Reference Group Data for the Functional Gait Assessment

Martha L Walker, Alvis G Austin, Gina M Banke, Suzanne R Foxx, Lynn Gaetano, Laurie A Gardner, Jill McElhiney, Kisiah Morris, Liz Penn

Background and Purpose

The Functional Gait Assessment (FGA) is a clinical tool for evaluating performance in walking. The purpose of this study was to determine age-referenced norms for performance on the FGA in community-living older adults.

Subjects

Subjects were 200 adults, ages 40 to 89 years, living independently.

Methods

Each subject completed the FGA one time and was scored simultaneously by 2 testers.

Results

The intraclass correlation coefficient for interrater reliability was .93. Mean scores for the FGA ranged from 29/30 for adults in their 40s to 21/30 for adults in their 80s.

Discussion and Conclusion

Patient performance on the FGA can be compared with age-referenced norms for expected performance. Further research is needed to determine the FGA's usefulness in tracking clinical changes or predicting falls. The FGA is a reliable test for people without disease, and it is able to detect decreases in gait performance among typical older adults.

ML Walker, PT, PhD, is Associate Professor and Chair, School of Physical Therapy, Health Sciences Building, Room 3118, Old Dominion University, Norfolk, VA 23529 (USA). Address all correspondence to Dr Walker at: mlwalker@odu.edu.

AG Austin, PT, DPT, is Staff Physical Therapist, RehabCare Group Inc, St Francis Nursing Center, Newport News, Va.

GM Banke, PT, DPT, is Staff Physical Therapist, Sentara Bayside Therapy Center, Virginia Beach, Va.

SR Foxx, PT, DPT, is Physical Therapist and Owner, Physical Therapy & Sports Medicine, Newport News, Va.

L Gaetano, PT, DPT, is Staff Physical Therapist, Therapy Associates, Chesapeake, Va.

LA Gardner, PT, DPT, is Staff Physical Therapist, Sentara Louise Obici Memorial Hospital, Suffolk, Va.

J McElhiney, PT, DPT, is Staff Physical Therapist, RehabCare Group Inc, Norfolk, Va.

K Morris, PT, DPT, is Staff Physical Therapist, Lewis Gale Medical Center, Salem, Va.

L Penn, PT, DPT, is Staff Physical Therapist, Sentara Williamsburg Regional Medical Center, Williamsburg, Va.

Dr Austin, Dr Banke, Dr Foxx, Dr Gaetano, Dr Gardner, Dr McElhiney, Dr Morris, and Dr Penn were DPT candidates at Old Dominion University during their participation in this study.

[Walker ML, Austin AG, Banke GM, et al. Reference group data for the Functional Gait Assessment. *Phys Ther*. 2007;87:1468–1477.]

The Functional Gait Assessment (FGA) is a standardized test for assessing postural stability during various walking tasks.[1] The test is a modified version of the Dynamic Gait Index (DGI), which was developed to assess gait and risk of falling in adults over 60 years of age by testing their ability to respond to changing gait tasks and requirements.[2,3] A score of 19 or less out of a possible 24 on the DGI has been found to be associated with an increased risk for falling in community-dwelling older adults and in individuals with vestibular dysfunction.[4,5] Hall et al[6] found the DGI to be useful in an equation to predict fall risk outcome based on patient assessment at the beginning of rehabilitation. The DGI total score has been shown to be reliable in patients with confirmed peripheral vestibular disorders[7,8] and in those with multiple sclerosis,[9] although individual items have not been shown to be uniformly reliable. It also has been shown to be useful as an outcome measure in people with stroke.[10] In concurrent validity tests, DGI scores have been correlated with Berg Balance Scale scores ($r \geq .78$)[10,11] and measurements of walking speed ($r \geq .85$),[10] Timed "Up & Go" Test scores ($r \geq .78$),[11] Dizziness Handicap Inventory scores ($r = .69$),[12] and Activities-specific Balance Confidence Scale scores ($r \geq .58$).[13] The DGI has been shown to exhibit sound psychometric properties and is considered an appropriate test for assessing community-dwelling older subjects with balance problems.[14]

In testing patients with vestibular disorders, one group of researchers[15] found that some patients scored well even though they felt they had impairments in walking and reported moderate disability due to dizziness. The DGI, therefore, was not considered sensitive enough to detect subtle changes in gait performance that these patients with vestibular disorders experienced.[15] To address this ceiling effect, Wrisley et al[1] revised the test to create the FGA, which consists of 7 of the 8 tasks from the DGI and 3 new tasks: gait with narrow base of support, ambulating backward, and gait with eyes closed. The task that was deleted was stepping around an obstacle. The added tasks were chosen to increase the challenge of the assessment so that the test would be more sensitive to minor changes in gait stability during walking.

In designing the FGA, Wrisley et al[1] also attempted to clarify administrative instructions and operational definitions to improve reliability of individual test items. The result was the development of a test that offers a broad composite assessment of gait and dynamic stability while remaining easy to administer in most clinical settings. All that is required is a stopwatch, a marked walking area, shoeboxes for obstacles, and a set of steps. Scoring for each FGA item ranges from 0 for severe impairment to 3 for normal performance. The highest score possible is 30.

Wrisley et al[1] tested the reliability, internal consistency, and validity of data obtained from the FGA when used with people with vestibular disorders. Six patients were rated with the FGA instrument by 10 raters twice. All raters were present and rated each patient at the same time. They had no training ahead of time, but were given 10 minutes to read the instructions, test items, and grading criteria. Intraclass correlation coefficients (ICCs) were .84 and .83 for interrater and intrarater reliability. Internal consistency of FGA scores, as measured by the Cronbach alpha, was .79. To test for validity, several other balance and gait tests were administered after the FGA by a different rater who was not involved in scoring the FGA. The FGA was found to have Spearman correlations (r) of $-.64$ with the Dizziness Handicap Inventory and of $-.66$ with reported number of falls.[1] The Pearson correlation (r) with the Timed "Up & Go" Test was .50. The authors felt that these findings demonstrated acceptable reliability, internal consistency, and concurrent validity with some of the balance measures used for patients with vestibular disorders.

Although the FGA was developed to address a ceiling effect in patients with vestibular disorders, it is essentially a revised DGI and may be useful in the same population of older adults. Testing is necessary, however, to determine reliability on a larger group of subjects and the expected range of scores for adults throughout the life span. Because the test may be used with older adults, we need to understand whether older adults are expected to achieve a score at or near 30. If the FGA is sensitive enough to detect gait changes with age, then we may see a decrease in total scores on the test across the decades.

The purpose of this study was to establish reference group data for the FGA by decade cohorts including people 40 to 89 years of age. We expected that average performance on the test would not decrease until middle age or later, so we did not include young adults in this study. The results of a reference group study should allow for patient scores to be compared with age-referenced norms rather than with a standard of perfect performance. Another component of this study was a test of interrater reliability for total FGA scores and percentage of agreement for individual item scores.

Method

Participants

Subjects were community-dwelling adults aged 40 to 89 years. For the purposes of our study, *community dwellers* were defined as people living independently with no assistance in activities of daily living. We relied on subjects' self-reports to determine whether criteria were met. We recruited our subjects from various settings, including churches, retirement communities with independent apartments, a local police department, funeral home employees, people attending a university-sponsored community health fair, and a Jewish Community Center. Individuals were invited to participate if on a written questionnaire they reported having no history of vestibular problems or dizziness, neurological disorders, cerebral palsy, stroke or amputation, or any other serious medical conditions that limited their mobility. None of our subjects used an assistive device during testing. Subjects were required to understand and sign a consent form approved by the Old Dominion University Institutional Review Board and to follow verbal commands.

Data collection was performed at different locations as a convenience to subjects and to reach the largest number of test subjects. A vinyl gait grid (Gait Grid*) was rolled out and taped to the floor to prevent sliding so that subjects tested at different locations would have the same surface on which to walk. A tape measure was used to determine boundaries for the test, and masking tape was used to mark these boundaries on the floor.

Once we determined that a participant had met our inclusion criteria, a gait belt was placed around the subject's waist and testing was initiated. Although we did not expect subjects to fall, we wanted to be cautious with subjects who were up to 89 years of age. We also reasoned that a clinician using the test might use a gait belt, so we made that a standard condition of testing for all subjects. Testers used standby guarding, and no tester had to actually touch the gait belt. A total of 8 testers worked in teams of 2 to test the 200 subjects. Before data collection, all testers underwent a training session and a separate practice testing session using 5 to 10 volunteers of various ages. Two testers evaluated each participant. One tester was stationary off to the side of the pathway, giving the verbal instructions to the participant directly from the FGA instrument. The second tester demonstrated the tasks when needed and walked with the participant. Both testers had stopwatches and score sheets, and they scored each participant independently. Each participant performed the tasks of the FGA as instructed by the tester without practice. Each item was performed once.

Data Analysis

We used the ICC (model 2) to evaluate interrater reliability for the total score. Percentage of agreement and kappa values also were calculated to determine variability between raters for individual test items. With reliability established, the rest of the statistics were calculated using only the first tester's scores rather than averaging the scores, because normally a therapist would have only one score for each patient. A Spearman rho correlation coefficient was calculated to determine whether age was significantly related to total score. Data then were divided into groups based on age by decades (40–49, 50–59, 60–69, 70–79, and 80–89). The mean, range, and standard deviation were calculated for subjects' scores for each test item by decade.

Results

The total number of subjects tested using the FGA was 200. The mean age of all subjects was 65.7 years (range=40–89). The subjects were 136 women (68%) and 64 men (32%). The mean score of all subjects combined was 26.1, with a standard deviation of 3.97 (95% confidence interval [CI]=25.5–26.6). Figure 1 shows a histogram of these scores. Fifteen of the subjects (10 men and 5 women) scored lower than 20 total points. The mean age of these subjects was 82.6 years (range=76–89).

Because twice as many women as men were entered into the study, we performed t tests to determine whether men and women scored differently on the FGA. We found that there were no significant differences in overall FGA scores between men and women when examined as a group ($t=1.48$, $P<.139$) or by decade (t values ranged from $-.14$ to 1.59; P values ranged from $<.118$ to .978) (Fig. 2).

Figure 1 illustrates the means and 95% confidence intervals of the total scores by decade. Table 1 shows the descriptive data and statistics for the total scores by decade cohorts. Mean total scores ranged from 28.9 for the fifth decade to 20.8 for the ninth decade. The 95% CIs also are reported in Table 1. Tables 2, 3, 4, 5, 6, and 7 show the means and 95% CIs for individual items for the total group and by decade. Figure 3 shows a graphic comparison of mean scores of each item by decade. Item 5 (gait and pivot turn) showed the highest mean scores, with subjects in all decades averaging greater than 2.5. Items 1 (gait on level surface), 7 (gait with narrow base of support), and 8 (gait with eyes closed) showed the lowest mean scores, with the means decreasing over the decades.

The ICC for intertester reliability was .93 ($P<.001$). Percentages of agreement between testers ranged from 78.5% to 96.0% (mean=87%) for the 10 items (Tab. 8). Kappa values for all items were significant at the $P>.001$ level. Individual kappa statistics, also listed in Table 8, ranged from .43 to .77, with an average of

* EFI Total Gym, 7755 Arjons Dr, San Diego, CA 92126.

.63. The Spearman rho correlation coefficient between age and total score was −.64 ($P<.001$, $r^2=.41$), indicating a significant negative relationship between the 2 variables.

Discussion

Mean total scores for the FGA showed a systematic decrease with increased age, as illustrated by the negative correlation between age and FGA total score. This decrease became more pronounced in subjects aged 70 years and older. At the same time, there was an increase in the standard deviation of total scores with each decade, demonstrating that variability of performance on the FGA increased with age. The clinical relevance of this is that clinicians should expect that subjects up to 59 years of age will score near the top of the total FGA scale. The 60 individuals in the 40- to 49-year and 50- to 59-year age groups did not show much variation, so if we see patients in this age range with scores lower than around 24 or 25, we may want to examine further to look for causes of poor performance. That is not to say that these patients will be at risk for falling, because we did not test for that.

The results of this study are in accordance with the findings of Lusardi et al.[16] They tested 76 community-living older adults (aged 66–101 years) using 7 measures of functional performance: comfortable gait speed, fast gait speed, Berg Balance Scale, 6-minute walk test, Timed "Up & Go" Test, Physical Performance Test, and timed sit-to-stand test. On all of the tests, performance decreased by decade, and on all tests except the 6-minute walk test, the standard deviation of the scores increased by decade. Lusardi et al also found that sex was not a significant predictor of functional performance in their sample. This finding is in accordance with our findings that there was no difference in perfor-

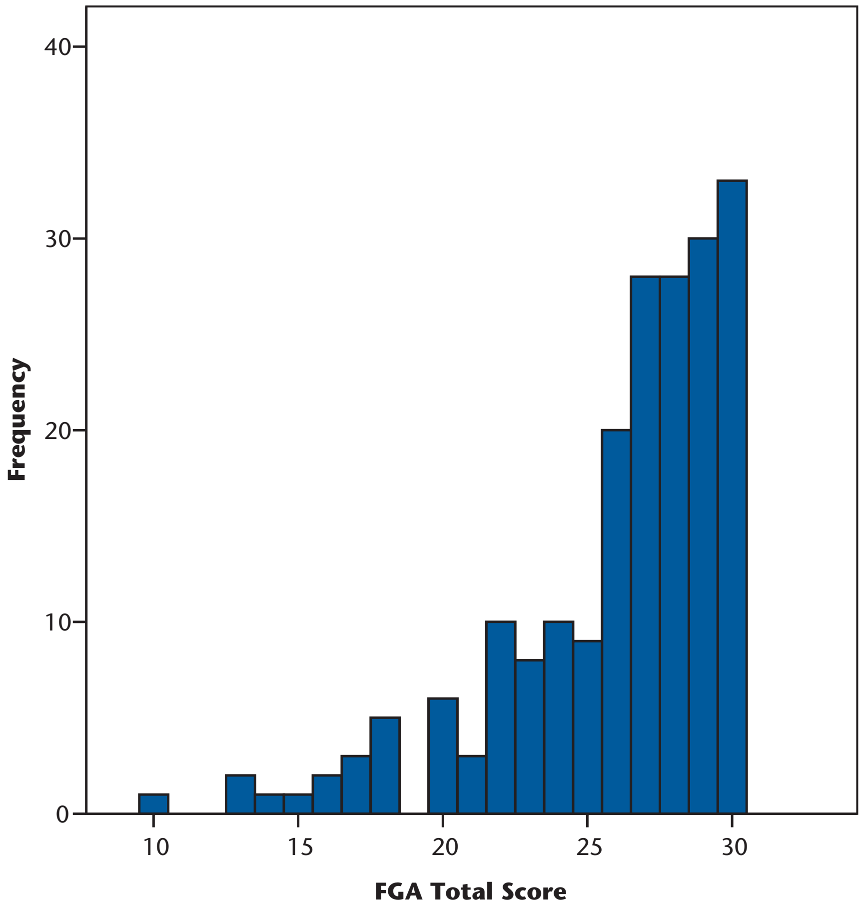

Figure 1.
Histogram of Functional Gait Assessment (FGA) total scores (N=200). Mean=26.1, SD=4.0. Fifteen participants had scores below 20.

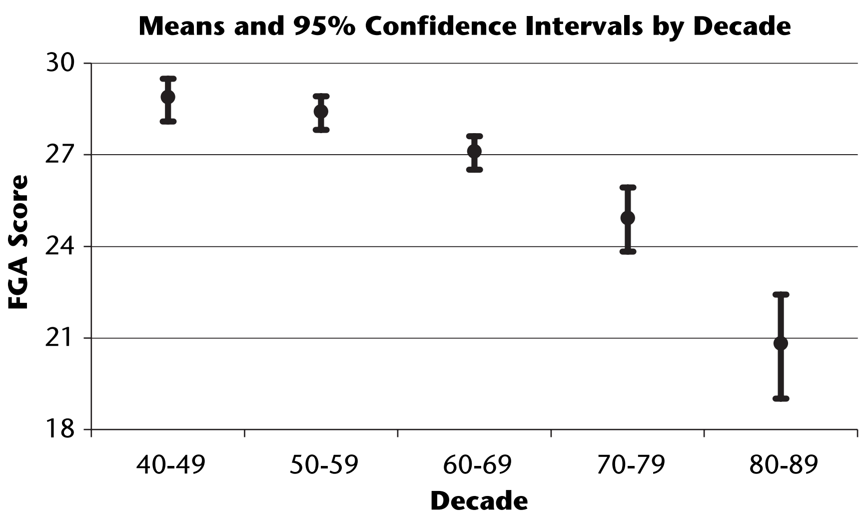

Figure 2.
Comparison of the mean Functional Gait Assessment (FGA) total scores and 95% confidence intervals by decade. A perfect score is 30.

Table 1.
Functional Gait Assessment Total Scores by Decade

Age (y)	N	Minimum Score	Maximum Score	Mean	SD	95% Confidence Interval
40–49	27	24	30	28.9	1.5	28.3–29.5
50–59	33	25	30	28.4	1.6	27.9–29.0
60–69	63	20	30	27.1	2.3	26.5–27.7
70–79	44	16	30	24.9	3.6	23.9–26.0
80–89	33	10	28	20.8	4.7	19.2–22.6
Total	200	10	30	26.1	4.0	25.5–26.6

Table 2.
Functional Gait Assessment Means, Standard Deviations, and 95% Confidence Intervals (CIs) for Total Group (N=200)[a]

Item	Mean	SD	95% CI
1. Gait on level surface	2.5	0.66	2.4–2.5
2. Change in gait speed	2.8	0.51	2.7–2.8
3. Gait with horizontal head turns	2.8	0.52	2.7–2.9
4. Gait with vertical head turns	2.9	0.36	2.8–2.9
5. Gait with pivot turn	3.0	0.22	2.9–3.0
6. Step over obstacle	2.6	0.69	2.5–2.7
7. Gait with narrow base of support	2.1	1.13	1.9–2.3
8. Gait with eyes closed	2.0	1.00	1.9–2.1
9. Ambulating backward	2.8	0.49	2.7–2.9
10. Steps	2.8	0.24	2.7–2.8

[a] Range of scores possible is 0 to 3.

Table 3.
Functional Gait Assessment Means, Standard Deviations, and 95% Confidence Intervals (CIs) for Subjects Aged 40 to 49 Years (n=27)[a]

Item	Mean	SD	95% CI
1. Gait on level surface	2.7	0.47	2.5–2.9
2. Change in gait speed	3.0	0.0	3.0–3.0
3. Gait with horizontal head turns	3.0	0.0	3.0–3.0
4. Gait with vertical head turns	3.0	0.0	3.0–3.0
5. Gait with pivot turn	3.0	0.0	3.0–3.0
6. Step over obstacle	3.0	0.19	2.9–3.0
7. Gait with narrow base of support	2.9	0.53	2.6–3.0
8. Gait with eyes closed	2.4	0.93	2.0–2.7
9. Ambulating backward	3.0	0.0	3.0–3.0
10. Steps	3.0	0.0	3.0–3.0

[a] Range of scores possible is 0 to 3.

mance between male and female subjects in any of the decades tested.

Fregly et al[17] performed a battery of ataxia tests with 1,055 men, aged 16 to 60 years, who were healthy. The tests included eyes-open and eyes-closed standing and walking heel-to-toe on a ¾-in (1.9-cm) rail, single-leg stance on the floor, and Sharpened Romberg test. They found a decrease in performance with age beginning at age 35 years. The current study of the FGA used 40 years as the youngest age of subjects. Middle-aged adults were included in data collection to establish a baseline from which decline in scores could be measured. Had we included individuals aged 20 to 39 years, we may have seen even slightly higher scores than from those in their 40s. However, variation in performance at the very top of the FGA scale was not the primary focus of our study.

Whitney et al[18] did not find age to be a significant factor in vestibular rehabilitation outcomes. In their study, patients with vestibular problems were divided into 2 groups: one with an average age of 33 years and the other with an average age of 70 years. The younger patients started out with lower DGI scores (15 versus 18), a larger percentage of them had abnormal caloric test results, and more of them reported falls than the older group. There was no statistically significant difference in number of treatment visits between age groups; younger patients were treated for an average of 4 visits, and older patients were treated for an average of 3.2 visits. Both groups made significant clinical improvement, and both groups had an average total score of nearly 21 out of 24 after intervention. The lack of difference in outcome scores, despite the difference in age between the 2 groups, may demonstrate that vestibular dysfunction was a bigger factor in performance than age. In addition, the DGI may not be as sensitive to age-related differences in performance as is the FGA.

In the older 3 age groups of our study, not only were the mean scores lower, but the variation increased tremendously. Clinicians, therefore, should expect lower scores and increased variability when testing older people who are healthy. Some of the scores in our test group were surprisingly low. Fifteen of the 200 participants had a total score below 20, which was the original DGI cutoff score for risk for falling. The mean age of these 15 low-scoring participants was 82.6 years (range=76–89). It could be argued that these subjects were not healthy; their functional limitations, as shown by low scores, may indicate some underlying disorder. We did not perform an examination on any of the subjects. We did make it clear to them that they were required to be living in the community and functioning independently, and all of the subjects indicated that they were. Although no subjects used an assistive device during the FGA, we do not know whether any of them used furniture or a wall for support in their daily activities.

If we take the data for these lowest-scoring subjects out of the calculations, the mean score for the subjects aged 70 to 79 years becomes 25.5 (n=41) rather than 24.9 (n=44), and the mean score for the subjects aged 80 to 89 years becomes 23.8 (n=21) rather than 20.8 (n=33). These values may be a more accurate reflection of a healthy aging population. Clinicians should be aware, however, that some people who state that they are independent may be functioning at a lower level than they admit to.

Some of the 10 tasks proved more difficult for older adults. Subjects aged 70 to 89 years scored lower on item 1 (gait on level surface). Winter et al[19] documented that older adults walk with decreased speed compared with younger adults, and Lusardi et al[16] also found this to be true. Lower scores on item 1 confirm that the FGA is sensitive to this relationship between age and gait speed. Subjects aged 60 to 89 years also had difficulty with items 7 (gait with narrow base of support) and 8 (gait with eyes closed). Loss of balance occurs when a person's body sway exceeds the limits of stability.[20] The decreased base of support that occurs in tandem stance or walking decreases the limits of stability, which increases the difficulty of maintaining balance. Murphy et al[21] found that the ability to perform tandem stance was negatively associated with age. Similarly, Fregly et al[17] found that, as age increased, there was a drop in ability to perform heel-to-toe walking on a ¾-in-wide rail. They also documented a decrease in performance with age for eyes-closed tests of both double-leg tan-

Table 4.
Functional Gait Assessment Means, Standard Deviations, and 95% Confidence Intervals (CIs) for Subjects Aged 50 to 59 Years (n=27)[a]

Item	Mean	SD	95% CI
1. Gait on level surface	2.7	0.47	2.5–2.9
2. Change in gait speed	2.9	0.36	2.7–3.0
3. Gait with horizontal head turns	2.9	0.24	2.9–3.0
4. Gait with vertical head turns	3.0	0.17	2.9–3.0
5. Gait with pivot turn	3.0	0.0	3.0–3.0
6. Step over obstacle	2.7	0.51	2.5–2.9
7. Gait with narrow base of support	2.7	0.68	2.5–2.9
8. Gait with eyes closed	2.7	0.54	2.5–2.7
9. Ambulating backward	3.0	0.17	2.9–3.0
10. Steps	2.9	0.29	2.8–3.0

[a] Range of scores possible is 0 to 3.

Table 5.
Functional Gait Assessment Means, Standard Deviations, and 95% Confidence Intervals (CIs) for Subjects Aged 60 to 69 Years (n=63)[a]

Item	Mean	SD	95% CI
1. Gait on level surface	2.6	0.30	2.5–2.8
2. Change in gait speed	2.9	0.36	2.8–3.0
3. Gait with horizontal head turns	2.9	0.46	2.8–3.0
4. Gait with vertical head turns	3.0	0.18	2.9–3.0
5. Gait with pivot turn	3.0	0.0	3.0–3.0
6. Step over obstacle	2.8	0.22	2.6–2.9
7. Gait with narrow base of support	2.2	1.04	2.0–2.5
8. Gait with eyes closed	2.0	1.03	1.7–2.2
9. Ambulating backward	2.9	0.27	2.9–3.0
10. Steps	2.9	0.36	2.8–3.0

[a] Range of scores possible is 0 to 3.

Table 6.
Functional Gait Assessment Means, Standard Deviations, and 95% Confidence Intervals for Subjects Aged 70 to 79 Years (n=44)[a]

Item	Mean	SD	95% CI
1. Gait on level surface	2.2	0.71	2.0–2.4
2. Change in gait speed	2.7	0.56	2.5–2.9
3. Gait with horizontal head turns	2.7	0.56	2.5–2.9
4. Gait with vertical head turns	2.8	0.37	2.7–3.0
5. Gait with pivot turn	2.9	0.36	2.8–3.0
6. Step over obstacle	2.5	0.88	2.2–2.7
7. Gait with narrow base of support	1.9	1.09	1.6–2.3
8. Gait with eyes closed	1.8	0.91	1.6–2.1
9. Ambulating backward	2.8	0.44	2.6–2.9
10. Steps	2.6	0.62	2.4–2.8

[a] Range of scores possible is 0 to 3.

dem stance and single-leg stance. The ability to maintain postural stability is dependent on the interaction of 3 sensory inputs: vestibular, visual, and somatosensory.[2] Of these 3 inputs, vision appears to play a larger role in postural stability in older adults than in younger adults.[22] This would explain why an older adult would have relatively greater difficulty on an eyes-closed task. The 15 subjects who scored below 20 performed the worst on gait with a narrow base of support (mean=0.2 of a possible 3) and on gait with eyes closed (mean=0.73).

In contrast to our findings, Chiu et al,[14] in their analysis of the DGI, found that the 3 most difficult tasks for their subjects were horizontal head turns, steps, and vertical head turns. The differences in their findings and ours are probably due to the differences in subjects and the differences in tests. They studied 84 older adults who had been treated for balance problems. Their sample included subjects with dizziness and vestibular dysfunction, which could have accounted for their subjects experiencing greater difficulties with tasks involving head turns. Chiu et al also were using the DGI, rather than the FGA. Items 7, 8, and 9 were added to the DGI to form the FGA in the hope that a ceiling effect that was seen in younger patients with mild vestibular disorders might be reduced.[1] The results of the current study indicate that items 7 and 8 are quite challenging, particularly for older adults. It remains to be seen whether patients with vestibular disorders are similarly challenged and whether the test is sensitive enough to detect improvements as patients progress through treatment.

Richardson et al[23] found that, in order to distinguish fallers from nonfallers among older adults, it was necessary to present them with a challenging environment. The DGI was able to identify people at risk for falling among older adults and patients with vestibular disorders.[4,5] The FGA, with many of the same tasks, should be examined to determine whether it is useful in predicting risk for falling.

We concur with Wrisley et al[1] that the FGA total score demonstrates acceptable reliability. We obtained an ICC of .93, which is higher than in the study by Wrisley et al (ICC=.86). Raters in our study were trained and participated in practice sessions prior to data collection. The study was conducted over a 2-year period, with 4 students collecting data each year. At the start of the study, the supervising faculty member trained the first 4 students, and the students then practiced with 10 subjects and discussed their results to improve consistency. The students did not establish any new rules for the test, as their intention was to use the FGA as written. When 4 students were recruited for the second year of the

Table 7.
Functional Gait Assessment Means, Standard Deviations, and 95% Confidence Intervals (CIs) for Subjects Aged 80 to 89 Years (n=33)[a]

Item	Mean	SD	95% CI
1. Gait on level surface	1.9	0.70	1.7–2.2
2. Change in gait speed	2.3	0.74	2.1–2.6
3. Gait with horizontal head turns	2.4	0.70	2.1–2.6
4. Gait with vertical head turns	2.6	0.62	2.3–2.8
5. Gait with pivot turn	2.9	0.33	2.8–3.0
6. Step over obstacle	2.1	0.89	1.8–2.4
7. Gait with narrow base of support	0.8	0.98	0.5–1.2
8. Gait with eyes closed	1.2	0.86	0.9–1.5
9. Ambulating backward	2.2	0.79	2.0–2.5
10. Steps	2.4	0.60	2.2–2.6

[a] Range of scores possible is 0 to 3.

study, students from the first year of the study trained them in the use of the FGA, and again, the students practiced scoring 10 subjects. Once the training was completed and testing began, the students did not discuss their results with each other to try to increase reliability. The training and practice, which were not done in the study by Wrisley et al, probably improved interrater reliability. Including a large number of subjects of varying abilities also may have improved reliability, because between-subject variability increased relative to within-subject variability, and that affects the ICC calculation.[24] Reliability of individual pairs of testers was not examined.

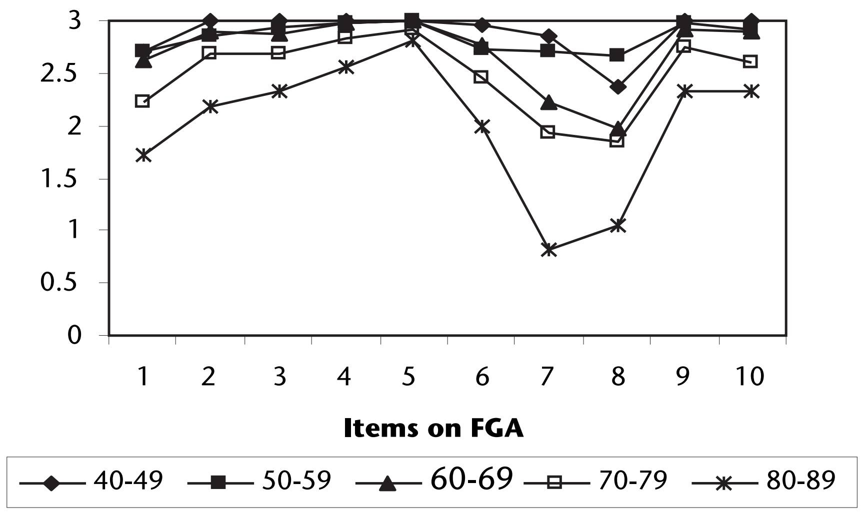

Figure 3.
Mean score of each Functional Gait Assessment (FGA) item by decade. On the Y axis, scores for each item can range from 0 (severe impairment) to 3 (normal). On the X axis are the 10 items of the FGA: 1=gait on level surface, 2=change in gait speed, 3=gait with horizontal head turns, 4=gait with vertical head turns, 5=gait with pivot turn, 6=step over obstacle, 7=gait with narrow base of support, 8=gait with eyes closed, 9=ambulating backward, 10=steps.

Although reliability of the FGA total score was high, reliability of scores for individual test items was more varied. Percentage of agreement between raters ranged from 78.5% to 96%. The lowest percentage of agreement was obtained on item 8 (gait with eyes closed). This item, which older adults tended to score worse on, required the tester to judge whether the subject deviated from a straight pathway and whether that deviation was up to 10 in (25.4 cm) (score 2), 10 to 15 in (38.1 cm) (score 1), or greater than 15 in. In our study, older subjects often deviated from straight ahead. The difficulty of a tester accurately determining the exact inches of deviation probably contributed to decreased reliability of this item. Gait with eyes closed was not one of the lowest percentage of agreement items in Wrisley and colleagues' study of people with vestibular disorders.[1] They found the lowest agreement (58%) on item 3 (gait with horizontal head turns). The differences in subjects between the 2 studies may account for differences in test performance.

Kappa values were statistically significant and ranged from .43 to .77. Using a previously published rating system, these values would be considered as having fair to good agreement.[25] The lowest kappa value that we found was .43 on item 2 (change in gait speed). We found an interrater agreement of 82.5% on this item. This item also received a low intertester kappa score of .37 by Wrisley et al[1] and an interrater agreement of 60%. Some of the lower reliability rating for this item may be due to the multidimensional judgments required during scoring. The score for this item is determined by the tester making a judgment of whether the change in speed was significant and without gait deviations (3), the change in speed was significant but

Table 8.
Mean Interrater Percentage of Agreement and Kappa Statistics for Each Item of the Functional Gait Assessment

Item	Percentage of Agreement	Kappa
Mean	87.0	.63
1. Gait on level surface	87.0	.77
2. Change in gait speed	82.5	.43
3. Gait with horizontal head turns	88.5	.64
4. Gait with vertical head turns	89.0	.45
5. Gait with pivot turn	96.0	.52
6. Step over obstacle	87.0	.71
7. Gait with narrow base of support	83.0	.73
8. Gait with eyes closed	78.5	.70
9. Ambulating backward	87.0	.57
10. Steps	91.5	.74

with mild gait deviations (2), the change in speed was not significant enough or without significant gait deviations (1), or the subject cannot change speeds (0). With both degree of change of speed and amount of gait deviation determined by the tester, this item has several dimensions on which testers can disagree and that may contribute to the lower scoring reliability.

There are 2 potential limitations in using kappa scores to determine reliability. The first potential limitation is data distribution and, in particular, score variability. In the example cited above (ie, change in gait speed), the percentage of agreement was 82.5%, and the kappa score was .43. In contrast to this, item 7 (gait with narrow base of support) had a similar percentage of agreement of 83%, but the kappa score was .73. The difference is that item 7 had a standard deviation more than twice as large: 1.131 versus 0.514. Kappa is purported to account for the amount of agreement that is due to chance. If most of the scores are similar (ie, low standard deviation), then it is presumed that a greater amount of agreement that occurs does so due to chance, and the kappa score will be lower than with the same percentage of agreement and more variability. A second potential limitation of kappa is that it was developed for binomial (yes/no) data and may not be valid for use with a test such as the FGA, which has 4 scoring categories.[26] Despite this controversy, kappa scores are commonly used measures of reliability in medical research and are presented here alongside percentages of agreement for the reader's interpretation.

A limitation of this study is that we did not know the activity levels of the participants, and so we could not determine a mean score for an active older adult who was healthy versus an activity-limited individual. A limitation of the FGA is that item 1 requires 20 feet (6.1 m) of clear walkway in order to test walking speed. This distance may not be available in a home health environment, limiting the therapist's ability to score the test correctly. Because many patients at risk for falling receive therapy in the home setting, future research should address whether a 3.05-m (10-ft) walkway is adequate to determine deficits in gait speed. If the test could use a 3.05-m walkway and still maintain validity, it could be much more useful to home health therapists.

Conclusion

The FGA can be used reliably as a clinical test of postural stability during walking in community-dwelling older adults. These results demonstrate that decreased performance is expected in typical older adults. Clinical research is needed to determine whether it is reasonable to use age-referenced data from subjects without impairments as target endpoints for patients with known dysfunction. Future research should include testing subjects with a history of falling to determine the FGA's ability to predict risk for falling, testing the usefulness of all items on the test, and addressing whether a 3.05-m (10-ft) walkway is adequate for determining gait speed.

Dr Walker provided concept/idea/research design, project management, and facilities/equipment. Dr Austin, Dr Banke, Dr Foxx, Dr Gaetano, Dr Gardner, Dr McElhiney, Dr Morris, and Dr Penn provided data collection and subjects. All authors provided writing and data analysis.

This study was approved by the Old Dominion University Institutional Review Board.

This article was submitted November 10, 2006, and was accepted July 12, 2007.

DOI: 10.2522/ptj.20060344

References

1 Wrisley DM, Marchetti GF, Kuharsky DK, Whitney SL. Reliability, internal consistency, and validity of data obtained with the Functional Gait Assessment. *Phys Ther.* 2004;84:906–918.

2 Shumway-Cook A, Woollacott MH. *Motor Control: Theory and Practical Applications.* Baltimore, Md: Lippincott Williams & Wilkins; 1995.

3 Whitney SL, Wrisley DM, Furman JM. Concurrent validity of the Berg Balance Scale and the Dynamic Gait Index in people with vestibular dysfunction. *Physiother Res Int.* 2003;8:178–189.

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

5 Whitney SL, Hudak MK, Marchetti GF. The Dynamic Gait Index relates to self-reported fall history in individuals with vestibular dysfunction. *J Vestib Res.* 2000;10:99–105.

6 Hall CD, Schubert MC, Herdman SJ. Prediction of fall risk reduction as measured by Dynamic Gait Index in individuals with unilateral vestibular hypofunction. *Otol Neurotol.* 2004;25:746–751.

7 Hall CD, Herdman SJ. Reliability of clinical measures used to assess patients with peripheral vestibular disorders. *Journal of Neurologic Physical Therapy*. 2006;30(2):74–81.

8 Wrisley DM, Walker ML, Echternach JL, Strasnick B. Reliability of the dynamic gait index in people with vestibular disorders. *Arch Phys Med Rehabil*. 2003;84:1528–1533.

9 McConvey J, Bennett SE. Reliability of the dynamic gait index in individuals with multiple sclerosis. *Arch Phys Med Rehabil*. 2005;86:130–133.

10 Galgon AK, Berg balance score, Dynamic Gait Index, and walking speed as outcome measures for individuals after stroke in an outpatient physical therapy setting [abstract]. *Neurol Rep*. 2003;26:196–197.

11 Gill-Body KM, Murphy MP, Sullivan PE. Concurrent validity of the Dynamic Gait Index in community dwelling elderly [abstract]. *Neurol Rep*. 2003;26:196–197.

12 Vereek L, Truijen S, Wuyts FL, Van de Heyning PH. The Dizziness Handicap Inventory and its relationship with functional balance performance. *Otol Neurol*. 2006;28:87–93.

13 Legters K, Whitney SL, Porter R, Buczek F. The relationship between the activities-specific balance confidence scale and the dynamic gait index in peripheral vestibular dysfunction. *Physiother Res Int.* 2005;10:10–22.

14 Chiu Y, Fritz SL, Light KE, Velozo CA. Use of item response analysis to investigate measurement properties and clinical validity of data for the dynamic gait index. *Phys Ther.* 2006;86:778–787.

15 Whitney SL, Wrisley DM, Brown KE, Furman JM. Is perception of handicap related to functional performance in persons with vestibular dysfunction? *Otol Neurol*. 2004;25;139–145.

16 Lusardi MM, Pellecchia GL, Schulman M. Functional performance in community living older adults. *Journal of Geriatric Physical Therapy*. 2003;26(3):14–22.

17 Fregly AR, Smith MJ, Graybiel A. Revised normative standards of performance of men on a quantitative ataxia test battery. *Acta Otolaryngol*. 1973;75:10–16.

18 Whitney SL, Wrisley DM, Marchetti GF, Furman JM. The effect of age on vestibular rehabilitation outcomes. *Laryngoscope*. 2002;112:1785–1790.

19 Winter DA, Patla AE, Frank JS, et al. Biomechanical walking pattern changes in the fit and healthy elderly. *Phys Ther*. 1990;70:340–347.

20 Wolfson L, Whipple R, Derby CA, et al. A dynamic posturography study of balance in healthy elderly. *Neurology*. 1992;42: 2069–2075.

21 Murphy MA, Olson SL, Protas EJ, Overby AR. Screening for falls in community dwelling elderly. *Journal of Aging and Physical Activity*. 2003;11:66–80.

22 Woollacott MH, Shumway-Cook A, Nashner L. Aging and posture control: changes in sensory organization and muscular coordination. *Intl J Aging Hum Dev.* 1986; 23:97–114.

23 Richardson JK, Thies SB, DeMott TK, Ashton-Miller JA. Gait analysis in a challenging environment differentiates between fallers and nonfallers among older patients with peripheral neuropathy. *Arch Phys Med Rehabil.* 2005;86:1539–1544.

24 Shrout PE, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420–428.

25 Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 2nd ed. Upper Saddle River, NJ: Prentice Hall Health; 2000:565.

26 Kraemer HC, Periyakoil VS, Noda A. Tutorial in biostatistics: kappa coefficients in medical research. *Stat Med*. 2002;21: 2109–2129.