P = 0.007$), 6MWT score ($r = 0.58$, $P < 0.001$), SCT score ($r = -0.62$, $P < 0.001$), KOOS pain score ($r = 0.52$, $P < 0.001$), KOOS symptoms score ($r = 0.46$, $P = 0.001$), KOOS activities of daily living score ($r = 0.42$, $P = 0.002$), KOOS sport and recreation function score ($r = 0.55$, $P < 0.001$), KOOS knee-related quality of life score ($r = 0.33$, $P = 0.019$), and Lysholm score ($r = 0.35$, $P = 0.013$, Table 4). The actual correlation coefficients for the confirmation of the initial hypotheses (Table 1) in construct validity are given in Table 4 and all hypotheses were confirmed. The confirmation rate was 100.0%.

The analysis revealed no evidence of multicollinearity. The regression model included all correlated variables as independent variables. The SCT score ($\beta = -1.97$, $P < 0.001$) and VAS rest score ($\beta = -3.37$, $P = 0.013$) were found to be associated with 2MST performance, explaining 42.5% of the variance (Table 3). The number of steps during the 2MST increased as the duration of the SCT and VAS rest score decreased.

Discussion

The aim of the present study was to investigate the feasibility, test–retest reliability, and construct validity of the 2MST in measuring aerobic exercise capacity and functional endurance in participants with ACLR. The test was found to be feasible. The 2MST demonstrated excellent test–retest reliability. The reliability of the test was also confirmed with an SEM%. SEM and MDC_{95} for 2MST were 2.96 and 8.20, respectively. There was a moderate-to-strong correlation between the average 2MST performance and the VAS rest score, VAS activity score, VAS night score, 6MWT score, SCT score, KOOS pain score, KOOS symptoms score, KOOS activities of daily living score, KOOS sport and recreation function score, KOOS knee-related quality of life score, and Lysholm score in terms of construct validity. The SCT score and VAS rest score were the two most significant determinants of 2MST performance. The 2MST is a feasible, reliable, and valid test to measure aerobic exercise capacity and functional endurance in participants with ACLR.

There is currently no consensus on the most appropriate performance-based test to measure physical performance after ACLR [3]. Hop tests, isokinetic tests, functional movement screen, 6MWT, and SCT are commonly used [3,4]. The ability of a test to mimic activities of daily living may be a better indicator of whether a participant returning from the clinical setting to daily life has regained their preinjury activity level. The 2MST offers the benefit of simulating activities of daily living, such as walking and climbing stairs.

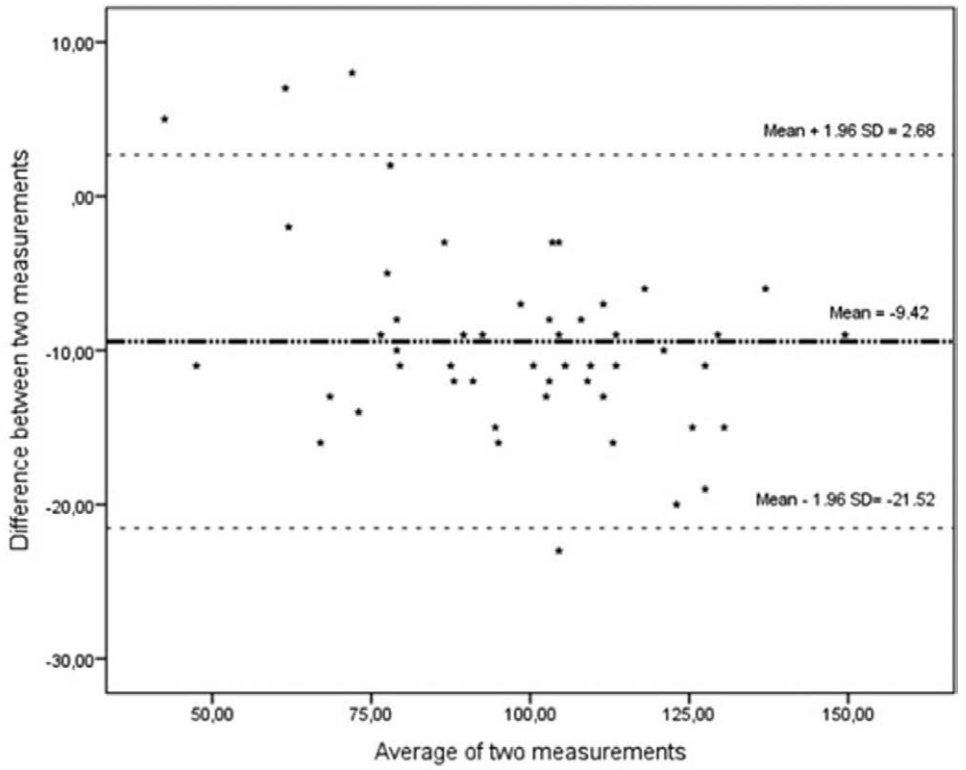
Previous studies have reported ICC values for the 2MST as follows: 0.94 in participants with knee osteoarthritis [6], 0.97 in participants with TKA [7], 0.94 in peripheral arterial disease [9], 0.83 in sedentary individuals [22], and

Table 3 **Test–retest reliability, standard error of measurement, and minimal detectable changes of the 2MST**

	First session Mean (SD)	Second session Mean (SD)	$ICC_{2,1}$ (95% of CI)	SEM	SEM (%)	MDC_{95}
2MST	93.2 (22.0)	102.6 (24.5)	0.98 (0.96–0.99)	2.96	2.99	8.20 (2.41–14.00)

CI, confidence interval; ICC, intraclass correlation coefficient; MDC_{95}, minimal detectable change at the 95% confidence level; SEM, SEM with a 95% confidence interval.

Fig. 2



Bland–Altman plot.

Table 4 **Correlation between 2MST and clinical variables**

	VAS rest	VAS activity	VAS night	6MWT	SCT	KOOS Pain	KOOS Symptoms	KOOS ADL	KOOS Sports	KOOS QoL	Lysholm score
2MST	$r=-0.46$ $P=0.001$*	$r=-0.46$ $P=0.001$*	$r=-0.37$ $P=0.007$*	$r=0.58$ $P<0.001$*	$r=-0.62$ $P<0.001$*	$r=0.52$ $P<0.001$*	$r=0.46$ $P=0.001$*	$r=0.42$ $P=0.002$*	$r=0.55$ $P<0.001$*	$r=0.33$ $P=0.019$*	$r=0.35$ $P=0.013$*

Regression equation formula is: 121.06 + [(–1.97 × SCT score, $P<0.001$) + (–3.37 × VAS rest score, $P=0.013$), F: 19.091 (dF: regression: 2; residual: 47), adjusted $R^2=0.425$.
2MST, 2-min step test; 6MWT, 6-min walk test; ADL, activities of daily living; dF, degrees of freedom; KOOS, Knee injury and Osteoarthritis Outcome Score; QoL, knee-related quality of life; SCT, stair climbing test; Sports, Sport and recreation function; VAS, visual analog scale.
*Statistically significant.

0.90 in participants with low back pain [5]. Our results are consistent with the existing literature, indicating that the 2MST can be administered reliably in ACLR participants between 3 and 12 months postoperatively, considering the excellent test–retest reliability. There is no consensus in the literature regarding the optimal timing for retesting. Braghieri *et al.* [9] reported that a 7-day interval is the most suitable duration of time. In this study, participants were retested 7 days after the first session. The retest scores of the participants were higher than their first session, possibly because the test was well learned.

de Jesus *et al.* examined the reliability of interrater evaluations for the 2MST in individuals with low back pain. They showed that the evaluations were reliable with ICC of 0.98 [5]. Akkan *et al.* [7] did not perform an intrarater evaluation similar to the methodology used in our study. The 2MST is a test that involves stepping while standing

and is scored by counting how much the right knee is raised. Therefore, we did not do an interrater reliability as the influence of the evaluator during this test is low.

SEM and MDC_{95} for 2MST were 2.96 and 8.20, respectively. The SEM% of 2.99 was considered to be at a very good level [7]. The SEM and MDC_{95} values require disease-specific analysis, and the lack of previous studies of 2MST performance in ACLR participants makes it challenging to discuss the results. SEM is a method for quantifying the error in a test. Although there is no definite evidence in the literature about the best cutoff point for SEM, Ostelo *et al.* [23] suggested that a 10% error rate is considered as acceptable. The results showed that the error rate for 2MST in ACLR participants was less than 10% [23]. The 2MST may be used to determine aerobic exercise capacity and functional endurance improvement in participants with ACLR between 3 and 12 months after surgery if there is a minimum change in performance of 8.2 steps.

Bland–Altman analysis showed a wide 95% LoA, with limits larger than the reported SEM and MDC_{95}, which were 2.96 and 8.2, respectively. This wide LoA appears to be inconsistent with the high ICC, low SEM, and relatively small MDC_{95}. The analysis of variance supports the ICC. Its strong dependence on the variance of the assessed population is the main limitation of this method. Therefore, higher ICC values could be achieved if applied to a more heterogeneous population (e.g. there were 50 participants with quadriceps tendon autograft, bone-patellar tendon-bone graft, or hamstring autograft in our study) than to a more homogeneous one, despite a similar degree of agreement [24–26]. LoA, on the other hand, is not affected by the variance of the population being assessed, but the interpretation of the significance of their results is highly subjective as it depends on the individual's understanding of the clinical significance of the range obtained [24,26]. A high ICC indicates that the between-subject variability is well explained, but does not necessarily reflect the variability of the measurement differences. The ICC may be high even when individual differences are large. A wide LoA is an indication of considerable variability in the differences between measurements. This can occur even with a high ICC and low SEM if there are significant individual differences [24,26].

The 2MST and 6MWT are both designed to assess the same underlying construct, namely aerobic exercise capacity. In line with our results, it has been reported that 2MST is well tolerated in healthy older adults, TKA subjects, older adults after coronary revascularization, and heart failure subjects on the basis of aerobic exercise capacity and could be considered as an alternative to 6MWT in terms of construct validity [7,27–29]. We found that higher levels of pain at rest, during activity, and at night were associated with worse 2MST performance. It seems reasonable in the light of the literature that higher pain levels in participants during 2MST, which is a prolonged, high-intensity, and weight-bearing activity, are associated with worse performance [7,10]. To assess construct validity, we used the SCT including an activity frequently used in real life. The fact that continuous stepping during the performance of the 2MST represents the neuromuscular demands of stair climbing and the strong correlation between 2MST-SCT supports the existence of construct validity [22]. In terms of KOOS and Lysholm scores, less functional problems and participant-reported symptoms in the knee were associated with better performance in 2MST. The moderate-to-severe correlations between the 2MST, KOOS, and Lysholm imply that participants' performance on the 2MST tended to improve as their self-reported knee function and symptoms, as measured by KOOS and Lysholm, improved. Our findings appear to be consistent with previous research emphasizing the utility of functional tests in reflecting participants' subjective assessments of their knee condition [1,2]. The SCT score and VAS rest score were the two most significant determinants of 2MST performance. When evaluating construct validity, it is important to use tests that are valid and include real-life tasks. Therefore, a clear correlation of 2MST with SCT is important for the validity of our results. The 2MST is recommended as a simple test to use in the clinics compared to the SCT. It appears that the high level of pain has a negative impact on 2MST performance, which involves constant stepping.

The study has limitations, including not assessing responsiveness and not including comorbidities in the analysis. As the study has a cross-sectional design, it is only possible to infer noncausal associations between the 2MST performance and clinical variables in participants with ACLR. In conclusion, the 2MST is a feasible, reliable, and valid test to measure aerobic exercise capacity and functional endurance in participants with ACLR.

Acknowledgements

The study was funded by the research team.

Each author has contributed substantially to the study protocol.

This study was approved by Selçuk University Medical Faculty Clinical Research Ethics Committee (2023/1221).

The data that support the findings of this study are available from the corresponding author, (C.K.), upon reasonable request.

Written and verbal consent were obtained from the participants.

The authors affirm that participants provided informed consent for publication.

Conflicts of interest

There are no conflicts of interest.

References

1 Glattke KE, Tummala SV, Chhabra A. Anterior cruciate ligament reconstruction recovery and rehabilitation: a systematic review. *J Bone Joint Surg Am* 2022; **104**:739–754.

2 Barber-Westin SD, Noyes FR. Factors used to determine return to unrestricted sports activities after anterior cruciate ligament reconstruction. *Arthroscopy* 2011; **27**:1697–1705.

3 Roe C, Jacobs C, Hoch J, Johnson DL, Noehren B. Test batteries after primary anterior cruciate ligament reconstruction: a systematic review. *Sports Health* 2022; **14**:205–215.

4 Thambyah A, Thiagarajan P, Goh Cho Hong J. Knee joint moments during stair climbing of patients with anterior cruciate ligament deficiency. *Clin Biomech (Bristol, Avon)* 2004; **19**:489–496.

5 de Jesus SFC, Bassi-Dibai D, Pontes-Silva A, da Silva de Araujo A, de Freitas Faria Silva S, Veneroso CE, *et al*. Construct validity and reliability of the 2-minute step test (2MST) in individuals with low back pain. *BMC Musculoskelet Disord* 2022; **23**:1062.

6 de Morais Almeida TF, Dibai-Filho AV, de Freitas Thomaz F, Lima EAA, Cabido CET. Construct validity and reliability of the 2-minute step test in patients with knee osteoarthritis. *BMC Musculoskelet Disord* 2022; **23**:159.

7 Akkan H, Kaya Mutlu E, Kuyubasi SN. Reliability and validity of the two-minute step test in patients with total knee arthroplasty. *Disabil Rehabil* 2024; **46**:3128–3132.

8 Ricci PA, Cabiddu R, Jürgensen SP, André LD, Oliveira CR, Di Thommazo-Luporini L, *et al*. Validation of the two-minute step test in obese with comorbidities and morbidly obese patients. *Braz J Med Biol Res* 2019; **52**:e8402.

9 Braghieri HA, Kanegusuku H, Corso SD, Cucato GG, Monteiro F, Wolosker N, *et al*. Validity and reliability of 2-min step test in patients with symptomatic peripheral artery disease. *J Vasc Nurs* 2021; **39**:33–38.

10 Bohannon RW, Crouch RH. Two-minute step test of exercise capacity: systematic review of procedures, performance, and clinimetric properties. *J Geriatr Phys Ther* 2019; **42**:105–112.

11 Kotsifaki R, Korakakis V, King E, Barbosa O, Maree D, Pantouveris M, *et al*. Aspetar clinical practice guideline on rehabilitation after anterior cruciate ligament reconstruction. *Br J Sports Med* 2023; **57**:500–514.

12 Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, De Vet HC., *et al*. COSMIN study design checklist for patient-reported outcome measurement instruments. *Qual Life Res*; 2019. pp. 1–32.

13 Vidmar MF, Baroni BM, Michelin AF, Mezzomo M, Lugokenski R, Pimentel GL, Silva MF. Isokinetic eccentric training is more effective than constant load eccentric training for quadriceps rehabilitation following anterior cruciate ligament reconstruction: a randomized controlled trial. *Braz J Phys Ther* 2020; **24**:424–432.

14 Ishigaki T, Kubo H, Yoshida K, Shimizu N, Ogawa T. Validity and reliability of the 2-min step test in individuals with stroke and lower-limb musculoskeletal disorders. *Front Rehabil Sci* 2024; **5**:1384369.

15 Tveter AT, Dagfinrud H, Moseng T, Holm I. Measuring health-related physical fitness in physiotherapy practice: reliability, validity, and feasibility of clinical field tests and a patient-reported measure. *J Orthop Sports Phys Ther* 2014; **44**:206–216.

16 Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, *et al*. 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* 2010; **19**:539–549.

17 Hoogeslag RAG, Brouwer RW, Boer BC, de Vries AJ, Huis In 't Veld R. Acute anterior cruciate ligament rupture: repair or reconstruction? Two-year results of a randomized controlled clinical trial. *Am J Sports Med* 2019; **47**:567–577.

18 Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). *Arthritis Care Res (Hoboken)* 2011; **63**(Suppl 11(0 11):S208–S228.

19 Chen YT, Murphy SL. Associations of coexisting pain and fatigue severity with physical performance and quality of life among middle-aged and older ındividuals with chronic knee pain: secondary analysis of a randomized clinical trial. *Pain Med* 2021; **22**:2575–2583.

20 PW S. Estimating the standard error of measurement from reliability studies. *Physiother Can* 2004; **56**:27–30.

21 Kararti C, Bilgin S, Dadali Y, Büyükturan B, Büyükturan O, Özsoy I, Bek N. Does plantar pressure distribution influence the lumbar multifidus muscle thickness in asymptomatic individuals? A preliminary study. *J Manipulative Physiol Ther* 2020; **43**:909–921.

22 Nogueira MA, Almeida TDN, Andrade GS, Ribeiro AS, Rêgo AS, Dias RDS, *et al*. Reliability and accuracy of 2-minute step test in active and sedentary lean adults. *J Manipulative Physiol Ther* 2021; **44**:120–127.

23 Ostelo RW, de Vet HC, Knol DL, van den Brandt PA. 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. *J Clin Epidemiol* 2004; **57**:268–276.

24 Costa-Santos C, Bernardes J, Ayres-de-Campos D, Costa A, Amorim-Costa C. The limits of agreement and the intraclass correlation coefficient may be inconsistent in the interpretation of agreement. *J Clin Epidemiol* 2011; **64**:264–269.

25 Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: measures of agreement. *Perspect Clin Res* 2017; **8**:187–191.

26 Haghayegh S, Kang HA, Khoshnevis S, Smolensky MH, Diller KR. A comprehensive guideline for Bland-Altman and intra class correlation calculations to properly compare two methods of measurement and interpret findings. *Physiol Meas* 2020; **41**:055012.

27 Węgrzynowska-Teodorczyk K, Mozdzanowska D, Josiak K, Siennicka A, Nowakowska K, Banasiak W, *et al*. Could the two-minute step test be an alternative to the six-minute walk test for patients with systolic heart failure? *Eur J Prev Cardiol* 2016; **23**:1307–1313.

28 Berlanga LA, Matos-Duarte M, Abdalla P, Alves E, Mota J, Bohn L. Validity of the two-minute step test for healthy older adults. *Geriatr Nurs* 2023; **51**:415–421.

29 Chow JJL, Fitzgerald C, Rand S. The 2 min step test: a reliable and valid measure of functional capacity in older adults post coronary revascularisation. *Physiother Res Int* 2023; **28**:e1984.