NAJSPT

ORIGINAL RESEARCH

RELIABILITY AND VALIDITY OF A PARTIAL WEIGHT BEARING MEASURE OF LOWER EXTREMITY PERFORMANCE

Daniel J Cipriani, PT, PhD[a]
Michelle Haines, MA[a]
Denise O'Rand, ATC, PhD[a]
Susan Levy, PhD[a]

ABSTRACT

Background. Methods of measuring lower extremity function is limited for those with partial weight bearing (PWB) status in early phases of a lower extremity rehabilitation program.

Objectives. The purpose of this study was to measure intra-rater reliability of two lower extremity PWB performance measures using an incline exercise apparatus and to evaluate the concurrent validity and responsiveness to change of these two measures.

Methods. Thirty-seven adult patients with lower extremity injuries were measured on two PWB measures (PWB20 and PWB30) of lower extremity performance as well as several common measures of LE function. After initial testing, subjects were asked to return for retesting, following four to six weeks of rehabilitation intervention. Reliability of the data from the measures was tested using intraclass correlation coefficients (ICC); validity was based on bivariate correlations of the measures. The minimal detectable change (MDC) value and limb symmetry index (LSI) were used to study the responsiveness of the PWB measures.

Results. The ICC for the PWB20 and PWB30 were 0.95 and 0.98, respectively. The bivariate correlations of the PWB20 with stair climbing and walking speed were greater than those of the PWB30. Correlations ranged from r = 0.49 to 0.72 between the PWB measures and the functional measures. For most patients, their change in score between initial testing and follow-up exceeded the MDC; the LSI improved for all patients.

Conclusion. Using the incline apparatus yielded reliable PWB data. In addition, performance on the PWB measures correlated fairly well with common measures of function.

Key words: partial weight bearing, incline apparatus, reliability, validity

CORRESPONDENCE
Daniel Cipriani
San Diego State University
5500 Campanile Drive
San Diego, CA 92182-7251
619-594-1922
cipriani@mail.sdsu.edu

[a] San Diego State University
San Diego, CA

INTRODUCTION

Functional performance tests (dynamic full weight bearing tests) are useful predictors of lower extremity performance, which in turn allows for development of a realistic prognoses.[1,2] Functional performance tests include the single legged hop for distance tests, stair climbing tests, and walk tests, among others. Each of these tests is supported by research in terms of the reliability of the measures and the validity of the inferences made from these tests.[3-13] The common denominator is that the patient must have full weight bearing status in order to perform these tests. Currently few options exist to measure lower extremity performance of individuals with less than full weight bearing ability. For instance, individuals recovering from surgery (lower extremity total joint replacement, anterior cruciate ligament reconstruction, fixation after fracture, etc.) frequently initiate rehabilitation under weight bearing restrictions. In other cases, because of pain or weakness, performing traditional functional weight bearing tests is not feasible early in the recovery. A useful measure of lower extremity performance ability is essential for rehabilitation, treatment progression, and development of accurate prognoses of individuals with limited weight bearing status (i.e., partial weight bearing, PWB). Unfortunately, clinicians are limited to either subjective evaluation, self-report, or non weight bearing measures to estimate performance in individuals with restricted weight bearing ability.

Common forms of measurement of non weight bearing performance include manual muscle testing, joint range of motion (ROM), joint integrity measures (e.g., ligament laxity testing), and isokinetic testing. While these traditional clinical examination techniques provide reliable data, their reported predictive validities are low, as tests generally demonstrate poor correlation to lower extremity functional performance.[1,2,8,10] For example, Kea et al[8] examined the relationship between isokinetic testing of hip abduction and adduction movements to a lateral hop test for distance in elite hockey players. The relationship between isokinetic measures of hip strength and the hop tests was slight to poor ($r = -.26$ to .27). Kea et al[8] concluded that function should not be predicted by joint-specific strength tests. Additional studies reported a wide range of correlations between isokinetic test measures and functional performance measures ranging from r-values of 0.26 to 0.63, with most of these studies testing the correlation of isokinetic measures to hop tests measures. In general, the authors concluded that care must be exercised when interpreting isokinetic measures of muscle performance in terms of functional performance.[14-17]

Indirectly measuring the responsiveness of various lower extremity performance measures, Worrell et al[18] used isokinetic testing, along with maximum lateral step-up repetitions, a leg press test, and two hop tests to measure changes in subjects following a six week lower extremity strengthening protocol. The protocol involved lateral step-up exercises in full weight bearing. At the conclusion of the study, all lower extremity performance measures improved with the exception of the isokinetic test measures. Worrell et al[18] concluded that the non weight bearing isokinetic measure was not responsive to the changes gained in a weight bearing exercise program.

While isokinetic tests provide reliable measures of muscle strength, these tests do not show evidence of predictive validity for weight bearing functional performance ability.[18,19] The manual muscle test, while being an accepted measure of leg strength, only measures static muscle strength and does not predict dynamic activity of the lower extremity.[20] Hence, clinicians have limited options to measure lower extremity performance in individuals with PWB status.

An option worth consideration for measuring lower extremity performance in PWB is a sliding incline device, called the Total Gym, that was originally designed for partial weight bearing exercise. Using the Total Gym, Munich et al[21] examined two lower extremity performance measures in a PWB position. The two lower extremity performance tests evaluated by Munich et al[21] included the following: 1) the number of one-legged squats performed in 20 seconds on the Total Gym; and 2) the time required to perform 50 one-legged squats on the Total Gym. The intention of these two measures was to evaluate lower extremity performance, using the one-legged squat test as the definition of performance, in partial weight bearing. According to Munich et al,[21] the test of one-legged squats in 20 seconds was designed to indirectly measure power of the lower extremity, and the 50 one-legged squats test was designed to indirectly measure local muscle endurance. All subjects were healthy young adults.

Munich et al[21] concluded that the sliding incline apparatus was able to yield reliable data, with ICC values for intratester reliability exceeding 0.80. However, the test-retest reliability and inter-tester reliability of these measures, on

a non-healthy population, has not been evaluated. In addition, the validity or application of this protocol to an injured population has not been studied.

In the rehabilitation setting, clinicians are limited in their ability to measure lower extremity PWB performance of individuals with limited weight bearing ability. The PWB tests studied by Munich et al[21] may provide an option for the clinician in order to provide early assessment of lower extremity performance in individuals with limited weight bearing. However, before test efficacy can be assumed, further data regarding the reliability and validity of these tests in an injured sample needs to be determined. Therefore, the purpose of this study was to measure the intratester reliability of the partial weight bearing tests described by Munich et al[21] with individuals recovering from lower extremity injuries or surgery. In addition, this study evaluated the validity of these PWB performance tests, in terms of concurrent validity evidence and responsiveness to change, in patients with lower extremity dysfunction.

METHODS

Subjects

Subjects were recruited from area orthopaedic surgery offices and physical therapy offices by way of information flyers that were distributed to these offices. Inclusion criteria for this study were adult individuals, 21-65 years of age, with a unilateral lower extremity dysfunction resulting from an injury or surgery. Subjects needed to be currently involved in some form of physical therapy or a home program for rehabilitation. In addition, subjects needed to be willing to report to the University's research laboratory for all data collection, on at least two separate occasions, at least four weeks apart. Subjects were excluded if they were not able to walk independently and ascend/descend stairs with full weight bearing, or if subjects were non weight bearing on the involved lower extremity. In addition, subjects were excluded if at least 0-90 degrees of knee flexion range of motion was not available at the time of testing. All subjects signed an informed consent document, and this study was approved by the University of San Diego State Institutional Review Board for Human Subjects Research. Subjects were compensated monetarily for the time and expenses required to participate in this study.

Apparatus

For the PWB performance tests, this study used a Total Gym 26000 (Engineering Fitness International, San Diego, CA). This device consists of a sliding board apparatus that is mounted to a rail system. The rail system is fixed to a vertical upright stand and the rail can be positioned at angles of 10 degrees to 50 degrees to the horizontal (floor surface). Positioning the sliding board at an angle of 50 degrees to the horizontal provides approximately 65% of the individual's body weight as resistance *(Figures 1 and 2)* according to manufacturer's specifications and based on the following.

The slide distance regulator (Engineering Fitness International, San Diego, CA) was used to restrict the displacement of the sliding board apparatus in the downward direction. The slide distance regulator was also used in order to control the total knee ROM during a single squat repetition. A standard stop watch was used to record time for all tests. A standard goniometer was used to record

Figure 1. *The sliding incline apparatus with subject in full extension.*

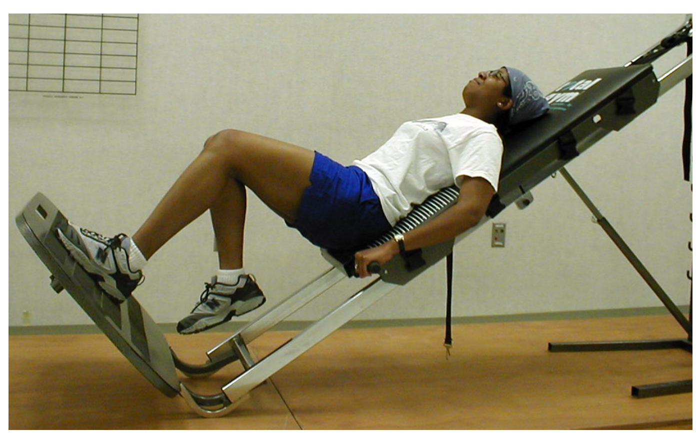

Figure 2. *The sliding incline apparatus with subject at full allowable flexion of the knee.*

knee joint ROM during the PWB performance tests. A standard tape measure was used to measure linear distance of the one-legged hop tests.

Procedures

Prior to any testing, subjects first performed single knee squat repetitions on the Total Gym device with the involved lower extremity, in order to determine an appropriate level for testing (i.e., angle of inclination and, therefore, appropriate body weight resistance), as well as to assure that subjects could perform a knee squat through a range of 0-90 degrees of knee flexion. The criteria for inclination level was an ability to perform ten consecutive one-legged squats without pause, through a ROM of at least 0-60 degrees of knee flexion *(Figures 1 and 2)*. This ROM was selected because the range is necessary for normal stair climbing. However, the preferred test range of motion was 0-90 degrees of knee flexion, in accordance with Munich et al.[21] Ten repetitions were chosen for the screening based on the investigators' clinical experience and the opinion that this would be a safe level for testing. Once the appropriate inclination level, and the comfortable knee ROM were determined, the slide distance regulator was secured to the sliding board in order to assure that knee flexion ROM did not exceed the maximum knee flexion available for that patient. Knee ROM was measured with the goniometer using accepted procedures.[22]

Subjects were then randomly assigned to a sequence of lower extremity performance tests. The tests included the following:

1) Repetitions completed during the 20 second test of single leg squats on the Total Gym (PWB20)
2) Time (seconds) to complete the 30 repetition test on the Total Gym (PWB30)
3) Time (seconds) to ascend a flight of stairs
4) Time (seconds) to descend a flight of stairs
5) Time (seconds) to walk 15 meters
6) Distance (centimeters) of a single one-legged hop

The sequence of tests were randomly determined using a 5 x 5 design; the stair climbing tests were considered as one test in this design, given that subjects would naturally need to descend and ascend stairs during a test. However, for the purpose of data analysis, ascent and descent scores were considered separately. Prior to testing, subjects were provided warm-up times with either walking or performing two-legged squat exercises, as appropriate, on the Total Gym apparatus. Subjects were provided three to five minutes of rest between each test. On day one, subjects performed all tests in random order. A subset of 15 subjects was selected randomly to perform the tests a second time, on the first testing day, for test-retest reliability analysis. All subjects were asked to return for follow-up testing four to six weeks following the first day of testing. The follow-up testing was to examine the responsiveness, of the PWB performance tests, change in status of the subjects. It was expected that changes would occur in the patients, following four to six weeks of physical therapy or home exercise intervention. Given these improvements, the PWB performance tests should also reflect this improvement. While this study did not control the interventions that were provided, it is reasonable to expect that patients would improve over time. The incline level and knee ROM of the PWB20 and PWB30 tests were maintained for the follow-up testing. The same licensed physical therapist (20 years outpatient clinical experience), trained in the administration of the PWB20 and PWB30 tests, performed all the measures of all the patients.

Measures

Twenty-second Squat Repetition test on the Total Gym (PWB20)

This test involved the subject performing as many single-leg squat repetitions as possible in a 20 second time period. The subject squatted from 0 degrees of knee extension to a maximum of 90 degrees of knee flexion *(Figures 1 and 2)*. If the subject was unable to flex the knee to 90 degrees the subject was asked to flex to a comfortable position. This position was measured with a 12-inch goniometer, using standard procedures as described by Norkin and White,[22] for knee ROM measurement, and the slide distance regulator was used to assure that knee ROM did not exceed the maximum comfortable level of knee flexion. The subject was instructed on how to perform the proper squatting technique and he/she was asked to practice the squatting technique prior to beginning the test. Subjects were instructed to move from an extended knee position into a flexed knee position until the subject felt minor resistance from the slide distance regulator. After becoming familiar with the test, the subject rested for one minute before starting the actual test. Subjects were encouraged to squat at the fastest pace they felt safe performing. During the test, the researcher counted the number of squats performed by the subject in a 20-second time period.

Timed 30 single leg squat repetition test (PWB30)

This test required the subject to perform 30 single leg squat repetitions on the Total Gym. Thirty repetitions were chosen in place of the original 50 repetitions because pilot testing demonstrated that 50 repetitions required too much effort from a patient in the early stages of recovery. The actual procedures are identical to the PWB20, with the exception that subjects were instructed to continue squatting until 30 full repetitions were completed. The time (seconds) required to complete 30 repetitions was recorded by the investigator. In the event that a subject needed to stop and rest or slow down, the time continued to be recorded until all 30 repetitions were completed. Five subjects needed to stop and rest momentarily during the first day of testing; rest was not needed for any subject during the follow-up testing four to six weeks later. Subjects were encouraged to squat at the fastest pace they felt safe performing.

Timed ascending stair test (Stair UP)

This test required the subject to ascend a single flight of stairs (24 steps). The subject was instructed to ascend the stairs as rapidly as possible while remaining safe. The researcher used a stopwatch to determine the amount of time (seconds) the subject took to ascend the fight of stairs. Subjects could use an assistive device (straight cane or quad cane) and the railing, if needed.

Timed descending stair test (Stair DOWN)

This test required the subject to descend a single flight of stairs (24 steps). This test followed the Stair UP test for all participants. On completion of the Stair UP test, the subject was then instructed to descend the stairs as rapidly as possible while remaining safe. The researcher used a stopwatch to determine the amount of time (seconds) the subject took to descend the fight of stairs. Subjects could use an assistive device (straight cane or quad cane) and the railing, if needed.

Walk test

For this test, the subject was asked to walk 30 meters at a comfortable pace. During the 30-meter walk, two distinct points, 15 meters apart, were used for measurement of walking speed. When the subject's heel reached the first mark, the researcher started the stopwatch. The stopwatch was stopped when the subject's heel reached the second mark, and the time (seconds) was recorded.

Single-leg hop test

To complete this test, subjects performed a maximal single-leg hop. The subject was instructed how to properly perform the test. Prior to the test, the subject performed two practice hops. The subject began the test with toes behind a starting line, and a maximal hop was performed. Upper extremity movement and position were not controlled by the researcher. The researcher then measured the distance from the starting line to the subject's heel. The single-leg maximal hop was conducted two separate times during the actual test. The maximal distance of these two, or best score, was used for data analysis.

Additional data

The subject's age, gender, diagnosis, onset of injury (i.e., time since injury), and treatment type (i.e., home program or formal clinical physical therapy) were recorded. This information was self-reported by the subject.

Data Analysis

Reliability study

Relative and absolute reliability of the data from the PWB performance tests (PWB20 and PWB30 tests) and functional tests was estimated using the test-retest data of day one. Relative reliability measures the test-retest consistency of the data by establishing a coefficient value (intraclass correlation coefficient). This coefficient value is then compared to an established criteria for acceptable reliability. Absolute reliability involved estimating the actual error in the measure, in the original units of measure. The absolute reliability provides information regarding the expected error in the measure. Re-testing occurred approximately 30 minutes following the initial test. In order to evaluate the relative reliability of the data the Intraclass Correlation Coefficient (ICC 3,1) was used to estimate intrarater reliability.[23,24] A lower one-sided 95% confidence value was constructed using SPSS version 11.0. It is the lower bound value of the 95% confidence interval that is of clinical importance for the ICC, because this represents the lowest possible relative reliability. In order to estimate absolute reliability of the measures, the standard error of measurement (SEM) was estimated based on:

s_x is the standard deviation of the measure and rx_x was the ICC derived in the test-retest portion of the study. An upper one-sided 95% confidence value was constructed for the SEM.[25] It is the upper bound value of the 95% confidence interval that is of importance clinically for the SEM, because this represents the highest possible value of error in the measure. The SEM was then used for the calculations of the minimal detectable change (MDC) with a 95% level of confidence, based on the procedures described by Stratford et al.[26] The MDC is an estimate of the absolute change in a measure that is required to be clinically meaningful. The MDC95 was estimated using the following formula:

In this case, z = 1.96 is the z-score associated with a 95% confidence interval, and the value of 2.0 is a correction factor accounting for error over two testing occasions.[26] The MDC was used to estimate the 95% confidence in the data that a clinically significant change occurred over time.

Validity study

Two elements of validity evidence were examined: concurrent validity evidence and responsiveness to change validity evidence. Concurrent validity evidence was assessed by comparing the PWB performance tests of the involved leg with known measures of function that included walking speed, stair ascending/descending speed, and hop performance. The values of all tests were evaluated using the Pearson's product moment correlation coefficient for bivariate correlations. Responsiveness validity evidence was examined using two procedures: 1) a two-factor (2x2) analysis of variance (ANOVA) compared the rate of change of the involved leg with the uninvolved leg (i.e., known groups method) on the PWB performance tests; and 2) a simple repeated measures ANOVA compared the relative change scores of the involved leg on the PWB performance tests with the relative change scores of the known measures of function (walking speed, stair ability speed, and hop distance). Paired t-tests, comparing the initial measurement values with the follow-up values, were also used to test whether subjects improved on the four measures of function.

Finally, the limb symmetry index was calculated, by obtaining the ratio of the involved leg raw score with the uninvolved leg score on the PWB performance tests, for the PWB performance tests. The limb summary index is a useful measure in that it accounts for changes in both lower extremities (i.e., involved and uninvolved) over time, to estimate the relative performance of the involved limb compared with the uninvolved limb.[27] The limb summary index at initial test was then compared with the limb summary index at follow-up using a simple repeated measures ANOVA, and planned repeated contrasts were used to test for differences between the PWB20 and PWB30 limb symmetry index (LSI) values.

RESULTS

Subject Demographics

Forty-four subjects originally volunteered to participate in this study. Seven subjects were excluded because they presented with bilateral lower extremity symptoms. Thirty-seven subjects completed the initial testing. Fifteen of these subjects were retested on the initial day to assess reliability of the data. Data from the 15 subjects were used for the reliability study and data from the 37 subjects were used for the correlation matrix. Of the original 37 subjects, only 23 subjects completed the second phase of testing after four to six weeks for follow-up. Data from the 23 subjects who completed both the initial and follow-up testing were used for the responsiveness to change analysis. The 14 subjects who did not complete the follow-up were excluded because they were not involved in any form of rehabilitation (i.e., formal clinical therapy or home therapy, n = 12) or they did not return for follow-up (n = 2). The two subjects who did not return for follow-up did not want to travel the distance for the follow-up test. Data of the remaining 23 subjects were then used for the responsiveness to change analysis. Eight of these 23 subjects maintained a regular physical therapy rehabilitation program, while 15 subjects continued with a home exercise program.

Table 1 provides the demographic information of the subjects, including age, sex, and time since original injury/dysfunction. Table 2 provides a distribution of the self-reported diagnoses of the subjects

Table 1: Subject descriptive statistics

Session	n	Female	Age (sd)	Days since onset (sd)	Receiving physical therapy
Initial	37	62.2%	49.8 (14.7)	155.6 (164.9)	10/37
Follow-up	23	61.0%	48.4 (13.1)	153.9 (164.3)	8/23
Reliability	15	60.0%	49.1 (14.2)	154.8 (164.4)	5/15

as a list of the physical diagnoses by self-report of all subjects. Knee joint pain refers to those subjects who reported either "arthritis" or "internal knee pain" as their reason for physical therapy consultation. The time, in days, of onset was estimated by each subject. For surgical cases (i.e., knee joint replacement, ACL surgery, etc.), the date of surgery served as the time since onset. For all other conditions, acute and chronic, the subject provided a best estimate of duration of symptoms.

Table 2: Distribution of diagnoses, by self-report, at initial test of study and at follow-up.

Self-Report Diagnosis	Initial (n)	Follow-up (n)
Knee joint pain	10	5
Patella-femoral dysfunction with/without lateral release	8	4
Anterior cruciate ligament reconstruction	7	6
Total knee arthroplasty	6	4
Achilles tendon rupture and repair	2	2
Iliotibial band syndrome	2	0
Tibial plateau fracture	1	1
Hamstring tear	1	1
Total	37	23

Finally, an insufficient number of subjects completed the hop test (n = 8), prohibiting any meaningful statistical analyses. Hence, findings on the hop tests are not included.

Reliability Study

The ICC's for the PWB performance tests, walking time, stair times, and hop test are reported in Table 3. All point estimates for the ICC's were greater than 0.90, and the lower bound of the 95% confidence interval exceeded 0.70 for all measures.[16] In addition, Table 3 provides the SEM for the PWB performance tests. Based on the SEM, the MDC90 is also presented in Table 3 for the two PWB performance tests.

Validity Study

Concurrent validity

The results of the bivariate correlation analysis testing between the two PWB performance tests and the four measures of function, on the initial day of testing, are presented in Table 4. All correlation coefficients were significant at $p < 0.05$. Negative correlations were identified between the number of one-legged squat repetitions that a subject could perform in 20 seconds with the time needed to walk or ascend/ descend stairs (r = -0.72 to -0.60). Hence, repetitions were inversely related to time; more repetitions in 20 seconds were related to less time needed to walk or ascend/ descend. Similarly, a positive correlation was identified between the time required to complete 30 one-legged squat repetitions and the time needed to walk and ascend/descend stairs (r = 0.61 to 0.49). In all cases, the PWB20 test had slightly higher bivariate correlation coefficient values than the PWB30 test, with the three measures of function. As expected, the two PWB performance tests were correlated with each other (inversely), and the stair climbing tests were correlated with each other. Walking was also correlated with stair climbing.

Table 3: Intraclass correlation coefficients (ICC), standard error of measurement (SEM), and minimal detectable change (MDC) for select measures. (n = 15)

Measure	ICC (lower 95% CI)	SEM (upper 95% CI)	MDC
PWB20	0.95 (0.83)	1.35 (1.86)	3.74
PWB30	0.98 (0.88)	1.23 (1.78)	3.41
Walk Time	0.99 (0.97)	---	---
Stairs UP	0.98 (0.96)	---	---
Stairs DOWN	0.96 (0.92)	---	---
Hop Distance	0.97 (0.78)	---	---

Responsiveness to change

Follow-up testing of the four performance tests and the two PWB performance tests occurred on average 30.27 days (sd = 2.94 days) post initial test with a minimum and maximum of 27 and 36 days, respectively. Subjects were tested at the same level on the Total Gym as their initial test Total Gym level. All subjects were initially tested at either level 8, 9, or 10, which coincided with 50-65% of body weight, on the Total Gym. Follow-up testing was performed at the same level. Average knee flexion for all PWB performance tests was 70.22 degrees (sd = 4.07 degrees) with a minimum and maximum of 60 degrees and 83 degrees. The results of the two-way

repeated measures ANOVA for the PWB20 test and for the PWB30 test revealed that the involved leg demonstrated significantly greater changes in performance compared with the uninvolved leg ($p < .05$ for the interaction term in both PWB performance tests). Table 5 provides the mean values at initial and at follow-up for the two PWB performance tests. In addition, pairwise t-tests revealed that all subjects improved in walking speed, stair climbing speed, and hop distance ($p < .05$). Mean measures for initial and follow-up are also displayed in Table 5 for these measures. Finally, the absolute change values are reported, for comparison to the MDC95.

Table 4. Correlation matrix for all measures. (n = 37)

	PWB20	PWB30	Stairs UP	Stairs DOWN	Walk
PWB20	-	-0.78	-0.60	-0.61	-0.72
PWB30			0.49	0.50	0.61
Stairs UP				0.96	0.78
Stairs DOWN					0.78
Walk					

Relative change of each of the functional tests was similar to the relative change in the PWBP tests for the involved lower extremity. Simple repeated measures to test these values revealed no significant differences in relative change scores ($p > .05$). Table 5 provides the relative change for each measure, expressed as a percentage. The uninvolved limb relative change scores were significantly less than the other relative change scores ($p < .05$). Finally, the limb symmetry index for the PWB performance tests changed significantly when tested with a simple repeated measure ANOVA ($p < .05$). Based on the planned repeated contrasts, the LSI increased from initial test values to the follow-up values *(Table 5)*, for both the PWB20 test and the PWB30 test ($p < .05$). The LSI of the PWB20 was greater than the LSI of the PWB30 at both the initial test and at follow-up ($p < .05$).

Table 5. Mean measures for the partial weight bearing performance tests and the measures of function (sd).

Test	Initial	Follow-up	Absolute Change	Relative Change (%)
PWB20 Uninvolved	22.1 (5.5)	23.0 (6.2)	1.1 (1.6)	4.1†
PWB20 Involved	17.1 (5.7)	21.9 (6.7)*	4.8 (2.7)	28.1
PWB30 Uninvolved	30.4 (9.1)	29.3 (9.3)	1.1 (1.4)	3.7†
PWB30 Involved	43.2 (24.3)	33.9 (12.9)*	9.2 (6.3)	27.4
Stair UP	13.3 (5.9)	10.7 (4.1)*	2.7 (3.0)	24.3
Stair DOWN	14.0 (7.3)	11.5 (6.5)*	2.6 (2.5)	21.7
Walk	10.2 (2.7)	8.6 (2.8)*	1.8 (1.2)	19.0
LSI PWB20	77.3 (9.7)	95.2 (7.9)*	17.9 (8.9)	23.0
LSI PWB30	70.3 (21.1)	86.4 (17.4)*	16.1 (7.7)	23.0

PWB20 = repetitions
PWB30, Stair, Walk = time (sec.)
LSI = limb symmetry index (%)
* Significantly different from initial values, $p < .05$.
† Significantly less than all other relative change estimates, $p < .05$.

DISCUSSION

The purpose of this study was to evaluate the measurement properties of two PWB measures of lower extremity performance. The two tests, both involving a single legged squat, were performed on a Total Gym, a device that allowed the measures to be performed at less than 100% of the subject's body weight. In fact, all subjects performed the PWB tests at approximately 65% of body weight. The measurement properties evaluated included absolute and relative reliability, as well as validity evidence in the form of concurrent validity and responsiveness to change. A heterogeneous sample of patients participated in this study, with lower extremity conditions ranging from patellofemoral dysfunction to total knee arthroplasty surgery.

The relative intra-tester test-retest reliability of the two PWB measures assures reliability, with ICC values exceeding 0.90.[23] In addition, the absolute reliability, as estimated with the SEM, was also excellent, with upper 95% SEM values less than 2.0 for either measure (i.e., PWB20 or PWB30). The ICC values exceed those reported by Munich et al.[21] It is likely that the heterogeneous sample in this current study contributed to the improved ICC values. Munich et al[21] used a homogeneous sample of college-aged healthy adults. In addition, the test-retest ses-

sion for this current study was separated by only 30 minutes, whereas the Munich et al[21] study performed re-testing one full week later. Regardless, the ICC values and the low SEM suggest that the two PWB performance tests provide reliable data in terms of intra-tester reliability.

Concurrent validity evidence was estimated by correlating the PWB performance test measures with measures of walking speed, stair climbing and descending speed, and maximum one-legged hop distance. The data from the initial test were used to evaluate these relationships. This study found correlations between the PWB performance tests and the functional tests. In fact, the correlation coefficients were higher than those found for isokinetic testing.[14-17] This finding is not surprising given that the PWB performance tests involve the entire lower extremity (ankle, knee, hip), whereas the isokinetic tests used in previous studies were measures of isolated lower extremity muscle group function. The PWB performance tests more closely replicate the interaction between multiple joint systems during a functional activity, and hence the measures of the PWB performance tests better correlate with walking and stair climbing speed, compared with single joint system tests. In addition, as noted by Aasa et al,[28] body size influences muscle strength assessment. Isokinetic tests are dependent on leg/limb mass whereas the PWB performance tests are dependent on total body mass.

The direction of the correlations for the PWB performance tests with the functional measures of walking and stair climbing also make sense. For instance, the PWB20 test, which is a measure of maximum repetitions, demonstrated negative correlations with walking speed and stair climbing speed. The more repetitions a person could complete in 20 seconds, the less time that person would need to walk the established walk distance and to ascend/descend a flight of stairs. Conversely, the PWB30 test positively correlated with these functional measures. The less time needed to complete 30 repetitions of the PWB30 test, the less time needed to also walk a set distance and ascend/descend a flight of stairs.

This study was not able to evaluate the relationship between the PWB performance tests and one-legged hop ability. Only eight subjects were able to complete the hop trials. The hop tests were considered too advanced by most of the subjects, at their present stage of recovery.

Additional validity evidence was estimated in the form of responsiveness to change. Worrell et al[18] noted that isokinetic testing may not be responsive to changes in lower extremity function, when weight bearing exercise protocols are involved in the rehabilitation. The two PWB performance tests demonstrated good responsiveness to change. The MDC values of 3.74 and 3.41, respectively, of the PWB20 and PWB30 were exceeded by the average absolute changes on both measures (4.8 and 9.2, respectively). Hence, the PWB performance tests are able to measure improvement/change in patients' lower extremity function, if indeed changes have occurred because of rehabilitation and/or time. Initial test and follow-up test measures on the PWB performance tests were significantly different from each other, as were walking speed and stair climbing speed. All measures improved significantly. In fact, relative changes for all measures were similar ($p > .05$).

Additional responsiveness evidence is provided by the change in the limb symmetry index (LSI). It was expected that the LSI would improve, given that the uninvolved lower extremity was not expected to improve as well as the involved lower extremity. In this study, the uninvolved lower extremity did not demonstrate any significant changes on the one-legged tests (PWB performance tests). Thus, the LSI improved significantly, based on both PWB performance tests. The initial test LSI of our subjects are comparable with those reported by Wilk et al,[27] when testing leg strength isokinetically. In their study, the LSI based on isokinetic testing was less than 85% in the majority of their subjects. Following four to six weeks of time, the subjects in this study demonstrated improved performance of the involved lower extremity, greater than the changes in the uninvolved lower extremity, as evidenced by the improved LSI.

Several limitations exist to this current study. The type of rehabilitation that each patient received was not controlled. In addition, the influence of rehabilitation approaches and lack of formal rehabilitation in terms of the outcomes achieved was not accounted for. In fact, as noted in the results, only eight of the 24 returning subjects received formal physical therapy. The majority performed physical therapy prescribed home exercise programs. Another limitation is the wide range of diagnoses included in this study. Given that the average time since onset was nearly four months, whether the PWB performance tests are better suited for acute or chronic conditions could not be determined. Future research should evaluate the PWB perform-

ance tests on individuals with acute conditions separately from chronic status, as well as analyze patients by diagnostic groups. Finally, even though the PWB performance tests are intended for the early stages of recovery and for individuals with PWB status, the subjects in this study had full weight bearing (FWB) status. The FWB status was necessary in order to test walking speed and stair climbing speed for the concurrent validity evaluation. Yet, even with FWB status, only eight of the original 37 subjects were willing to perform the hop tests. The hop tests were either considered too aggressive or subjects were afraid to try to hop. While this is a limitation to our study, it is also evidence of the need for a controlled weight bearing measure of performance.

Hence, in order to determine if the PWB performance tests are appropriate for a patient population with acute presentation or PWB status, future research is needed that involves this population. Our study included individuals with lower extremity pathology, however, all were FWB. Further research is needed on a sample of patients with actual PWB limitations.

CONCLUSION

Two partial weight bearing one-legged squats tests were evaluated for measurement properties of reliability and validity. The two partial weight bearing performance tests, the PWB20 and PWB30, demonstrated sufficient intrarater test-retest reliability. In addition, this study provides evidence of validity of these measures to estimate lower extremity performance. The two tests correlate with walking and stair climbing speed. In addition, the two tests are responsive to changes in condition and provide an indication of leg symmetry. These partial weight bearing performance tests might be suitable for the orthopaedic setting, as a means of patient examination of function in a partial weight bearing position. Clinicians should use caution in interpreting the findings of an evaluation with the partial weight bearing performance tests until further research with specific patient populations and acute status have been completed.

REFERENCES

1. Augustsson J, Thomee R. Ability of closed and open kinetic chain tests of muscular strength to assess functional performance. *Scand J Med Sci Sports.* 2000;10:164-168.
2. Pfeifer K, Banzer W. Motor performance in different dynamic tests in knee rehabilitation. *Scand J Med Sci Sports.* 1999;9:19-27.
3. Ageberg E, Zatterstrom R, Moritz U. Stabilometry and one-leg hop test have high test-retest reliability. *Scand J Med Sci Sports.* 1998;8:198-202.
4. Bergmann G, Deuretzbacher G, Heller M, et al. Hip contact forces and gait patterns from routine activities. *J Biomech.* 2001;34:859-871.
5. Bolgla L, Keskula D. Reliability of lower extremity functional performance tests. *J Orthop Sports Phys Ther.* 1997;26:138-142.
6. Fitzgerald G, Lephart S, Hwang J, et al. Hop tests as predictors of dynamic knee stability. *J Orthop Sports Phys Ther.* 2001;31:588-597.
7. Jarvela T, Kannus P, Latvala K, et al. Simple measurements in assessing muscle performance after an ACL reconstruction. *Int J Sports Med.* 2002;23:196-201.
8. Kea J, Kramer J, Forwell L, et al. Hip abduction-adduction strength and one-leg hop tests: Test-retest reliability and relationship to function in elite ice hockey players. *J Orthop Sports Phys Ther.* 2001;31:446-455.
9. Mittlmeier T, Weiler A, Sohn T, et al. Functional monitoring during rehabilitation following anterior cruciate ligament reconstruction. *Clin Biomech.* 1999;14:576-584.
10. Radcliffe S, Smith D. The Mennen plate in periprosthetic hip fractures. *Injury.* 1996;27:27-30.
11. Rudolph K, Axe M, Snyder-Mackler L. Dynamic stability after ACL injury: Who can hop? *Knee Surg Sports Traumatol Arthrosc*. 2000;8:262-269.
12. Tegner Y, Lysholm J, Lysholm M, et al. A performance test to monitor rehabilitation and evaluate anterior cruciate ligament injuries. *Am J Sports Med.* 1986;14:156-159.
13. Yu B, Growney E, Hansen D, et al. Abduction/adduction moment at the knee in stair climbing and level walking. *Gait Posture.* 1996;4:190.
14. Andrade MS, Cohen M, Picarro IC, et al. Knee performance after anterior cruciate ligament reconstruction. *Isokinet Exerc Sci.* 2002;10:81-86.
15. Greenberger HB, Paterno MV. Relationship of knee extensor strength and hopping test performance in the assessment of lower extremity function. *J Orthop and Sports Phys Ther.* 1995;22:202-206.
16. Ozcakar L, Kunduracyoolu B, Cetin A, et al. Comprehensive isokinetic knee measurements and quadriceps tendon evaluations in footballers for assessing functional performance. *Br J Sports Med.* 2003;37:507-510.
17. Petschnig R, Baron R, Albrecht M. The relationship between isokinetic quadriceps strength test and hop tests for distance and one-legged vertical jump test following anterior cruciate ligament reconstruction. *J Orthop and Sports Phys Ther.* 1998;28:23-31.

18. Worrell TW, Borchert B, Erner K, et al. Effect of a lateral step-up exercise protocol on quadriceps and lower extremity performance. *J Orthop and Sports Phys Ther.* 1993;18: 646-653.

19. Sole G, Milosavljevic S, Nicholson H, Sullivan S. Test-retest reliability of isokinetic knee extension and flexion. *Arch Phys Med Rehabil.* 2007;88:626-631.

20. Bohannon R. Manual muscle test scores alone provide limited indication of the functional independence of home care patients. *Int J Rehabil Res.* 1999;22:151-152.

21. Munich H, Cipriani D, Hall C, et al. The test-retest reliability of an inclined squat strength test protocol. *J Orthop and Sports Phys Ther.* 1997;26:209-213.

22. Norkin C, White D. *Measurement of Joint Motion: A Guide to Goniometry.* 3rd ed. Philadelphia, PA: FA Davis; 2003.

23. Portney LG, Watkins MP. *Foundations of Clinical Research*. 2nd ed. Upper Saddle River, NJ: Prentice Hall Health; 2000.

24. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull.* 1979;86:420-428.

25. 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.

26. Stratford PW, Binkley JM, Solomon P, et al. Defining the minimum level of detectable change the Roland-Morris Questionnaire. *Phys Ther.* 1996;76:359-368.

27. Wilk KE, Romaniello WT, Soscia SM, et al. The relationship between subjective knee scores, isokinetic testing, and functional testing in the ACL-reconstructed knee. *J Orthop and Sports Phys Ther.* 1994;20:60-73.

28. Aasa U, Jaric S, Barnekow-Bergkvist M, et al. Muscle strength assessment from functional performance tests: Role of body size. *J Strength Cond Res.* 2003;17:664-670.