Physical Performance Assessment in Military Service Members

Jason M. Wilken, PT, PhD
Benjamin J. Darter, PT, PhD
Stephen L. Goffar, PT
Jesse C. Ellwein, PT
Rachel M. Snell, DPT, CSCS
Eric A. Tomalis, DPT, CSCS
Scott W. Shaffer, PT, PhD

Abstract

Few established measures allow effective quantification of physical performance in severely injured service members. We sought to establish preliminary normative data in 180 healthy, active-duty service members for physical performance measures that can be readily implemented in a clinical setting. Interrater and test-retest reliability and minimal detectable change (MDC) values were also determined. Physical performance testing included self-selected walking velocity on level and uneven terrain, timed stair ascent, the sit-to-stand five times test, the four-square step test, and the 6-minute walk test. Data analysis included descriptive statistics, intraclass correlation coefficients, and MDC. Interrater and test-retest reliability were excellent for all measures (intraclass correlation coefficients >0.75). MDC values for timed measures were <0.3 seconds for interrater comparisons and <1.5 seconds for between-day comparisons. Physical performance measures had a narrow range of normal performance and were reliable and stable between days.

From the Center for the Intrepid, Department of Orthopaedics and Rehabilitation, Brooke Army Medical Center (Dr. Wilken and Dr. Darter) and the US Army-Baylor University Doctoral Program in Physical Therapy, Fort Sam Houston, TX (Dr. Wilken, Dr. Darter, Mr. Goffar, Mr. Ellwein, Ms. Snell, Mr. Tomalis, and Dr. Shaffer).

This project was funded by a grant from the Military Amputee Research Program. The views expressed in this article are those of the authors and do not reflect the official policy or position of Brooke Army Medical Center, the US Army Medical Department, the US Army Office of the Surgeon General, the Department of the Army, the Department of Defense, or the US Government.

J Am Acad Orthop Surg 2012; 20(suppl 1):S42-S47

http://dx.doi.org/10.5435/JAAOS-20-08-S42

The prevalence of severe extremity trauma sustained during the conflicts in Iraq and Afghanistan as well as the potential long-term costs associated with these injuries have been well documented.[1-4] Understanding of the time course of recovery associated with these combat injuries is limited, and limited evidence specifically quantifies the effect of surgical and rehabilitative interventions on physical function in patients with these injuries.[5,6] Efforts are underway to quantify the efficacy of existing interventions to improve patient care. However, clinicians are confronted by the current paucity of established measures that allow for effective characterization of physical performance in both healthy and injured young adults.

Outcomes assessment following extremity trauma is typically performed using qualitative assessments and questionnaires that rely on patient self-report.[7-10] Although these measures are useful for characterizing vital global outcomes such as quality of life, they typically lack sufficient resolution to fully describe disability, track gradual improvements over time,[11] and link interventions to a change in a specific aspect of physical performance.

Quantitative assessment of physical performance in patients with severe extremity trauma typically includes basic functional activities such as rising from a chair and walking a short distance.[12-14] Although appropriate for older adults (age >65 years) or more impaired patients,

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
AUG 2012		**00-00-2012 to 00-00-2012**

4. TITLE AND SUBTITLE
Physical Performance Assessment in Military Service Members

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brooke Army Medical Center,Department of Orthopaedics and Rehabilitation,Fort Sam Houston,TX,78234

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Journal of the American Academy of Orthopaedic Surgeons, August 2012 vol. 20 no. suppl S42-S47

14. ABSTRACT
Few established measures allow effective quantification of physical performance in severely injured service members. We sought to establish preliminary normative data in 180 healthy, active-duty service members for physical performance measures that can be readily implemented in a clinical setting. Interrater and test-retest reliability and minimal detectable change (MDC) values were also determined. Physical performance testing included self-selected walking velocity on level and uneven terrain, timed stair ascent, the sit-to-stand five times test, the four-square step test, and the 6-minute walk test. Data analysis included descriptive statistics intraclass correlation coefficients, and MDC. Interrater and testretest reliability were excellent for all measures (intraclass correlation coefficients >0.75). MDC values for timed measures were <0.3 seconds for interrater comparisons and <1.5 seconds for between-day comparisons. Physical performance measures had a narrow range of normal performance and were reliable and stable between days.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**7**	

these measures typically demonstrate ceiling effects in more highly functioning persons (eg, athletes, military service members). In contrast, routine tests used to assess athletic populations often require significant areas of open space (eg, obstacle course, agility run) and can include activities that may place more severely involved patients at risk of further injury.[15]

To address the need for reliable and relevant physical performance assessment measures, we sought to establish preliminary normative data for young, healthy service members and determine the reliability for measures that are readily implemented in a clinical setting. Selected measures were thought to be likely to reflect anticipated improvements resulting from commonly used interventions; be usable across a broad continuum of injury severity; demonstrate excellent reliability and sufficient resolution to detect relevant changes in performance; be implementable with minimal cost and training; and enable effective characterization of physical performance across a range of functional domains, including agility, mobility, balance, power, and exercise capacity.

Methods

This descriptive and repeated-measures study included 180 healthy, active-duty service members training at Fort Sam Houston, Texas. Participants ranged in age from 18 to 43 years and had no current or recent history (within 6 months) of medical or neuromusculoskeletal disorders that limited participation in their military occupation specialty or physical training activities. Participants were excluded if they did not have full, pain-free motion of the spine and lower extremities or could not complete heel-toe walking, five deep squats, and five single-leg hops on each limb. The study was approved by the Brooke Army Medical Center Institutional Review Board, and all participants provided written consent.

Participants were randomly assigned to one of four physical performance testing stations, including self-selected walking velocity (SSWV) on level ground and on a loose-rock surface, timed stair ascent (TSA), the sit-to-stand five times (STS5) test, and the four-square step test (FSST). Testing order was counterbalanced and, after completing the exercise at each of the four physical performance stations, participants completed the 6-minute walk test (6MWT). Instruction, demonstration, and data collection for all measures were done by four physical therapists and five physical therapy students. To ensure uniformity of testing procedures, testers received 2 hours of training in the administration of the measures by two of the authors (S.W.S. and S.L.G.). To assess interrater reliability, two raters recorded performance for 25 participants as they completed all tests. To assess test-retest reliability, 20 participants returned for repeat testing 5 to 10 days after their initial assessment.

Study Measures

Measurements included in the study reflect activities required for basic mobility and participation in civilian and military environments. For example, the SSWV test allows direct assessment of mobility, fall risk, and disability.[16] SSWV on a loose-rock surface ($SSWV_{RS}$) was examined secondary to its importance as a critical task for optimal performance in military and athletic environments. In our study, participants were instructed to walk a distance of 20 meters at a normal, comfortable pace. Gait speed was calculated based on the time required to cross the middle 10 meters of the walking path. The rocky path consisted of stones measuring approximately 1 to 2 inches in diameter and ≥3 inches deep. The averages of three walking tests over ground ($SSWV_{OG}$) and three $SSWV_{RS}$ trials were used in the final analysis.

TSA is often used as an objective measure of mobility and power in older adults (age >65 years).[17] However, limited research exists regarding normal TSA in younger adults. We define TSA as the ability to ascend 12 steps as quickly as possible, without using the hand railing, touching every step with at least one foot (ie, alternating steps). Timing began on the word “go” and stopped when both of the participant’s feet were on top of the 12th step. Participants performed three test trials and were allowed 1 minute of rest between trials.

The STS5 test is commonly used to assess lower extremity strength and mobility in older adults (age >65 years),[18,19] but critical analysis of this measure in younger adults is lacking. Proper completion of the STS5 test requires participants to fully stand up and sit down five times as fast as possible with the arms crossed over the chest. Participants completed two trials and were allowed 1 minute of rest between trials.

Mobility and dynamic balance

None of the following authors or any immediate family member has received anything of value from or has stock or stock options held in a commercial company or institution related directly or indirectly to the subject of this article: Dr. Wilken, Dr. Darter, Mr. Goffar, Mr. Ellwein, Ms. Snell, Mr. Tomalis, and Dr. Shaffer.

were also assessed with the FSST. This measure has excellent interrater and test-retest reliability (intraclass correlation coefficient [ICC] = 0.99 and 0.98, respectively) and is a valid measure for identifying fall risk in older, community-dwelling adults, older adults who have sustained a transtibial amputation,[20] and adults with vestibular dysfunction.[21] In the FSST, the patient steps sequentially over four 1-in–diameter sticks or canes that are placed flat on the floor in the shape of a cross. In our study, participants began in the left rear square and were required to step over each cane as they moved as fast as possible in the following pattern: (1) forward, (2) sidestep to the right, (3) backward, and (4) sidestep to the left. They returned by sidestepping to the right, forward, left, and backward. Timing began when the participant's foot was placed in the box in front of him or her and was stopped when both feet were placed in the final box. Participants completed one practice trial followed by four test trials.

The 6MWT is frequently used to assess aerobic fitness, endurance, and mobility by measuring the distance a person can walk on a level surface in 6 minutes.[22,23] Normative times have been established across age spectrums,[23-25] but test-retest reliability remains unclear.[26,27] Participants were instructed to walk, not run, as far as they could for 6 minutes on a level, oval-shaped walking path.

Data Analysis

SPSS for Windows, version 17.0 (SPSS, Chicago, IL) was used for data analysis. Descriptive statistics included mean, standard deviation, quartiles, and 5th and 95th percentiles for each measure. With the exception of the 6MWT, for which only one trial was conducted per participant, mean values were used for data analysis. Interrater and test-retest reliability was analyzed with ICC models 2,1 and 2,k (ie, two-way random, single measure and two-way random, average measure, respectively), and minimal detectable change (MDC) was determined at the 95% confidence level (MDC_{95}). The standard error of the mean (SEM) represents the standard deviation (SD) of the measurement error and was calculated using the operation

$$SD \times \sqrt{1 - ICC},$$

with SD representing the pooled variance.[28] The MDC_{95} is an extension of the

$$SEMMDC_{95} = SEM \times \text{z-score (95\% confidence interval)} \times \sqrt{2}$$

and provides a boundary for the minimum amount of variation that is not due to chance.[28] Assessment of ICC statistics was conducted using criteria described by Fleiss,[29] with reliability coefficients ≥0.75 rated as excellent, 0.40 to 0.74 rated as fair to good, and <0.40 rated as poor.

Results

Participants' demographic characteristics are reported in Table 1. Participants ranged in age from 18 to 43 years (mean ± SD = 24.5 ± 5.4), and 72.2% were men. Measures were reported for both men and women secondary to significant ($P < 0.05$) differences in height and weight, as well as in $SSWV_{RS}$ and TSA physical performance measures (Table 1). Participants had a high level of performance and limited variability in all measures except the 6MWT, in which the 5th percentile was 578.5 meters and the 95th percentile was 860.1 meters. Of 180 participants, 90% scored between 2.3 and 3.7 seconds on the TSA, 4.7 and 8.9 seconds on the STS5 test, and 3.9 and 7.8 seconds on the FSST. Participants also exhibited a limited range of performance on $SSWV_{OG}$ and $SSWV_{RS}$, with 90% of scores ranging from 1.27 to 1.80 m/s and 0.95 to 1.56 m/s, respectively.

Interrater reliability was assessed in 25 of 180 participants and was found to be nearly perfect for all measures (ICC = 0.97 to 0.99, Table 2). In addition, MDC_{95} values reflected limited error between raters, with 0.05 m/s each for $SSWV_{OG}$ and $SSWV_{RS}$, 0.19 seconds for the TSA, 0.27 seconds for the STS5 test, and 0.30 seconds for the FSST. Table 3 demonstrates differences between the mean and SD of the first and second sessions for the four performance measures as well as ICC and MDC_{95} values. All physical performance measures exhibited excellent test-retest reliability, with ICC ranging from 0.86 to 0.93. For each measure, the mean improved between sessions, with an MDC_{95} value of 81.25 meters for the 6MWT and MDC_{95} values of 0.37 seconds for TSA, 1.12 seconds for the STS5 test, and 1.41 seconds for the FSST.

Discussion

We sought to establish preliminary normative data and determine the reliability of and MDC values for physical performance measures that can be readily implemented in a clinical setting. The young, healthy cohort that we examined reflected demographics similar to those of injured service members described in earlier reports.[1-4] The selected measures were found to reliably allow for rapid assessment of a range of functional domains, including agility, mobility, balance, power, and exercise capacity.

In general, the mean and SD values

Table 1

Baseline Characteristics of Participants

Measure	Mean ± SD (range)	95% CI	Median (quartiles)	5th, 95th Percentiles
Age (yr)[a]				
Men	24.5 ± 5.5 (18–43)	23.3–25.4	24 (20–27)	18, 36
Women	24.5 ± 5.1 (18–40)	22.9–26.0	24 (22–27)	18, 38
Height (cm)[a]				
Men	175.9 ± 7.4 (152–193)	174.3–177.0	176 (171–182)	163, 188
Women	163.2 ± 6.3 (153–176)	161.3–165.2	162 (158–168)	153, 175.5
Weight (kg)[a]				
Men	80.1 ± 12.4 (52.3–111.8)	78.0–82.6	79.3 (70.9–87.7)	60.9, 102.6
Women	64.7 ± 8.2 (50–88.2)	63.5–68.4	64.1 (59–70.6)	52.2, 82.4
$SSWV_{OG}$ (m/s)[b]				
Men	1.51 ± 0.17 (1.0–1.93)	1.48–1.54	1.52 (1.41–1.64)	1.27, 1.76
Women	1.48 ± 0.16 (1.1–1.8)	1.43–1.53	1.45 (1.37–1.60)	1.27, 1.80
$SSWV_{RS}$ (m/s)[b]				
Men	1.26 ± 0.18 (0.66–1.70)	1.23–1.29	1.25 (1.17–1.39)	0.98, 1.56
Women	1.19 ± 0.14 (0.86–1.54)	1.15–1.24	1.19 (1.11–1.24)	0.95, 1.47
Timed Stair Ascent (s)[a]				
Men	2.82 ± 0.37 (2.27–4.5)	2.75–2.89	2.7 (2.6–3.0)	2.3, 3.5
Women	3.14 ± 0.36 (2.4–4.0)	3.03–3.25	3.2 (2.9–3.4)	2.5, 3.7
Sit-to-stand 5 Times Test (s)[a]				
Men	6.0 ± 1.0 (4.4–11.0)	5.8–6.1	5.8 (5.3–6.4)	4.8, 8.9
Women	6.24 ± 1.1 (4.35–9.14)	5.9–6.6	6.1 (5.4-6.8)	4.7, 8.8
Four-square Step Test (s)[a]				
Men	5.7 ± 1.0 (3.7–8.6)	5.5–5.9	5.6 (4.9–6.4)	3.9, 7.5
Women	6.0 ± 1.0 (4.0–8.1)	5.7–6.3	6.1 (5.4–6.7)	4.1, 7.8
Six-minute Walk Test (m)[a]				
Men	724.9 ± 84.1 (509–1,007)	710.3–739.5	722.2 (670.3–778.5)	581.2, 860.1
Women	707.4 ± 69.8 (555–856)	687.5–727.2	707.8 (655.3–760.5)	578.5, 820.5

CI = confidence interval, SD = standard deviation, $SSWV_{OG}$ = self-selected walking velocity over ground, $SSWV_{RS}$ = self-selected walking velocity over loose-rock surface
[a] N = 180 (130 men, 50 women)
[b] N = 160 (117 men, 43 women)

reflect a consistent and high-level baseline physical ability in the tested cohort compared with assessments of physical ability in older or impaired cohorts described by other authors (FSST,[21,30,31] STS5 test,[18] 6MWT,[12,32,33] SSWV[32,33]). Our participants demonstrated limited variability and a narrow range of performance for all measures except the 6MWT.

Overall, our study provides initial evidence regarding expected range of performance in young, healthy adults for six physical performance measures: $SSWV_{OG}$, $SSWV_{RS}$, STS5, FSST, TSA, and 6MWT. These measures demonstrated excellent interrater reliability and reliability between days (ICC >0.75) that was equal to or greater than that reported in earlier studies.[12,32,33] However, reliability of measures other than the 6MWT in young, active persons has not been published to date. Interrater MDC values were <0.3 second for simple timed measures and <0.1 m/s for walking velocity measures. Between-day MDC values ranged from 0.37 to 1.41 seconds for simple timed measures. Standardized instructions with clearly defined start and stop points were used to minimize the MDC values for timed measures. Mean values for between-day comparisons demonstrated small but significant improvements between

Table 2

Interrater Reliability of Physical Performance Measures[a] (n = 25)

Measure	Mean ± SD (rater 1)	Mean ± SD (rater 2)	ICC (95% CI range)	SEM	MDC_{95}
$SSWV_{OG}$ (m/s)	1.51 ± 0.19	1.50 ± 0.19	0.99 (0.99–0.99)	0.02	0.05
$SSWV_{RS}$ (m/s)	1.27 ± 0.17	1.28 ± 0.17	0.99 (0.99–0.99)	0.02	0.05
Timed Stair Ascent (s)	3.06 ± 0.40	3.03 ± 0.37	0.97 (0.92–0.98)	0.07	0.19
Sit-to-stand 5 Times Test (s)	6.51 ± 0.96	6.51 ± 0.90	0.99 (0.97–0.99)	0.10	0.27
Four-square Step Test (s)	6.57 ± 1.08	6.53 ± 1.08	0.99 (0.99–0.99)	0.11	0.30

CI = confidence interval, ICC = intraclass correlation coefficient, MDC_{95} = minimal detectable change 95% CI, SD = standard deviation, SEM = standard error of measurement, $SSWV_{OG}$ = self-selected walking velocity over level ground, $SSWV_{RO}$ = self-selected walking velocity over rock surface

[a] Reliability calculated with ICC (2,k)

Table 3

Test-retest Reliability of Physical Performance Measures (n = 20)

Measure	Mean ± SD (session 1)	Mean ± SD (session 2)	ICC (95% CI range)	MDC_{95}
Timed Stair Ascent (s)[a]	2.84 ± 0.43	2.77 ± 0.41	0.90 (0.76–0.96)	0.37
Sit-to-stand 5 Times Test (s)[a,b]	6.56 ± 1.41	6.17 ± 1.44	0.92 (0.75–0.97)	1.12
Four-square Step Test (s)[a,b]	5.12 ± 1.36	4.60 ± 1.03	0.86 (0.75–0.96)	1.41
6-minute Walk Test (m)[b,c]	733.45 ± 85.93	757.20 ± 110.65	0.93 (0.79–0.97)	81.25

CI = confidence interval, ICC = intraclass correlation coefficient, MDC_{95} = minimal detectable change 95% CI, SD = standard deviation

[a] Reliability calculated with ICC (2,k)

[b] Significant between-session changes ($P < 0.05$)

[c] Reliability calculated with ICC (2,1)

sessions for all measures except TSA. Potentially more important is the relative magnitude of the MDC values compared with the types of changes expected in patients during the recovery process. Although it is necessary to determine population-specific MDC values for injured service members, these early results suggest that the measures tested likely provide sufficient resolution to detect meaningful change in performance.

Several factors should be taken into consideration when using our results to interpret physical performance data. Data were collected in a cohort of physically fit persons who were expected to perform similarly over time, thereby increasing the likelihood of obtaining high ICC and low MDC values. The time between testing days was approximately 1 week; therefore, it remains unknown whether the same level of stability will be observed over much longer periods of time.

Summary

To our knowledge, our study is the first to provide preliminary normative data, reliability, and MDC values for each of six physical performance measures (ie, $SSWV_{OG}$, $SSWV_{RS}$, TSA, STS5, FSST, 6MWT) in a cohort of young, healthy, active-duty service members. The measures are easy to understand, require little equipment, and demonstrate excellent reliability. Additionally, they can be used to assess a range of functional domains that are thought to be impaired in injured service members; these domains are amenable to intervention. Further research is required on the predictive validity, responsiveness, and influence that surgical interventions and rehabilitation have on each of these six physical performance measures to enhance the current level of care for younger civilians and military service members who have sustained polytraumatic extremity injuries.

References

1. Cross JD, Ficke JR, Hsu JR, Masini BD, Wenke JC: Battlefield orthopaedic

injuries cause the majority of long-term disabilities. *J Am Acad Orthop Surg* 2011;19(suppl 1):S1-S7.

2. Owens BD, Kragh JF Jr, Macaitis J, Svoboda SJ, Wenke JC: Characterization of extremity wounds in Operation Iraqi Freedom and Operation Enduring Freedom. *J Orthop Trauma* 2007;21(4): 254-257.

3. Owens BD, Kragh JF Jr, Wenke JC, Macaitis J, Wade CE, Holcomb JB: Combat wounds in operation Iraqi Freedom and operation Enduring Freedom. *J Trauma* 2008;64(2): 295-299.

4. Masini BD, Waterman SM, Wenke JC, Owens BD, Hsu JR, Ficke JR: Resource utilization and disability outcome assessment of combat casualties from Operation Iraqi Freedom and Operation Enduring Freedom. *J Orthop Trauma* 2009;23(4):261-266.

5. Owens JG: Physical therapy of the patient with foot and ankle injuries sustained in combat. *Foot Ankle Clin* 2010;15(1):175-186.

6. Patzkowski JC, Blanck RV, Owens JG, Wilken JM, Blair JA, Hsu JR: Can an ankle-foot orthosis change hearts and minds? *J Surg Orthop Adv* 2011;20(1):8-18.

7. Hammarlund CS, Carlström M, Melchior R, Persson BM: Prevalence of back pain, its effect on functional ability and health-related quality of life in lower limb amputees secondary to trauma or tumour: A comparison across three levels of amputation. *Prosthet Orthot Int* 2011;35(1):97-105.

8. Pape HC, Probst C, Lohse R, et al: Predictors of late clinical outcome following orthopedic injuries after multiple trauma. *J Trauma* 2010;69(5): 1243-1251.

9. Johanson NA, Liang MH, Daltroy L, Rudicel S, Richmond J: American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments: Reliability, validity, and sensitivity to change. *J Bone Joint Surg Am* 2004; 86(5):902-909.

10. Ly TV, Travison TG, Castillo RC, Bosse MJ, MacKenzie EJ, LEAP Study Group: Ability of lower-extremity injury severity scores to predict functional outcome after limb salvage. *J Bone Joint Surg Am* 2008;90(8):1738-1743.

11. Cecchi F, Molino-Lova R, Di Iorio A, et al: Measures of physical performance capture the excess disability associated with hip pain or knee pain in older persons. *J Gerontol A Biol Sci Med Sci* 2009;64(12):1316-1324.

12. Resnik L, Borgia M: Reliability of outcome measures for people with lower-limb amputations: Distinguishing true change from statistical error. *Phys Ther* 2011;91(4):555-565.

13. Archer KR, Castillo RC, Mackenzie EJ, Bosse MJ: Gait symmetry and walking speed analysis following lower-extremity trauma. *Phys Ther* 2006;86(12):1630-1640.

14. Castillo RC, MacKenzie EJ, Archer KR, Bosse MJ, Webb LX, LEAP Study Group: Evidence of beneficial effect of physical therapy after lower-extremity trauma. *Arch Phys Med Rehabil* 2008; 89(10):1873-1879.

15. Getchell B: *Physical Fitness: A Way of Life*, ed 2. New York, John Wiley and Sons, 1979.

16. VanSwearingen JM, Brach JS: Making geriatric assessment work: Selecting useful measures. *Phys Ther* 2001;81(6): 1233-1252.

17. Hughes C, Osman C, Woods AK: Relationship among performance on stair ambulation, functional reach, and timed up and go in older adults. *Issues on Aging* 1998;21:18-22.

18. Whitney SL, Wrisley DM, Marchetti GF, Gee MA, Redfern MS, Furman JM: Clinical measurement of sit-to-stand performance in people with balance disorders: Validity of data for the Five-Times-Sit-to-Stand Test. *Phys Ther* 2005; 85(10):1034-1045.

19. Wolinsky FD, Miller DK, Andresen EM, Malmstrom TK, Miller JP: Reproducibility of physical performance and physiologic assessments. *J Aging Health* 2005; 17(2):111-124.

20. Dite W, Temple VA: A clinical test of stepping and change of direction to identify multiple falling older adults. *Arch Phys Med Rehabil* 2002;83(11): 1566-1571.

21. Whitney SL, Marchetti GF, Morris LO, Sparto PJ: The reliability and validity of the four square step test for people with balance deficits secondary to a vestibular disorder. *Arch Phys Med Rehabil* 2007; 88(1):99-104.

22. Bunc V: A simple method for estimating aerobic fitness. *Ergonomics* 1994;37(1): 159-165.

23. Cooper KH: A means of assessing maximal oxygen intake: Correlation between field and treadmill testing. *JAMA* 1968;203(3):201-204.

24. Cooper KH, Zechner A: Physical fitness in United States and Austrian military personnel: A comparative study. *JAMA* 1971;215(6):931-934.

25. Gibbons WJ, Fruchter N, Sloan S, Levy RD: Reference values for a multiple repetition 6-minute walk test in healthy adults older than 20 years. *J Cardiopulm Rehabil* 2001;21(2):87-93.

26. Butland RJ, Pang J, Gross ER, Woodcock AA, Geddes DM: Two-, six-, and 12-minute walking tests in respiratory disease. *Br Med J (Clin Res Ed)* 1982;284(6329):1607-1608.

27. Drinkard B, McDuffie J, McCann S, Uwaifo GI, Nicholson J, Yanovski JA: Relationships between walk/run performance and cardiorespiratory fitness in adolescents who are overweight. *Phys Ther* 2001;81(12): 1889-1896.

28. Haley SM, Fragala-Pinkham MA: Interpreting change scores of tests and measures used in physical therapy. *Phys Ther* 2006;86(5):735-743.

29. Fleiss JL: *The Design and Analysis of Clinical Experiments*. New York, NY, John Wiley and Sons, 1986.

30. Blennerhassett JM, Jayalath VM: The four square step test is a feasible and valid clinical test of dynamic standing balance for use in ambulant people poststroke. *Arch Phys Med Rehabil* 2008;89(11):2156-2161.

31. McCulloch KL, Buxton E, Hackney J, Lowers S: Balance, attention, and dual-task performance during walking after brain injury: Associations with falls history. *J Head Trauma Rehabil* 2010; 25(3):155-163.

32. Ries JD, Echternach JL, Nof L, Gagnon Blodgett M: Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Phys Ther* 2009; 89(6):569-579.

33. Steffen T, Seney M: Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys Ther* 2008;88(6):733-746.