The Journal of Physical Therapy Science

SPTS

Original Article

Reliability and minimal detectable change of three functional tests: forward-lunge, step-up-over and sit-to-stand

Carmen Luque-Siles[1], Tomas Gallego-Izquierdo[2], Jose Jesus Jímenez-Rejano[3],
Susana Granados-de-la-Orden[4], Gustavo Plaza-Manzano[5],
Africa López-Illescas-Ruiz[6], Alejandro Ferragut-Garcías[7],
Natalia Romero-Franco[2], Patricia Martín-Casas[8], Daniel Pecos-Martín[2, 9]*

[1] High Performance Sports Medicine Center of the High Council for Sports, Spain
[2] Physiotherapy Department, University of Alcalá: Madrid-Barcelona Rd., Km. 33.600. Alcalá de Henares, Madrid, Spain
[3] Physiotherapy Department, University of Seville, Spain
[4] Sub-department of Prevention and Health Promotion, Council for Health, Spain
[5] Departamento de Medicina Física y Rehabilitación, Facultad de Medicina, Universidad Complutense de Madrid. Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Spain
[6] Health Sciences Department, University Alfonso, Spain
[7] Physiotherapy and Nursery Department, University of the Islas Baleares, Spain
[8] Nursery and Physiotherapy Department, University of the Balearic Islands, Spain
[9] Departamento de Medicina Física y Rehabilitación, Facultad de Enfermería, Fisioterapia y Podología, Universidad Complutense de Madrid, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Spain

Abstract. [Purpose] To examine the intrasession and intersession reliability and the absolute reliability of three functional dynamic tests—forward-lunge, step-up-over and sit-to-stand tests—using computerized dynamic posturography. [Subjects and Methods] An intra-test and test-retest, repeated measure study was designed. Forty-five healthy subjects twice carried out the forward-lunge test, step-up-over test, and sit-to-stand test on two days, one week apart. The intrasession and intersession reliabilities as judged by the intraclass correlation coefficient (ICC) and the minimal detectable change of the three functional tests were calculated. [Results] Excellent to very good intrasession reliability of the forward-lunge test (ICC range of 0.9–0.8) was found. Very good to good intrasession reliability of the step-up-over test (ICC range of 0.9–0.5) was found and very good intrasession reliability of the sit-to-stand test (ICC range of 0.8–0.7) was found. The minimal detectable change at the 95% confidence level of most of the measures was lower than 30%. [Conclusion] The forward-lunge, step-up-over and sit-to-stand tests are reliable measurement tools.
Key words: Postural balance, Reproducibility, Activities of daily living

INTRODUCTION

Static and dynamic balance are milestones in the proper development of daily living. Thanks to dynamic stability, a subject displaces and controls the center of gravity (COG) within the base of support in dynamic functional conditions, performing

*Corresponding author. Daniel Pecos-Martín (E-mail: daniel.pecos@uah.es)

complex combined movements such as striding forward (forward-lunge, FL), going up and down a step (step-up-over, SUO), and standing up from a seated position (sit-to-stand, STS)[1].

Many researchers have studied the ability of the computerized dynamic posturography to characterize several features of balance in different populations[2, 3]. This test provides accurate values regarding parameters such as impact, rising, distance, time, impulse and COG sway, during the aforementioned movements in dynamic functional conditions[3]. Using computerized dynamic posturography, these tests show the balance impairments of patients who could benefit from rehabilitation treatment during activities of living involving dynamic conditions[4]. However, only one study has reported the reliability of the FL and SUO dynamic functional tests and none have reported the reliability of the STS test[5], which makes the clinical applicability of these dynamic balance tests controversial. In this study, the intrasession and intersession reliability of measurements obtained in the FL, SUO and STS tests were assessed. This study also calculated the standard error of measurement (SEM), which shows the ability of the tests to exclude the noise from measurements[6, 7]. The absolute and relative minimal detectable changes (MDC) were also calculated to estimate the real change between two measurements of the same subject across time, as this parameter is clinically more relevant[8].

SUBJECTS AND METHODS

An intra-test and test-retest, repeated measures study was designed. The sample size was determined considering a two-tailed hypothesis, a 0.05 type I error, a 0.93 intraclass correlation coefficient (ICC) estimated for the impulse variable in the FL test and a 0.075 accuracy level[9]. The required sample size was 30 subjects.

Forty-five healthy subjects were randomly selected from the general university population (24 males and 21 females; 23.89 ± 5.49 years; 1.69 ± 0.08 cm; 68.43 ± 10.97 kg; 32 right-handed, 12 left-handed, 2 ambidextrous). Subjects with balance affectation[10], musculoskeletal diseases or related surgery, or previous balance training were excluded. Before beginning the tests, each participant signed an informed consent form. The Ethics Committee of Seville University approved this study (Code: M2012/024/20130226), which was carried out at the Physiotherapy Unit of a Sports Medicine Center.

Measurements were carried out using the platform Balance Master® System 8.2 (USA) under controlled laboratory conditions by an experienced examiner, and following a standardized protocol so all the participants were given directions in the same way. Data were collected in two assessment sessions, 6–8 days apart. During the first session, the three tests were carried out consecutively with 20–30-minute rest periods. The three tests were repeated in the same order as the first session. Initially, the participants practiced each dynamic test to acquaint themselves with the movement exertion when receiving a visual prompt from a screen[10]. Each test was performed three times. In the second session, the three tests were performed only once, with three trials performed for each.

Forward-lunge (FL) test: The participants stayed as upright and as still as possible on the platform. On seeing the prompt GO on the screen, they moved their foot forward as far and fast as they could, bending the knee. Then, they quickly returned to the starting position and stood still. The trunk was kept in a vertical position while the COG transferred from one foot to the other quickly. The subjects were instructed not to use their arms for help during the movement. The left foot went first. The outcome variables were the distance reached (distance) defined as the stride length, measured from the COG location during the standing position to the furthest point reached by the COG while stepping forward, expressed as a percentage of the subject's height (%). The impact index was defined as the maximal reaction force generated by the swinging forward leg when it contacted the ground, expressed as a percentage of the body weight (%). The contact time (contact time) was defined as the lapse of time (s) between the moment when the swinging forward leg left the ground to when it made contact with the ground again; and the impulse force was defined as the total amount of work done by the swinging forward leg during the stance phase, and the push back force of the movement, expressed as a percentage of the body weight multiplied by the time that the force was exerted (% × s). The mean of three trials per test was used following the manufacturer's protocol and previous recommendations regarding the highest relative and absolute reliabilities of biomechanical variables in the forward-lunge test[11].

Step-up-over (SUO) test: The participants stayed as upright and as still as possible on the platform. On seeing the prompt GO on the screen, they put one foot on a 21-cm high step and then lifted the other foot, passed it over the step without touching it and placed it back on the ground. Then, they placed the foot on the step and stood still. The left foot went first. The outcome variables were: the lift-up index, defined as the maximal strength exerted by the lower limb on the step during the lift-up phase, expressed as a percentage of the body weight (%); the movement time, defined as the lapse of time when it contacted the ground again; and the impact index, defined as the maximal reaction force generated by the swinging forward leg when it contacted the ground again, expressed as a percentage of the body weight (%). Although the platform manufacturer states that this test should be carried out "quickly", in our prior experience we observed that this instruction resulted in performance of the movement while "running". This caused the other variables—the lift-up index and impact index—to exhibit characteristics not related to the subjects' behaviors in activities of daily living. As a consequence, the participants were instructed to mount the step in a "spontaneous" way, as they would in normal daily life.

Sit-to-stand (STS) test: The participants sat quickly on a 41-cm high box placed on the platform. On seeing the prompt GO on the screen, they stood up as quickly as possible without using their arms and stood still. The outcome variables were the rising index, defined as the maximal strength exerted by the lower limbs during the rising phase, expressed as a percentage of

Table 1. Mean values, standard deviations and ranges of unilateral and bilateral variables in the forward-lunge, step-up-over and sit-to-stand tests

		Session 1: trial 1		Session 1: trial 1		Session 2	
Forward-lunge		Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Distance (%)	Right	56.5 (6.3)	(40–69)	56.6 (6.2)	(39–69)	56.8 (6.2)	(43–69)
	Left	56.3 (5.8)	(41–67)	55.8 (6.0)	(42–69)	56.3 (6.0)	(44–70)
Impact index (%)	Right	27.4 (7.6)	(12–50)	26.3 (7.1)	(15–46)	27.3 (6.8)	(15–49)
	Left	26.7 (6.7)	(12–41)	25.8 (6.3)	(15–41)	27.9 (6.8)	(17–45)
Contact time (s)	Right	1.0 (0.2)	(0.5–1.3)	1.0 (0.2)	(0.6–1.4)	1.0 (0.2)	(0.6–1.4)
	Left	1.0 (0.2)	(0.6–1.4)	1.0 (0.2)	(0.6–1.3)	1.0 (0.2)	(0.6–1.4)
Impulse force (%xs)	Right	106.8 (16.3)	(69–139)	106.6 (17.2)	(64–144)	107.4 (16.1)	(73–148)
	Left	104.3 (16.4)	(71–142)	103.8 (15.3)	(72–134)	104.6 (15.4)	(75–145)
Step-up-over							
Lift-up index (%)	Right	43.0 (9.3)	(18–60)	48.9 (11.2)	(23–77)	48.3 (11.8)	(22–80)
	Left	40.2 (9.7)	(18–61)	43.8 (10.0)	(19–68)	45.4 (12.0)	(19–79)
Movement time (s)	Right	1.3 (0.2)	(1.0–1.7)	1.2 (0.2)	(0.9–1.5)	1.2 (0.2)	(1.0–1.6)
	Left	1.3 (0.2)	(1.0–1.7)	1.2 (0.1)	(0.9–1.5)	1.2 (0.2)	(0.9–1.6)
Impact index (%)	Right	42.1 (13.1)	(11–74)	44.4 (13.7)	(15–79)	43.5 (13.6)	(11–76)
	Left	44.1 (12.7)	(20–70)	48.0 (13.6)	(15–80)	47.6 (14.4)	(17–85)
Sit-to-stand							
Rising index (%)		29.6 (8.5)	(12–50)	31.9 (8.4)	(17–49)	32.3 (8.9)	(18–54)
COG swayvelocity (°/s)		3.1 (1.1)	(1.0–5.4)	3.4 (1.0)	(1.2–5.3)	3.4 (1.0)	(1.2–5.4)

COG: center of gravity; SD: standard deviation; Distance (% subject's height: cm), Impact Index (% body weight: kg), Contact Time (s), Force Impulse (% body weight: kg × Time: s) Lift-Up Index (% body weight: kg), Movement Time (s), Rising Index (% body weight: kg), COG Sway Velocity (°/s)

the body weight (%), and the COG sway velocity (COG sway velocity), defined as the velocity of the COG during the rising phase and the next five seconds (°/s).

SPSS 18 for Windows was used for statistical analysis. Three tests and nine variables were analyzed (seven of them bilateral). Each test was carried out three times and the mean was used the statistical analysis.

The normality of the variables of the three measurements was tested. The mean as well as the differences between each pair of measurements were calculated (the first was compared to the second one for intrasession reliability, and the first to the third, for intersession reliability) giving the sixteen variables, and the skewness coefficient and the kurtosis were also calculated also.

The relative reliability was analyzed considering the ICC, which was calculated following two procedures: a two factor mixed-effect model (ICC A) with absolute agreement; and a one factor random-effect model (ICC B). The absence of a systematic error in the measurements was verified due to the equality of ICC A and the ICC B[12].

The SEM was calculated following the Stratford and Goldsmith procedure[7]. The absolute reliability or "repeatability" was calculated following the Bland and Altman procedure[12, 13]. The repeatability, also known as minimal detectable change (MDC$_{95}$) because it refers to "the minimal change that falls outside the measurement error of an instrument used to measure a symptom"[14], was calculated using the equations: SEM × 1.96 × √2, 95% confidence level. According to one-way analysis of variance (one-way ANOVA), the square root of the quadratic mean of the residual from the intrasubject effect test is the intrasubject standard deviation, which allows calculation of the systematic error[12, 13]. The SEM and the MDC$_{95}$ were also calculated as percentages of the measurement mean[6, 8]. In this way, the absolute reliability is obtained independent of the units of measurement. They are represented as SEM% and MDC$_{95}$%, respectively. Therefore, MDC$_{95}$ indicates a real change with 95% confidence for individual participants and demonstrates the repeatability of the test.

The statistical tests were performed using a 95% confidence interval (CI) (p<0.05).

RESULTS

Fifty subjects were enrolled, but five of them dropped out from the second session due to issues not related to the research. Thus, the final sample was 45 subjects. For that reason, the estimated ICC value for a 0.075 accuracy level was 0.902.

Table 1 gives a descriptive analysis of the collected data. The normality of all the variables registering a skewness coefficient and kurtosis ≤1.5 was calculated. In all cases, ICC A was not different from ICC B, implying that there were no systematic errors. Accordingly, only ICC A is shown in Tables 2, 3 and 4. The Weir classification for ICC was used[15]: scores

Table 2. Forward-lunge test

		Intrasession			Intersession		
		ICC (95% CI)	SEM (%)	MDC$_{95}$ (%)	ICC (95% CI)	SEM (%)	MDC$_{95}$ (%)
Distance (%)	Right	0.9 (0.9–0.9)	1.8 (3.1%)	4.9 (8.6%)	0.9 (0.8–0.9)	2.2 (3.8%)	6.0 (10.6%)
	Left	0.9 (0.9–1.0)	1.7 (3.0%)	4.7 (8.3%)	0.9 (0.7–0.9)	2.3 (4.2%)	6.5 (11.5%)
Impact index (%)	Right	0.8 (0.6–0.9)	3.4 (12.6%)	9.4 (34.8%)	0.6 (0.4–0.7)	4.3 (15.9%)	11.9 (44.1%)
	Left	0.8 (0.6–0.9)	3.1 (11.4%)	8.5 (31.7%)	0.7 (0.5–0.8)	3.8 (14.4%)	10.7 (39.8%)
Contact time (s)	Right	0.9 (0.8–0.9)	0.1 (6.1%)	0.2 (16.2%)	0.7 (0.5–0.8)	0.1 (10.1%)	0.3 (27.3%)
	Left	0.9 (0.8–0.9)	0.1 (5.2%)	0.1 (13.5%)	0.8 (0.6–0.9)	0.1 (8.3%)	0.2 (22.9%)
Impulse force	Right	0.9 (0.8–0.9)	5.8 (5.4%)	16.1 (15.1%)	0.7 (0.5–0.8)	9.4 (8.8%)	26.0 (24.3%)
(%xs)	Left	0.9 (0.8–0.9)	5.1 (4.9%)	14.2 (13.6%)	0.8 (0.7–0.9)	7.2 (6.9%)	20.0 (19.2%)

CI: confidence interval; ICC: intraclass correlation coefficient; MDC: minimal detectable change; SEM: standard error mean; Distance (% subject's height: cm), Impact Index (% body weight: kg), Contact Time (s), Impulse Force (% body weight: kg × Time: s) $p < 0.001$

Table 3. Step-up-over test

		Intrasession			Intersession		
		ICC (95% CI)	SEM (%)	MDC$_{95}$ (%)	ICC (95% CI)	SEM (%)	MDC$_{95}$ (%)
	Right	0.6 (0.2–0.8)	5.1 (11.0%)	14.2 (30.5%)	0.7 (0.3–0.9)	5.0 (10.7%)	13.8 (29.5%)
Lift-up index (%)	Left	0.7 (0.4–0.8)	5.3 (12.4%)	14.7 (34.2%)	0.6 (0.3–0.8)	6.0 (14.0%)	16.7 (38.8%)
	Right	0.7 (0.3–0.9)	0.1 (5.7%)	0.2 (15.3%)	0.6 (0.4–0.8)	0.1 (8.1%)	0.3 (21.8%)
Movement time (s)	Left	0.5 (0.1–0.8)	0.1 (7.1%)	0.3 (19.8%)	0.5 (0.1–0.8)	0.1 (7.9%)	0.3 (21.4%)
	Right	0.8 (0.7–0.9)	5.0 (11.5%)	13.8 (31.8%)	0.9 (0.8–0.9)	5.0 (11.4%)	13.7 (31.7%)
Impact index (%)	Left	0.7 (0.5–0.8)	6.7 (14.3%)	18.5 (39.7%)	0.7 (0.5–0.8)	7.0 (15.0%)	19.3 (46.6%)

CI: confidence interval; ICC: intraclass correlation coefficient; MDC: minimal detectable change; SEM: standard error of mean; Lift-up index (% body weight: kg), Movement time (s), Impact index (% body weight: kg) $p < 0.001$

Table 4. Sit-to-stand test

	Intrasession			Intersession		
	ICC (95% CI)	SEM (%)	MDC$_{95}$ (%)	ICC (95% CI)	SEM (%)	MDC$_{95}$ (%)
Rising Index (%)	0.8 (0.6–0.9)	3.4 (10.8%)	9.4 (30.0%)	0.7 (0.5–0.8)	4.5 (14.5%)	12.5 (40.0%)
COG swayvelocity (°/s)	0.7 (0.5–0.8)	0.6 (17.7%)	1.6 (48.8%)	0.6 (0.3–0.7)	0.7 (21.7%)	2.0 (59.8%)

CI: confidence interval; COG: center of gravity; ICC: intraclass correlation coefficient; MDC: minimal detectable change; SEM: standard error of mean; Rising Index (% body weight: kg), COG Sway Velocity (°/s) $p < 0.001$

from 0.50 to 0.69, good; from 0.70 to 0.89, very good; and from 0.90, excellent; SEM% lower than 35%, good;[6] MDC$_{95}$% lower than 30%, reasonable; and MDC$_{95}$% lower than 10%, excellent[16].

Table 2 shows ICC and 95% confidence intervals of the FL test. They were excellent to very good for the intrasession reliabilities (range 0.9–0.8), and very good to good for the intersession reliabilities (range 0.9–0.7). The SEM% scores for all the parameters were much lower than 35% and the MDC$_{95}$% of most of the parameters were lower than 30% (except for the impact index), with excellent scores (around 10%) for the distance variable, which was the most reliable, followed by the impulse and the contact time.

Table 3 shows the ICCs (95% CI) of the SUO test. They were very good to good for the intra- and inter-session reliabilities (range 0.9–0.5). The SEM% scores of all the parameters were much lower than 35% and the MDC$_{95}$% scores of most of all the parameters were around 30%, though the lift-up and impact index of the SUO showed poorer ability to detect real change than the FL variables.

Table 4 shows the ICCs (95% CI) of the STS test, which were very good for the intrasession reliability (range 0.8–0.7) and very good to good for the intersession reliability (range 0.7–0.6). The SEM% scores of both parameters were much lower than 35% and the most reliable variable was the rising index, which had an intersession MDC$_{95}$% of 30%, whereas the COG sway exceeded this limit in intra- and inter-session reliabilities.

DISCUSSION

The main finding of this study was the good absolute reliability of most FL variables, which means that this test is a potential diagnosis tool for musculoskeletal diseases. However, comparisons with other research are difficult due to the ICC context-dependence[17, 18] the ICC being affected by factors such as the number of subjects, the number of repetitions, the lapse of time between assessment sessions, the protocol followed, the number of examiners and systematic errors[7]. However, the 0.9 ICC hypothesis for the impulse variable of the FL test was not been proven, as the intersession ICC for the impulse values of the FL in the present study were 0.7 and 0.8 for the right and left legs, respectively.

The relative variability of the FL measurements made with a 1-week lapse of time (intersession) is bigger than that observed between the measurements made during the same session (intrasession). This shows that time is a factor influencing on the FL parameters' intersubject variability. Furthermore, FL contact time may differ with dynamic movement exertion due to anterior cruciate ligament (ACL) pathology[11]. FL contact time data collected in this present study were longer than those of a study following the same protocol and methods[5], even though subjects were specifically asked in both studies to perform FL as quickly as possible. The age range and nature of the sedentary subjects enrolled in the present study (18–40 years compared to 18–26) might explain the longer contact time duration. Thus, these slight differences may show the impact of age or physical condition on this variable.

The FL contact time was longer in another previous study, possibly because a greater number of test repetitions and the absence of a specific instruction to perform FL as quickly as possible[11]. Also, a previous study reported a 270 ms difference in the FL contact time between control and experimental subjects who had not fully recovered from ACL surgery[19].

In our study, the absolute reliability of the FL contact time was calculated using the SEM values of 50 to 100 ms. The MDC_{95} was lower (130–270 ms) indicating the ability of the FL test ability to discriminate between subjects with or without ACL pathologies. The FL test was also able to highlight the variation in the FL contact time between two measurements made on the same subject suffering from this kind of pathology.

Furthermore, our results for the FL and SUO impact when down a step (similar to a short stride) are higher than the impact values of a long stride. Since a longer stride elicits less stress than a shorter one, previous studies have reported that the distances achieved (the stride length) in the FL test affects the impact force and the stress borne by the patellofemoral joint[20]. Therefore, the shorter the step, the bigger the force that the lower limb absorbs. Also, the FL test mainly involves the hip extensors; therefore, it would be worthwhile to investigate whether this test is sensitive and specific to pathologies involving these muscles[21].

Similarly, our results for the movement times of the SUO test were longer than those of similar studies[5]. In the present study, the subjects were not asked to perform this test quickly as the manufacturer recommends; rather, they were asked to perform it at that in their daily life. The difference in instructions may explain the longer times observed in the present study as well as the differences seen in the impact and rising indices, which were lower in the present study.

Parameters based on the vertical reaction forces during descending stairs tests are able to monitor the rehabilitation progress after ACL surgery[22]. Patients suffering from patellofemoral pain minimize the reaction forces both in ascent and descent phase of stairs to maintain normal levels of stress in the joint[23]. Accordingly, the SUO lift-up and impact index (derived from the reaction forces) may show these adaptations. It is our opinion that this test should be accompanied by other functional tests to help to determine their limitations. Although the relative variabilities of these parameters in our present research were high, their absolute reliabilities persuaded us that this test is a useful tool.

The STS COG velocity exhibited relative and absolute reliability values lower than the STS rising index. In agreement with our present findings, previous research has reported poor relative reliabilities for this test[24]. Goldie et al.[25] determined that measurements based on ground reaction forces in a unipedal stance were more reliable than those based on centers of pressure (COP)[25]. The STS and SUO rising and lift-up indices theoretically represent similar aspects of the ground reaction forces in two movements of daily living (going upstairs and standing up). Thus, they may complement each other and enhance the possibility of discriminating between subjects with and without a musculoskeletal disease.

A limitation of this study was that the intrasession and intratest reliabilities were calculated from measurements made within 1 hour, which features the internal consistency of the measurements (i.e., the ability of the tool to reproduce measurements and the accuracy of the measurements in stable subjects). Moreover, although the test-retest and intraexaminer reliabilities were calculated, the interexaminer reliability was not. However, this factor affects the variability of measurements made in tests that depend on the explanation that an examiner gives to subjects; i.e. results that are based on subjects' understanding[17]. Consequently, these aspects should be considered when discussing the reliability of the results.

In conclusion, the contact time, distance and force impulse variables of the FL test showed excellent absolute and relative reliabilities. The variables based on COP changes in the STS tests did not show enough reliability in a healthy population.

The collected data establishes a range of normal values for the variables of the tests, as well as relative and absolute reliabilities and $MDC_{95}\%$ (minimal detectable change) of dynamic postural stability measurements of a healthy population.

REFERENCES

1) Nashner LM, Peters JF: Dynamic posturography in the diagnosis and management of dizziness and balance disorders. Neurol Clin, 1990, 8: 331–349. [Medline]

2) Ford-Smith CD, Wyman JF, Elswick RK Jr, et al.: Test-retest reliability of the sensory organization test in noninstitutionalized older adults. Arch Phys Med Rehabil, 1995, 76: 77–81. [Medline] [CrossRef]

3) Liston RA, Brouwer BJ: Reliability and validity of measures obtained from stroke patients using the Balance Master. Arch Phys Med Rehabil, 1996, 77: 425–430. [Medline] [CrossRef]

4) Topp R, Mikesky A, Thompson K: Determinants of four functional tasks among older adults: an exploratory regression analysis. J Orthop Sports Phys Ther, 1998, 27: 144–153. [Medline] [CrossRef]

5) Naylor ME, Romani WA: Test-retest reliability of three dynamic tests obtained from active females using the Neurocom Balance Master. J Sport Rehabil, 2006, 15: 326–337. [CrossRef]

6) Wagner JM, Rhodes JA, Patten C: Reproducibility and minimal detectable change of three-dimensional kinematic analysis of reaching tasks in people with hemiparesis after stroke. Phys Ther, 2008, 88: 652–663. [Medline] [CrossRef]

7) Stratford PW, Goldsmith CH: Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. Phys Ther, 1997, 77: 745–750. [Medline]

8) Hiengkaew V, Jitaree K, Chaiyawat P: Minimal detectable changes of the Berg Balance Scale, Fugl-Meyer Assessment Scale, Timed "Up & Go" Test, gait speeds, and 2-minute walk test in individuals with chronic stroke with different degrees of ankle plantarflexor tone. Arch Phys Med Rehabil, 2012, 93: 1201–1208. [Medline] [CrossRef]

9) Morales AR, Zárate LEM: Epidemiología clínica: investigación clínica aplicada: Médica Panamericana, 2004.

10) Pickerill ML, Harter RA: Validity and reliability of limits-of-stability testing: a comparison of 2 postural stability evaluation devices. J Athl Train, 2011, 46: 600–606. [Medline]

11) Alkjaer T, Henriksen M, Dyhre-Poulsen P, et al.: Forward lunge as a functional performance test in ACL deficient subjects: test-retest reliability. Knee, 2009, 16: 176–182. [Medline] [CrossRef]

12) Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. Lancet, 1986, 1: 307–310. [Medline] [CrossRef]

13) Conradsson M, Lundin-Olsson L, Lindelöf N, et al.: Berg balance scale: intrarater test-retest reliability among older people dependent in activities of daily living and living in residential care facilities. Phys Ther, 2007, 87: 1155–1163. [Medline] [CrossRef]

14) Kovacs FM, Abraira V, Royuela A, et al. Spanish Back Pain Research Network: Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. BMC Musculoskelet Disord, 2008, 9: 43. [Medline] [CrossRef]

15) Weir JP: Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res, 2005, 19: 231–240. [Medline]

16) Huang SL, Hsieh CL, Wu RM, et al.: Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with Parkinson disease. Phys Ther, 2011, 91: 114–121. [Medline] [CrossRef]

17) Innes E, Straker L: Workplace assessments and functional capacity evaluations: current beliefs of therapists in Australia. Work, 2003, 20: 225–236. [Medline]

18) Bruton A, Conway JH, Holgate ST: Reliability: what is it, and how is it measured? Physiotherapy, 2000, 86: 94–99. [CrossRef]

19) Alkjaer T, Simonsen EB, Peter Magnusson SP, et al.: Differences in the movement pattern of a forward lunge in two types of anterior cruciate ligament deficient patients: copers and non-copers. Clin Biomech (Bristol, Avon), 2002, 17: 586–593. [Medline] [CrossRef]

20) Escamilla RF, Zheng N, Macleod TD, et al.: Patellofemoral joint force and stress between a short- and long-step forward lunge. J Orthop Sports Phys Ther, 2008, 38: 681–690. [Medline] [CrossRef]

21) Riemann B, Congleton A, Ward R, et al.: Biomechanical comparison of forward and lateral lunges at varying step lengths. J Sports Med Phys Fitness, 2013, 53: 130–138. [Medline]

22) Colby SM, Hintermeister RA, Torry MR, et al.: Lower limb stability with ACL impairment. J Orthop Sports Phys Ther, 1999, 29: 444–451, discussion 452–454. [Medline] [CrossRef]

23) Brechter JH, Powers CM: Patellofemoral joint stress during stair ascent and descent in persons with and without patellofemoral pain. Gait Posture, 2002, 16: 115–123. [Medline] [CrossRef]

24) Panzer VP, Wakefield DB, Hall CB, et al.: Mobility assessment: sensitivity and specificity of measurement sets in older adults. Arch Phys Med Rehabil, 2011, 92: 905–912. [Medline] [CrossRef]

25) Goldie PA, Bach TM, Evans OM: Force platform measures for evaluating postural control: reliability and validity. Arch Phys Med Rehabil, 1989, 70: 510–517. [Medline]