Test of Intrarater and Interrater Reliability for the Star Excursion Balance Test

In Hyouk Hyong, PhD, PT[1], Jae Hyun Kim, PhD, PT[1]*

[1] Department of Physical Therapy, Shinsung University: Daehak-ro, Jeongmi-myeon, Dangjin-si, Chungcheongnam-do, Republic of Korea

Abstract. [Purpose] The aim of this study was to examine the intrarater and interrater reliability of the Star Excursion Balance Test (SEBT), thereby increasing understanding of its efficient utilization. [Subjects and Methods] There were 67 subjects (49 female; 18 male). For the SEBT, eight lines were made using tape at 45-degree angles from the center of a circle. The experiment was conducted in the following order: the anterior, anterior-medial, medial, posterior-medial, posterior, posterior-lateral, lateral, and anterior-lateral directions. Intraclass correlation coefficients (ICC) (3,1) were used to evaluate the intrarater and interrater reliability (2,1) for each reach distance, while the standard error of measurement (SEM) and smallest detectable distance (SDD) were employed to assess absolute reliability. [Results] For intraratar reliability, the ICC values for all directions ranged from 0.88 to 0.96, SEM values ranged from 2.41 to 3.30, and SDD values ranged from 6.68 to 9.15. For interrater reliability, the ICC values for all directions ranged from 0.83 to 0.93, SEM values ranged from 3.19 to 4.26, and SDD values ranged from 8.85 to 11.82 [Conclusion] The SEBT is a highly reliable tool for measuring dynamic balance. Measurements for intrarater reliability are more reliable than measurements for interrater reliability. When measurement for eight directions was difficult, the SEBT was used. While the anterior, posteromedial, and posterolateral directions employed in the Y Balance Test Kit[TM] can be utilized, this study recommends using the reverse Y Balance Test Kit[TM] method with the posterior direction, not the anterior direction.
Key words: Star excursion balance test (SEBT), Dynamic balance, Reliability

INTRODUCTION

The Star Excursion Balance Test (SEBT) is a tool to assess the dynamic balance of healthy people and athletes[1–3]. This evaluation tool uses closed-kinetic chain exercises, specifically single-leg squat exercises which require appropriate range of motion in the hip joints, knees and ankle joints; and muscle strength; and proprioceptive and neuromuscular adjustments[4] dynamic balance is measured from eight directions which are highly related[5]: the anterior, anterior-medial, medial, posterior-medial, posterior, posterior-lateral, lateral, and anterior-lateral directions. The most widely used process for the SEBT was developed by Hertel et al[1]. Each exercise is conducted six times, and measurements are taken three times for each of the eight directions. Hertel et al.[1] and Plisky et al.[6] reported high intraclass correlation coefficients (ICC) ranging from 0.78 to 0.96.

Therefore, the SEBT offers high reliability in evaluating dynamic balance and has the advantage of the capability to measure the dynamic balance of both healthy people and athletes. However, practicing the exercise for measurement and measuring dynamic balance is time intensive[1], and the intrarater and interrater reliability differ. In addition, it is claimed that the dynamic balances of each of the eight directions are highly related[5], but no study has compared the dynamic balances of these directions. Therefore, the aim of this study was to analyze and compare intrarater and interrater reliability of the SEBT in order to increase understanding of its most efficient utilization.

SUBJECTS AND METHODS

This study involved 67 subjects, who understood its purpose and voluntarily consented to participate. There were 49 female subjects (age, weight, and height: 20.5±0.6, 54.2±7.5 kg, and 161.5±5.5 cm, respectively) and 18 male subjects (age, weight, and height: 21.4±1.6, 69.6±10.7 kg, 176.7±5.4 cm, respectively). The subjects had no musculoskeletal diseases or neurological problems that negatively influenced dynamic balance. All participants signed an informed consent form, and the study was approved by the institutional review board of the Catholic University of Busan (document number: CUPIRB-2013-042).

For the SEBT, eight lines were made using tape at 45 degrees from the center of a circle[7]. The experiment was conducted in the following order: the anterior, anterior-medial, medial, posterior-medial, posterior, posterior-lateral, lateral, and anterior-lateral directions. The subjects posi-

*Corresponding author. Jae Hyun Kim (E-mail: anatomy2@ shinsung.ac.kr)

tioned the bilateral arms on the iliac crest of the waist[7] and the heels at the intersection of the eight lines in the center of the circle. The subject stretched one leg to its maximal extent and lightly touched one line with the end of the big toe while maintaining his/her balance. After touching the line, the subject returned to the erect position with both legs in the center. Using the same method, measurements were taken for the other seven directions. The rater measured the distance from the center of the circle to the point where the subjects touched each line. Measurements were taken three times after the subjects practiced the exercise six times, in accordance with the method of Hertel et al[1]. Rater A conducted measurements twice to measure intrarater reliability, while raters. B and C conducted measurement once to measure interrater reliability.

ICCs were used to evaluate intrarater (3,1) and interrater reliability (2,1) for each reach distance. ICCs were evaluated according to the following standard: poor≤0.40, fair ≈ 0.40–0.70, good ≈ 0.70–0.90, excellent≥0.90[8]. The standard error of measurement (SEM) (standard deviation (SD)× $\sqrt{(1-ICC)}$)[9] and smallest detectable difference (SDD) (1.96 × $\sqrt{2}$ ×SEM)[10] were employed to assess absolute reliability. Data collected from this study were analyzed using SPSS 17.0. The significance level was set at α = 0.05.

RESULTS

For intrarater reliability, the ICC values for all directions ranged from 0.88 to 0.96, SEM values ranged from 2.41 to 3.30, and SDD values ranged from 6.68 to 9.15 (Table 1). Table 1 shows the mean, 95% confidence interval (CI), SEM, and SDD values. For interrater reliability, the ICC values for all directions ranged from 0.83 to 0.93, SEM values ranged from 3.19 to 4.26, and SDD values ranged from 8.85 to 11.82 (Table 2).

DISCUSSION

Plisky et al.[6] used the same method as the present study—practicing the exercise six times and measuring dynamic balance three times—and reported that the test-retest reliability ranged from 0.89 to 0.93. The present study also exhibits high reliability, ranging from 0.88 to 0.96, using the same method. In a recent study, however, the subjects practiced the exercise four times, and their dynamic balance was measured three times, resulting ICCs ranging from 0.84 to 0.92[11]. Therefore, if there is no difference in reliability, measuring dynamic balance after practicing the exercise four times is recommended, instead of measuring dynamic balance after practicing the exercise six times[11].

The SEM, a measure of absolute reliability, provides estimates for the error size of each measured score and is an indicator of the reliability of indexes[12]. The SDD[10] is another measure of absolute reliability and is used with the small reference difference (SRD)[12]. The SDD is defined as a reliability level of 95% of the SEM between measured scores. It measures the sensitivity of changes in measured values and, together with the SEM, is a change index reflecting the reliability of indicators[12]. Lower SEM and SDD values indicate higher reliability of the accuracy and precision of the measured values. When the SEM value is less

Table 1. Intrarater reliability (one later) for all reach directions

Direction	Mean (%)	ICC (3,1)	95%CI	SEM	SDD
Anterior	99.53	0.88	0.81–0.93	3.24	8.99
Anterior-Lateral	102.77	0.91	0.85–0.94	2.89	8
Lateral	100.13	0.96	0.93–0.97	2.41	6.68
Posterior-Lateral	94.1	0.93	0.88–0.96	3.3	9.15
Posterior	83.86	0.95	0.92–0.97	2.5	6.93
Posterior-Medial	77.7	0.94	0.91–0.97	2.41	6.68
Medial	72.2	0.96	0.94–0.98	2.49	6.9
Anterior-Medial	84.9	0.91	0.85–0.94	2.58	7.16

All values except the ICC are normalized to the excursion distance (excursion distance/leg length × 100).

Table 2. Interrater reliability for all reach directions

Direction	Mean (%)	ICC (2,1)	95%CI	SEM	SDD
Anterior	101.85	0.83	0.75–0.89	3.68	10.2
Anterior-Lateral	106.01	0.88	0.82–0.92	3.59	9.96
Lateral	103.22	0.93	0.89–0.95	3.19	8.85
Posterior-Lateral	97.46	0.88	0.82–0.92	4.26	11.82
Posterior	88.07	0.88	0.81–0.92	4.14	11.49
Posterior-Medial	81.44	0.9	0.85–0.94	3.76	10.41
Medial	76.05	0.93	0.89–0.95	3.7	10.24
Anterior-Medial	86.19	0.84	0.77–0.90	3.87	10.73

All values except the ICC are normalized to the excursion distance (excursion distance/leg length × 100).

than 10% of the average measured value or the highest measured value, the measurement error is small, and therefore, the measurement is reliable[12].

In the present study, the intrarater SEM, SDD, and ICC values were 2.41–3.30, 7.16–8.99, and 0.88–0.96, respectively, and the interrater SEM, SDD, and ICC values were 3.19–4.26, 8.85–11.82, and 0.88–0.93, respectively. Comparing the intrarater and interrater SEM and SDD values, the intrarater ICCs ranging from 0.88 to 0.96 were high. Therefore, using one rater to measure dynamic balance during the SEBT, is more than using multiple raters.

Plisky et al.[13] used the Y Balance Test Kit[TM] (Functional Movement Systems, Danvile, VA, USA) while conducting SEBTs to measure dynamic balance from three directions in order to resolve the difficulty of measuring dynamic balance from eight directions. The Y Balance Test Kit[TM] was designed to measure dynamic balance from the anterior, posteromedial, and posterolateral directions. Plisky et al.[13] reported that the ranged of the intrarater ICC and SEM were 0.85–0.89 and 2.01–3.11, respectively, and that those of the interrater ICC and SEM were 0.99–1.0 and 0.69–0.73, respectively. These results are similar to those of the present study in which SEBT was measured from eight directions. Therefore, the Y Balance Test Kit[TM] may be utilized when measuring SEBT from eight directions is difficult. In a study by Plisky et al.[13], the intrarater absolute reliability index SEM (2.01–3.11) was lower than the interrater SEM (0.69–0.73). As in the present study, intrarater measures are more reliable than interrater measures when the Y Balance Test Kit[TM] is used. However, when the Y Balance Test Kit[TM] was used in the present study, the ICC, SEM, and SDD values from the posterior directions were more reliable than those from the anterior directions. Therefore, the present study proposes a reverse Y Balance Test Kit[TM] method using the posterior directions, instead of the anterior directions. Herrington et al.[14] applied the SEBT to normal subjects and patients with anterior cruciate ligament injury and reported obvious differences in dynamic balance. It is concluded that the SEBT is a highly reliable tool to measure the dynamic balance of normal people and athletes; however, intrarater measurements are more reliable than interrater measurements. Research on the reliability of the SEBT when measuring dynamic balance from three directions using the Y Balance Test Kit[TM] is needed.

REFERENCES

1) Hertel J, Miller SJ, Denegar CR: Intratester and intertester reliability during the star excursion balance tests. J Sport Rehabil, 2000, 9: 104–116.
2) Kinzey SJ, Armstrong CW: The reliability of the star-excursion test in assessing dynamic balance. J Orthop Sports Phys Ther, 1998, 27: 356–360. [Medline] [CrossRef]
3) Gribble PA, Hertel J: Considerations for normalizing measures of the star excursion balance test. Meas Phys Educ Exerc Sci, 2003, 7: 89–100. [CrossRef]
4) Olmsted LC, Carcia CR, Hertel J, et al.: Efficacy of the star excursion balance tests in detecting reach deficits in subjects with chronic ankle instability. J Athl Train, 2002, 37: 501–506. [Medline]
5) Hertel J, Braham RA, Hale SA, et al.: Simplifying the star excursion balance test: analyses of subjects with and without chronic ankle instability. J Orthop Sports Phys Ther, 2006, 36: 131–137. [Medline] [CrossRef]
6) Plisky PJ, Rauh MJ, Kaminski TW, et al.: Star Excursion Balance Test as a predictor of lower extremity injury in high school basketball players. J Orthop Sports Phys Ther, 2006, 36: 911–919. [Medline] [CrossRef]
7) Robinson RH, Gribble PA: Support for a reduction in the number of trials needed for the star excursion balance test. Arch Phys Med Rehabil, 2008, 89: 364–370. [Medline] [CrossRef]
8) Coppieters M, Stappaerts K, Janssens K, et al.: Reliability of detecting 'onset of pain' and 'submaximal pain' during neural provocation testing of the upper quadrant. Physiother Res Int, 2002, 7: 146–156. [Medline] [CrossRef]
9) Thomas JR, Nelson JK, Silverman SJ: Research methods in physical activity, 3rd ed. Champaign: Human Kinetics, 2005.
10) Kropmans TJ, Dijkstra PU, Stegenga B, et al.: Smallest detectable difference in outcome variables related to painful restriction of the temporomandibular joint. J Dent Res, 1999, 78: 784–789. [Medline] [CrossRef]
11) Munro AG, Herrington LC: Between-session reliability of the star excursion balance test. Phys Ther Sport, 2010, 11: 128–132. [Medline] [CrossRef]
12) Liaw LJ, Hsieh CL, Lo SK, et al.: The relative and absolute reliability of two balance performance measures in chronic stroke patients. Disabil Rehabil, 2008, 30: 656–661. [Medline] [CrossRef]
13) Plisky PJ, Gorman PP, Butler RJ, et al.: The reliability of an instrumented device for measuring components of the star excursion balance test. N Am J Sports Phys Ther, 2009, 4: 92–99. [Medline]
14) Herrington L, Hatcher J, Hatcher A, et al.: A comparison of Star Excursion Balance Test reach distances between ACL deficient patients and asymptomatic controls. Knee, 2009, 16: 149–152. [Medline] [CrossRef]