S. O'Hoski, MScPT, Department of Physical Therapy, University of Toronto, Toronto, Ontario, Canada, and Department of Respiratory Medicine, West Park Healthcare Centre, Toronto, Ontario, Canada.

B. Winship, MScPT, Department of Physical Therapy, University of Toronto.

L. Herridge, MScPT, Department of Physical Therapy, University of Toronto, and Southlake Regional Health Centre, Newmarket, Ontario, Canada.

T. Agha, MScPT, Department of Physical Therapy, University of Toronto.

D. Brooks, PhD, Department of Physical Therapy, University of Toronto; West Park Healthcare Centre; and Toronto Rehabilitation Institute–University Health Network, Toronto, Ontario, Canada.

M.K. Beauchamp, PhD, Department of Physical Therapy, University of Toronto; West Park Healthcare Centre; and Harvard Medical School, Cambridge, Massachusetts.

K.M. Sibley, PhD, Department of Physical Therapy, University of Toronto, and Toronto Rehabilitation Institute–University Health Network, 550 University Ave, Room 11-173, Toronto, Ontario, Canada M5G 2A2. Address all correspondence to Dr Sibley at: k.sibley@utoronto.ca.

Post a Rapid Response to this article at:
ptjournal.apta.org

Increasing the Clinical Utility of the BESTest, Mini-BESTest, and Brief-BESTest: Normative Values in Canadian Adults Who Are Healthy and Aged 50 Years or Older

Sachi O'Hoski, Bonnie Winship, Lauren Herridge, Taimoor Agha, Dina Brooks, Marla K. Beauchamp, Kathryn M. Sibley

Background. Balance is a composite ability requiring the integration of multiple systems. The Balance Evaluation Systems Test (BESTest) and 2 abbreviated versions (the Mini-BESTest and the Brief-BESTest) are balance assessment tools that target these systems. To date, no normative data exist for any version of the BESTest.

Objective. The purpose of this study was to determine the age-related normative scores on the BESTest, Mini-BESTest, and Brief-BESTest for Canadians who are healthy and 50 to 89 years of age.

Design. A cross-sectional study design was used.

Methods. Seventy-nine adults who were healthy and aged 50 to 89 years (mean age=68.9 years; 50.6% women) participated. Normative scores were reported by age decade.

Results. Mean BESTest scores were 95.7 (95% confidence interval [CI]=94.4–97.1) for adults who were aged 50 to 59 years, 91.4 (95% CI=89.8–93.0) for those who were aged 60 to 69 years, 85.4 (95% CI=82.5–88.2) for those who were aged 70 to 79 years, and 79.4 (95% CI=74.3–84.5) for those who were aged 80 to 89 years. Similar results were obtained for the Mini-BESTest and the Brief-BESTest, and all 3 tests showed statistically significant differences in scores among the age cohorts.

Limitations. Because only adults who were 50 to 89 years of age were tested, there are still no normative data for people outside this age range. Also, the scores presented may not be generalizable to all countries.

Conclusions. These normative data enhance the clinical utility of the BESTest, Mini-BESTest, and Brief-BESTest by providing clinicians with reference points to guide treatment.

Approximately one third of people who live in the community and are more than 65 years of age fall each year.[1] Falls are associated with increased morbidity and mortality as well as high health care costs.[2] Many risk factors for falls have been identified, and an important modifiable risk factor is a deficit in balance.[3-6] Defined as the ability to maintain the body's center of mass over its base of support, balance is not a stand-alone skill; it is a composite ability involving the rapid, automatic anticipatory and reactive integration of information from several systems.[7,8] Many of the components that contribute to balance, such as strength and sensation, are impaired in elderly people.[3,4,6,9] Therefore, appropriate clinical assessment tools are necessary to screen for balance impairments.

Commonly used functional balance tests, including the Berg Balance Scale (BBS)[10] and the Timed "Up & Go" Test (TUG),[11] have been designed to identify balance problems and predict fall risk.[10,12-14] However, few balance tests have been developed to identify the underlying systems responsible for balance deficits. An understanding of the systems underlying deficits in postural control is critical for diagnosing specific impairments and developing individualized treatment plans.[8] The Balance Evaluation Systems Test (BESTest) is a recently developed standardized functional balance tool that is aimed at identifying the components contributing to dysfunctional balance; it targets 6 postural control subsystems (Tab. 1).[15] The BESTest has been shown to have high interrater reliability, high test-retest reliability, and very good validity in people with Parkinson disease (PD).[16] Performance on the BESTest has been shown to discriminate between people who have PD and experience falls and people who have PD and do

not experience falls[16,17] and between the impairments associated with several clinical diagnoses, including PD and vestibular dysfunction.[15] The BESTest also has been used in people with cerebral palsy, peripheral neuropathy, total hip replacements, fibromyalgia, and chronic obstructive pulmonary disease.[15,18-20]

Despite its validation and published findings, the BESTest is not often used in clinical practice,[21] perhaps because the administration time—which has been reported to range from 20 to 60 minutes[15,22]—may not be feasible in all clinical settings. Accordingly, an abbreviated version of the BESTest (Mini-BESTest) was developed as a brief test of dynamic balance that can be administered in less than half the time of the original BESTest.[23] The Mini-BESTest consists of 14 of the 36 items from the original BESTest, but the items are scored differently—on a 3-point rather than a 4-point scale.[23,24] Scores on the Mini-BESTest have been shown to correlate well with total BESTest scores,[17] balance confidence scores,[24] and BBS scores in people with PD.[25,26] It also has been shown to have high interrater and test-retest reliability.[17] The Mini-BESTest has been used to test balance in people with stroke, multiple sclerosis, vestibular disorders, and traumatic brain injury[23] and, like the BESTest, has been shown to discriminate between people who have PD and experience falls and people who have PD and do not experience falls.[17]

Although the Mini-BESTest fulfills the need for a shorter version of the BESTest, it only provides a total score for dynamic balance and does not identify the underlying impairment. Another abbreviated version, the Brief-BESTest,[22] was developed to maintain the theoretical basis of the original test. Padgett et al[22] examined the internal consistency of

each item of the BESTest and used item-total correlations to identify the most representative item of each section. The resulting Brief-BESTest consists of 1 item from each section of the original BESTest, with 2 items (single-leg stance and functional forward reach) being scored bilaterally. In preliminary testing, the Brief-BESTest was shown to have interrater reliability comparable to those of the BESTest and the Mini-BESTest and accuracy superior to those of the other tests for identifying people with and without a neurological diagnosis and people who fall versus people who do not fall.[22]

The BESTest has been used with people who were healthy serving as control participants in a few studies,[15,19,20] but the small sample sizes in those studies (ranging from 3 to 32 participants), as well as the failure to report scores based on age, limited the generalizability and interpretation of scores achieved by the participants. To date, no normative BESTest, Mini-BESTest, or Brief-BESTest data have been published. The ability to compare patients' scores on the BESTest, Mini-BESTest, and Brief-BESTest with a range of scores expected for people who are healthy and matched for age will be meaningful for clinicians and patients because it will provide a relative indication of balance performance and help guide treatment. Thus, the primary objective of this study was to determine the age-related normative scores on the BESTest, Mini-BESTest, and Brief-BESTest for Canadians who are healthy and 50 to 89 years of age. We hypothesized that balance scores would differ significantly among the age groups.

Method
Written informed consent was obtained from participants, and a copy of the consent form was pro-

Table 1.
Description of Sections and Items of the Balance Evaluation Systems Test[15]

Section	Description	Items (Number)
Biomechanical constraints	Items in this section evaluate constraints on standing balance, including posture, range of motion, and strength	Quality of base of support, postural alignment, ankle strength and range of motion, hip strength, ability to sit on floor and stand up (1–5)
Stability limits/ verticality	Items in this section evaluate how far the body can move over its base of support and the internal perception of gravitational vertical	Lateral lean in sitting position, verticality, and forward and lateral reach (6–8)
Anticipatory postural adjustments	Items in this section evaluate active movement of the center of mass in anticipation of positional changes	Sit-to-stand transfer, rising to toes, single-leg stance, stair tap, and standing arm raise (9–13)
Postural responses	Items in this section evaluate in-place and compensatory stepping responses to external perturbations	In-place resistance to perturbation and forward, backward, and lateral responses to "push and release" (14–18)
Sensory orientation	Items in this section evaluate increases in postural sway under different sensory conditions	Standing on flat ground and foam with eyes open or closed and standing on ramp with eyes closed (19 and 20)
Stability in gait	Items in this section evaluate stability while walking under conditions in which balance is challenged	Usual gait speed, change in speed, walking with head turns, quick turning and stopping, stepping over obstacle, Timed "Up & Go" Test (TUG), TUG with cognitive dual tasks (21–27)

vided to each participant. A cross-sectional study design was used.

Participants

Adults (50–89 years of age) who were healthy and living in the community were recruited through local advertisements in community centers, hospitals, and universities. Consistent with previous reports of normative scores,[27,28] we targeted a sample size of 80 participants (10 men and 10 women in each decade between 50 and 89 years of age). Assignment to an age cohort was determined by a participant's chronological age at the time of testing.

Interested people were screened over the telephone to determine eligibility for the study. People were included if they met the following 6 criteria: 50 to 89 years of age, living independently in the community, able to speak and read English, able to follow 3-step commands, able to provide written informed consent, and able to ambulate 6 m independently without a gait aid. People were excluded if they reported a history of dizziness or fainting; a past or current history of a cardiorespira-

tory, neurological, or musculoskeletal impairment that affected their balance; or current use of any medications that can cause dizziness or impair balance (eg, psychotropic medications).

Procedure

Each data collection session was completed within a 60-minute period in a quiet laboratory setting at the University of Toronto between January and July 2012. Participants were instructed to wear comfortable, flat shoes. Demographic data, including sex, age, height, and weight, were collected before the administration of the BESTest.

Four members of the research team who were master of physical therapy students (S.O., B.W., L.H., and T.A.) were trained to administer and score the BESTest by first observing the BESTest training DVD[29] and then receiving training from a registered physical therapist (M.K.B.) with extensive experience administering the test. To reduce errors in inter-rater reliability, all 4 testers scored the first 4 participants. The scores for each item of the BESTest were

then compared to ensure consistency of ratings. When discrepancies in scoring were evident, the testers discussed their rationales for the scores chosen and reached a unified conclusion on how to score future attempts of problematic tasks.

Two of the 4 testers were present for each testing session. For each item on the BESTest, 1 tester read the standardized instructions[29] to the participant while the second tester completed a demonstration of the task. The participant then attempted the task with close supervision provided by the second tester to ensure participant safety. If the participant's attempt indicated an obvious misunderstanding of the instructions, another demonstration was given, and the participant was allowed a second attempt at the task. Each task was scored immediately after completion, and participants were provided with a verbal summary of their BESTest results at the end of the session.

Scoring of the Mini-BESTest and Brief-BESTest occurred after the completion of all testing sessions and

was based on the performance of the BESTest tasks; participants did not complete Mini-BESTest or Brief-BESTest tasks separately. All scores were calculated by 2 of the testers and verified by the other 2 testers at the time of data entry.

Outcome Measures

BESTest. The BESTest[15] consists of 36 items grouped into 6 categories (Tab. 1). Each task is scored on an ordinal scale from 0 to 3, as judged by time or performance criteria. The total BESTest score is a sum of all of the individual items, resulting in a maximum score of 108 points. Scores are converted to percentages; higher scores indicate better balance performance. Materials needed to administer the BESTest, including a 10-degree incline ramp, a 60- × 60-cm block of approximately 10-cm (4-in), medium-density Tempur foam (Tempur-Pedic North America Inc, Medical Division, Lexington, Kentucky), and the BESTest training DVD, were purchased from the BESTest website.[29] All other materials were used in accordance with BESTest written standards. The stair height was 17 cm, and the obstacle (2 stacked shoe boxes) height was 25 cm. A 2.25-kg (5-lb) plate was used for the standing arm raise item in the anticipatory postural adjustments section.

Mini-BESTest. The Mini-BESTest[23] includes 14 items from 4 of the 6 sections of the BESTest. It includes 3 tasks for anticipatory postural adjustments, 3 tasks for postural responses, 3 tasks for sensory orientation, and 5 tasks for stability in gait. It does not include any items from the biomechanical constraints or stability limits/verticality sections because items from these sections were not deemed to measure dynamic balance. Items are scored from 0 to 2, and the scores are summed to obtain a total score out of a possible maximum score of 28

points.[24] Higher scores indicate better balance performance.

Brief-BESTest. The Brief-BESTest[22] was created from 6 items of the BESTest, 1 from each section, with 2 items (single-leg stance and functional forward reach) being scored bilaterally, resulting in an 8-item test. Items are scored from 0 to 3, and the scores are summed to obtain a total score out of a possible maximum score of 24 points. Higher scores indicate better balance performance. Because this test was created by compiling the most statistically representative item from each section of the BESTest, each item has its own section score.

Data Analysis

Descriptive statistics (mean, SD, and 95% confidence interval) were calculated for age, height, weight, body mass index, BESTest (total score and section scores), Mini-BESTest, and Brief-BESTest (total score and section scores). Box plots were used to show the median, minimum, and maximum values and the 25th to 75th percentiles for the BESTest, Mini-BESTest, and Brief-BESTest total scores for each age cohort. Both graphic and statistical methods (Shapiro-Wilk test) were used to determine normality. Because the data were not normally distributed, Kruskal-Wallis analyses were used to determine whether balance scores differed significantly across age groups within each of the balance tests. All statistical analyses were conducted with SPSS software (version 19.0 for Windows, SPSS Inc, Chicago, Illinois).

Results

The targeted sample size (n=10) was achieved in all age and sex cohorts except for men who were 80 to 89 years old (n=9), resulting in a total sample size of 79 participants. The descriptive characteristics of the participants are shown in Table 2. Mini-

BESTest scores are missing for 1 man in the 50- to 59-year-old cohort and for 2 men in the 60- to 69-year-old cohort because of differences in scoring of the BESTest and Mini-BESTest (Tab. 3). A score of 2 points on item 20 in the sensory orientation section of the BESTest could correspond to a score of either 1 or 2 on item 9 of the Mini-BESTest; therefore, these 3 tests were not included in the analyses.

Table 3 shows the normative scores on the BESTest (total score and section scores), Mini-BESTest, and Brief-BESTest (total score and section scores) for each age cohort. Figures 1, 2, and 3 show the box plots for each test's total score. Mean total scores decreased with age for all 3 tests. The Kruskal-Wallis analyses showed significant differences across age groups on the BESTest (χ^2=47.990, df=3, P<.001), Mini-BESTest (χ^2=41.662, df=3, P< .001), and Brief-BESTest (χ^2=37.608, df=3, P<.001) as well as all subscores on the BESTest and Brief-BESTest (Tab. 3).

Discussion

The present study provides BESTest, Mini-BESTest, and Brief-BESTest scores for a representative cohort of older adults who were healthy and living in the community and fills a gap in the literature because no normative data previously existed for these measures. The results of the present study can be used by clinicians to guide the interpretation of balance scores on the BESTest, Mini-BESTest, and Brief-BESTest. Furthermore, our data support our hypothesis that BESTest, Mini-BESTest, and Brief-BESTest scores would decrease with age.

Balance scores showed a significant decline with age, as expected from previous work.[28,30] Isles et al[30] found that balance performance, as measured with the TUG, the Step Test,

Table 2.
Participant Characteristics

Characteristic	Group	Age Cohort (y)											
		50–59			60–69			70–79			80–89		
		n	X̄ (SD)	Range	n	X̄ (SD)	Range	n	X̄ (SD)	Range	n	X̄ (SD)	Range
Age (y)	Total	20	55.5 (3.1)	50–59	20	63.5 (2.9)	60–69	20	74.0 (0.5)	70–79	19	82.5 (0.5)	80–87
	Men	10	54.6 (3.0)	50–58	10	64.2 (3.1)	60–69	10	73.4 (2.0)	70–77	9	82.3 (2.6)	80–87
	Women	10	56.4 (3.1)	50–59	10	62.8 (2.8)	60–67	10	74.5 (2.3)	71–79	10	82.6 (2.2)	80–87
Height (m)	Total	20	1.7 (0.1)	1.5–1.9	20	1.7 (0.0)	1.5–1.8	20	1.7 (0.0)	1.6–1.9	19	1.6 (0.0)	1.5–1.8
	Men	10	1.8 (0.1)	1.7–1.9	10	1.8 (0.0)	1.7–1.8	10	1.7 (0.1)	1.6–1.9	9	1.7 (0.0)	1.7–1.8
	Women	10	1.6 (0.1)	1.5–1.7	10	1.7 (0.1)	1.5–1.8	10	1.6 (0.0)	1.6–1.7	10	1.5 (0.1)	1.5–1.7
Weight (kg)	Total	20	72.5 (19.1)	52.3–116.0	20	75.8 (3.0)	52.2–100.2	20	70.3 (2.0)	50.5–83.0	19	67.1 (2.6)	50.5–83.0
	Men	10	81.6 (15.4)	59.1–105.2	10	81.4 (9.7)	64.0–100.2	10	77.6 (3.6)	72.0–83.0	9	75.1 (8.0)	61.0–83.0
	Women	10	63.5 (18.8)	52.3–116.0	10	70.2 (14.7)	52.2–97.5	10	63.1 (6.3)	50.5–72.6	10	59.9 (8.9)	50.5–79.5
Body mass index (kg/m²)	Total	20	25.2 (4.9)	20.1–31.8	20	26.1 (1.1)	19.1–35.8	20	25.0 (0.6)	18.6–30.1	19	25.6 (0.8)	20.2–32.3
	Men	10	26.0 (3.9)	20.1–31.8	10	26.5 (3.9)	21.4–35.1	10	25.8 (2.7)	20.7–30.1	9	26.1 (2.6)	21.9–29.0
	Women	10	24.4 (5.8)	20.1–30.7	10	25.8 (5.8)	19.1–35.8	10	24.2 (2.3)	18.6–26.7	10	25.2 (4.0)	20.2–32.3

the Functional Reach Test, and the Lateral Reach Test, gradually declined with age in women who were 20 to 80 years old, dwelled in the community, and were independently mobile. Similarly, Steffen et al[28] demonstrated a consistent trend for scores on the BBS and the TUG to decline with age in older adults dwelling in the community.

The BESTest was used to measure balance in people who were healthy in 3 previous studies.[15,19,20] However, the data were obtained for comparison with data from patients with a variety of health conditions, rather than with the specific purpose of providing normative scores that could be used as a reference for clinicians.[15,19,20] Therefore, the sizes of the healthy control group samples in those studies were small, and the authors did not provide scores by age decade. Overall mean BESTest scores in earlier work ranged from 90.6% (for people with a mean age of 65.7 years)[15] to 95.6% (for people with a mean age of 46.5 years).[19] These scores are similar to the scores that we obtained for the corresponding age groups in the present study

(95.7% for participants with a mean age of 55.5 years and 91.4% for participants with a mean age of 63.5 years).

Visual inspection of the box plots suggested a considerable increase in the variation across balance scores with age. Furthermore, although the variation in the BESTest scores of our participants who were 50 to 69 years old (SD=1.4–3.9) was similar to that reported in other studies (SD=2.9–4.8),[15,19,20] we found greater variation in scores for older participants (SD=4.6–10.8 for participants who were 70 years of age and older). These results may have been obtained because we did not control for participants' activity levels, which are known to be related to balance and to change with age.[31,32] We also did not control for the presence of comorbidities that were not thought to affect balance, and it is likely that our older participants had more comorbidities.[33] However, other normative studies of balance measures, such as single-leg stance,[34] BBS and TUG,[28] and lateral and forward reach,[30] have not yielded similar results. Another possibility is that

the BESTest[30] was able to detect a wider variety of impairments than other balance measures because of the wide variety of tasks included in the BESTest. The increased variability in BESTest, Mini-BESTest, and Brief-BESTest scores with age needs to be examined further.

Our findings fill an important knowledge gap and may facilitate the use of the BESTest, Mini-BESTest, or Brief-BESTest by clinicians. A recent survey showed that the 3 balance measures most commonly used by physical therapists in Ontario, Canada, were single-leg stance, BBS, and TUG[21]—all measures with normative data.[28,34] The reference data that we have provided by age decade for BESTest, Mini-BESTest, and Brief-BESTest scores will allow more widespread use of these tests, which are among the only tools that enable clinicians to distinguish specific subsystems contributing to impaired balance. This knowledge is essential for allowing clinicians to tailor treatment to target the specific deficits underlying the observed balance limitations in their patients.

Table 3.
BESTest, Mini-BESTest, and Brief-BESTest Scores for Canadians 50 to 89 Years of Age[a]

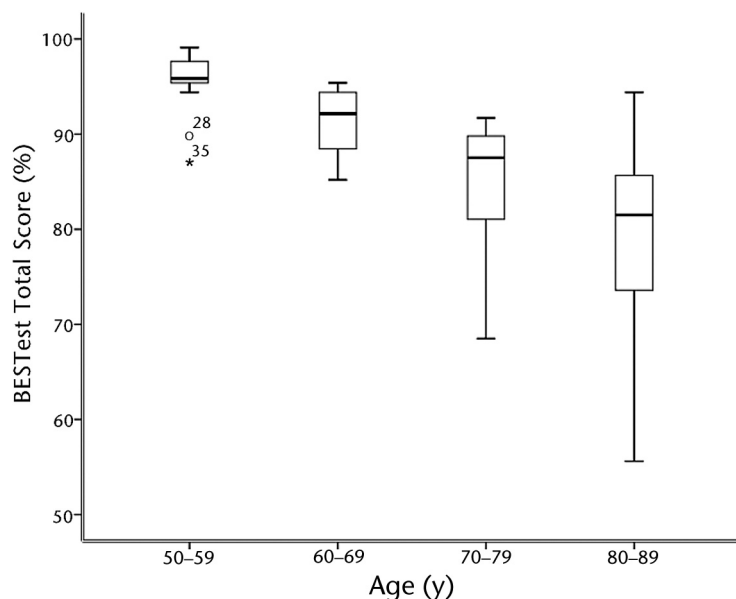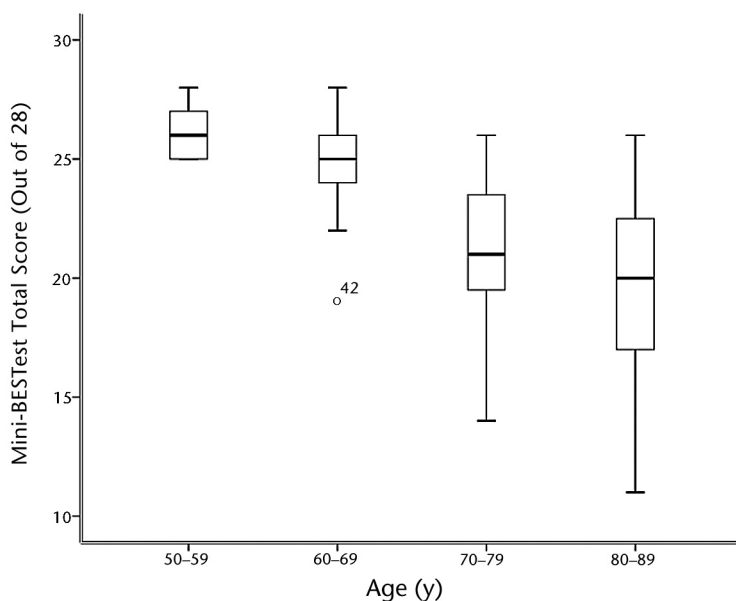| Test	Age Cohort (y)												Kruskal-Wallis Analysis		
	50–59			60–69			70–79			80–89					
	n	X̄ (SD)	95% CI	n	X̄ (SD)	95% CI	n	X̄ (SD)	95% CI	n	X̄ (SD)	95% CI	χ²	df	P
BESTest															
Total score (%)	20	95.7 (2.9)	94.4–97.1	20	91.4 (3.4)	89.8–93.0	20	85.4 (6.0)	82.5–88.2	19	79.4 (10.6)	74.3–84.5	47.990	3	<.001
Section score (%)															
Biomechanical constraints	20	96.3 (9.0)	92.1–100.6	20	89.0 (9.5)	84.6–93.4	20	83.7 (10.5)	78.8–88.6	19	78.6 (13.4)	72.1–85.1	28.843	3	<.001
Stability limits/verticality	20	94.8 (4.3)	92.7–96.8	20	92.1 (6.6)	89.1–95.2	20	87.1 (8.0)	83.4–90.9	19	85.2 (9.1)	80.8–89.6	17.463	3	.001
Anticipatory postural adjustments	20	97.8 (5.2)	95.3–100.2	20	94.4 (6.7)	91.3–97.6	20	85.6 (12.3)	79.8–91.3	19	75.1 (18.2)	66.4–83.9	28.401	3	<.001
Postural responses	20	96.9 (4.6)	94.8–99.1	20	88.3 (9.9)	83.7–92.9	20	85.5 (7.9)	81.8–89.3	19	76.9 (17.3)	68.5–85.2	29.460	3	<.001
Sensory orientation	20	98.3 (3.0)	96.9–99.7	20	96.7 (5.1)	94.3–99.0	20	94.7 (10.1)	90.0–99.4	19	88.8 (14.1)	82.0–95.5	11.782	3	.008
Stability in gait	20	92.6 (5.0)	90.3–94.9	20	90.0 (6.5)	86.9–93.1	20	77.8 (12.3)	72.1–83.6	19	73.1 (13.2)	66.8–79.5	36.571	3	<.001
Mini-BESTest (maximum score=28)	19	26.3 (1.1)	25.7–26.8	18	24.7 (2.2)	23.6–25.8	20	21.0 (3.1)	19.5–22.4	19	19.6 (4.2)	17.6–21.6	41.662	3	<.001
Brief-BESTest (maximum score)															
Total score (24)	20	22.7 (1.7)	21.9–23.5	20	20.5 (2.2)	19.5–21.6	20	18.8 (3.3)	17.3–20.4	19	15.0 (4.7)	12.8–17.3	37.608	3	<.001
Section score															
Biomechanical constraints (3)	20	2.9 (0.5)	2.6–3.1	20	2.1 (1.2)	1.5–2.6	20	2.2 (1.0)	1.8–2.7	19	1.8 (1.0)	1.4–2.3	14.392	3	.002
Stability limits/verticality (3)	20	2.7 (0.5)	2.5–2.9	20	2.5 (0.5)	2.3–2.7	20	2.4 (0.5)	2.1–2.6	19	2.0 (0.5)	1.8–2.2	16.710	3	.001
Anticipatory postural adjustments (6)	20	5.7 (0.9)	5.3–6.1	20	5.6 (0.8)	5.2–5.9	20	4.0 (2.0)	3.0–4.9	19	2.6 (2.1)	1.6–3.7	29.417	3	<.001
Postural responses (6)	20	5.7 (0.7)	5.3–6.0	20	4.9 (1.2)	4.3–5.4	20	4.8 (0.7)	4.4–5.1	19	3.8 (2.0)	2.8–4.8	17.413	3	.001
Sensory orientation (3)	20	2.8 (0.4)	2.6–3.0	20	2.6 (0.6)	2.3–2.9	20	2.7 (0.6)	2.4–2.9	19	2.1 (0.8)	1.7–2.5	10.566	3	.014
Stability in gait (3)	20	3.0 (0.0)	3.0–3.0	20	3.0 (0.0)	3.0–3.0	20	2.9 (0.3)	2.8–3.0	19	2.6 (0.5)	2.4–2.9	17.152	3	.001

[a] CI=confidence interval.

Figure 1.
Box plot comparing total BESTest scores for 4 age cohorts (*P*<.001; Kruskal-Wallis test). Minimum and maximum values, upper and lower quartiles, and the median (line inside the box) are depicted. The length of the box represents the interquartile range (IQR). Values more than 3 IQRs from either end of the box are considered extremes and are denoted by an asterisk. Values more than 1.5 IQRs but less than 3 IQRs from either end of the box are considered outliers and are denoted by "o."

Figure 2.
Box plot comparing Mini-BESTest scores for 4 age cohorts (*P*<.001; Kruskal-Wallis test). Minimum and maximum values, upper and lower quartiles, and the median (line inside the box) are depicted. The length of the box represents the interquartile range (IQR). Values more than 1.5 IQRs but less than 3 IQRs from either end of the box are considered outliers and are denoted by "o."

Limitations and Future Directions

A limitation of the present study is that it may not be generalizable because we tested only 79 Canadians who were 50 to 89 years of age. Normative scores for people outside this age range still do not exist. In addition, although our sample was representative of people who were healthy and living in the community in an urban area of Ontario, Canada, our results may not be reflective of populations in other countries. Furthermore, for men in the 80- to 89-year-old cohort, the mean age was 82.3 years and the sample size was 9. The difficulty in recruitment for this cohort could have been due to the increased number of comorbidities present in older people[32] affecting eligibility for the present study. Future studies including the administration of the BESTest in people who are healthy should aim to include larger samples and to recruit people across the lifespan and from a variety of countries.

A second limitation of the present study pertains to our inclusion and exclusion criteria, which relied solely on a participant's self-report of his or her own medical status. A more rigorous screening process involving medical examination or chart review may have increased the likelihoods of finding older adults who were healthier and observing higher scores on the balance tests. However, stricter criteria would have decreased the external validity of our findings.

A third limitation of the present study is that although we found that the scores on all of the balance tests differed significantly among the age groups, performing post hoc analyses to determine where those differences existed was beyond the scope of this study. Although a trend for scores to decrease with age was demonstrated visually in Figures 1, 2, and 3, further exploration is needed.

Although we took extra precautions to ensure consistency of scoring among testers in the present study, the training DVD that is available from the BESTest website[29] is a comprehensive training tool that should be used by clinicians before adopting this test as an outcome measure. Repeated administration of the BESTest highlighted an issue that clinicians should be aware of when interpreting scores for stability in gait. The scores in this section were the lowest among all of the sections for most of our age groups; we hypothesize that this finding may have been due to difficulties with the last item, the dual-task TUG. Participants in all age groups struggled with counting backward by 3 even before the secondary physical task was added, suggesting that this particular cognitive dual-task item may have been too difficult to distinguish among people with different levels of deficits. This observation is supported by the study of Padgett et al,[22] who found that the dual-task TUG item was the least representative item in the entire BESTest. Simplifying the cognitive task to counting backward by 2 or using a manual dual-task TUG[35] may be better alternatives to the current cognitive dual-task TUG, which can be influenced by practice and familiarity with numbers.

In summary, the present study is the first to provide for the BESTest, Mini-BESTest, and Brief-BESTest normative values for older adults who are healthy. The availability of normative values may enhance the utility of these tools as comprehensive measures of balance for clinicians to use with a wide variety of patients. Further research should focus on the predictive validity, reliability, and responsiveness of these tests in people who are healthy as well as the relationship between balance scores and physical activity levels.
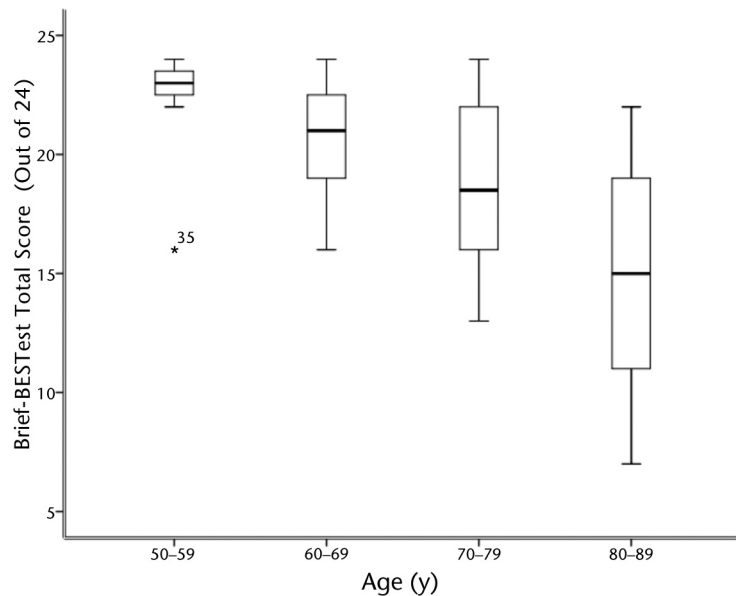

Figure 3.
Box plot comparing total Brief-BESTest scores for 4 age cohorts ($P<.001$; Kruskal-Wallis test). Minimum and maximum values, upper and lower quartiles, and the median (line inside the box) are depicted. The length of the box represents the interquartile range (IQR). Values more than 3 IQRs from either end of the box are considered extremes and are denoted by an asterisk.

References

1 Tinetti ME, Speechley M, Ginter S. Risk-factors for falls among elderly persons living in the community. *N Engl J Med.* 1988; 319:1701–1707.

2 SMARTRISK. *The Economic Burden of Injury in Canada.* Toronto, Ontario, Canada: SMARTRISK; 2009.

3 Rubenstein LZ. Falls in older people: epidemiology, risk factors and strategies for prevention. *Age Ageing.* 2006;35:37–41.

4 Sattin R. Falls among older persons: a public-health perspective. *Annu Rev Public Health.* 1992;13:489–508.

5 Tinetti ME, Kumar C. The patient who falls: "It's always a trade off." *JAMA.* 2010; 303:258–266.

6 Fabre JM, Ellis R, Kosma M, Wood RH. Falls risk factors and a compendium of falls risk screening instruments. *J Geriatr Phys Ther.* 2010;33:184–197.

7 Shumway-Cook A. *Motor Control: Translating Research Into Clinical Practice.* 3rd ed. Baltimore, MD: Lippincott Williams & Wilkins; 2007.

8 Horak FB. Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls? *Age Ageing.* 2006;35(suppl 2):7–11.

9 Newton RA. Balance and falls among older people. *Generations.* 2003;27:27–31.

10 Berg K, Wood-Dauphinée S, Williams J. Measuring balance in the elderly: validation of an instrument. *Can J Public Health.* 1992;83:S7–S11.

11 Podsiadlo D, Richardson S. The Timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc.* 1991;39:142–148.

12 Muir SW, Berg K, Chesworth B, Speechley M. Use of the Berg Balance Scale for predicting multiple falls in community-dwelling elderly people: a prospective study. *Phys Ther.* 2008;88:449–459.

13 Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther.* 2000; 80:896–903.

14 Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the probability for falls in community-dwelling older adults. *Phys Ther.* 1997;77:812–819.

15 Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. *Phys Ther.* 2009;89:484–498.

16 Leddy AL, Crowner BE, Earhart GM. Functional gait assessment and balance evaluation system test: reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson disease who fall. *Phys Ther.* 2011;91:102–113.

17 Leddy AL, Crowner BE, Earhart GM. Utility of the Mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with Parkinson disease. *J Neurol Phys Ther.* 2011;35:90–97.

18 Kurz MJ, Corr B, Stuberg W, et al. Evaluation of lower body positive pressure supported treadmill training for children with cerebral palsy. *Pediatr Phys Ther.* 2011; 23:232–239.

19 Jones KD, Horak FB, Winters-Stone K, et al. Fibromyalgia is associated with impaired balance and falls. *J Clin Rheumatol.* 2009;15:16–21.

20 Beauchamp MK, Sibley KM, Lakhani B, et al. Impairments in systems underlying control of balance in COPD. *Chest.* 2012; 141:1496–1503.

21 Sibley KM, Straus SE, Inness EL, et al. Balance assessment practices and use of standardized balance measures among Ontario physical therapists. *Phys Ther.* 2011;91: 1583–1591.

22 Padgett PK, Jacobs JV, Kasser SL. Is the BESTest at its best? A suggested brief version based on interrater reliability, validity, internal consistency, and theoretical construct. *Phys Ther.* 2012;92:1197–1207.

23 Franchignoni F, Horak F, Godi M, et al. Using psychometric techniques to improve the Balance Evaluation Systems Test: the Mini-BESTest. *J Rehabil Med.* 2010;42:323–331.

24 King L, Horak F. On the Mini-BESTest: scoring and the reporting of total scores [letter]. *Phys Ther.* 2013;93:571–575.

25 McNeely ME, Duncan RP, Earhart GM. Medication improves balance and complex gait performance in Parkinson disease. *Gait Posture.* 2012;36:144–148.

26 King LA, Priest KC, Salarian A, et al. Comparing the Mini-BESTest with the Berg Balance Scale to evaluate balance disorders in Parkinson's disease. *Parkinsons Dis.* 2012;2012:375–419.

27 Hill K, Wickerson LM, Woon LJ, et al. The 6-min walk test: responses in healthy Canadians aged 45 to 85 years. *Appl Physiol Nutr Metab.* 2011;36:643–649.

28 Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Phys Ther.* 2002;82:128–137.

29 Horak F. BESTest: Balance Evaluation Systems Test. Available at: http://www.ohsu.edu/tech-transfer/portal/technology.php?technology_id=217191. Accessed November 19, 2013.

30 Isles R, Choy N, Steer M, Nitz J. Normal values of balance tests in women aged 20–80. *J Am Geriatr Soc.* 2004;52:1367–1372.

31 Howe TE, Rochester L, Jackson A, et al. Exercise for improving balance in older people. *Cochrane Database Syst Rev.* 2007;(4):CD004963.

32 Milanovi Z, Pantelli S, Trajkovi N, et al. Age-related decrease in physical activity and functional fitness among elderly men and women. *Clin Interv Aging.* 2013;8: 549–556.

33 Fortin M, Bravo G, Hudon C, et al. Prevalence of multimorbidity among adults seen in family practice. *Ann Fam Med.* 2005;3: 223–228.

34 Springer BA, Marin R, Cyhan T, et al. Normative values for the unipedal stance test with eyes open and closed. *J Geriatr Phys Ther.* 2007;30:8–15.

35 Hofheinz M, Schusterschitz C. Dual task interference in estimating the risk of falls and measuring change: a comparative, psychometric study of four measurements. *Clin Rehabil.* 2010;24:831–842.