# Psychometric Comparisons of the Timed Up and Go, One-Leg Stand, Functional Reach, and Tinetti Balance Measures in Community-Dwelling Older People

*Mau-Roung Lin, PhD,* Hei-Fen Hwang, MS,† Ming-Hsia Hu, PhD,‡ Hong-Dar Isaac Wu, PhD,§ Yi-Wei Wang, BS,* and Fu-Chao Huang, MD‖*

**OBJECTIVES:** To compare the practicality, reliability, validity, and responsiveness of the timed up and go (TUG), one-leg stand (OLS), functional reach (FR), and Tinetti balance (TB) performance measures in people aged 65 and older.

**DESIGN:** A prospective study.

**SETTING:** Shin-Sher Township of Taichung County, west-central Taiwan.

**PARTICIPANTS:** Twelve hundred community-dwelling older people.

**MEASUREMENTS:** During an initial assessment at their residences, participants were interviewed for demographics, cognition, fall history, use of a walking aid, and activities of daily living (ADLs), in addition to completing the four balance tests. Falls were ascertained by telephone every 3 months for a 1-year follow-up; the four balance measures and ADLs were also reassessed at the end of the follow-up year.

**RESULTS:** Of the four balance measures, the OLS had the lowest participation rate, and participation of people who were cognitively impaired had fallen in the previous year, used a walking aid, or suffered from an ADL disability was lower than for their counterparts. The time to complete the tests ranged from 58 seconds for OLS, to 160 seconds for the TB. All four balance measures exhibited excellent test-retest reliability and discriminant validity but poor responsiveness to fall status. The TB showed better discriminant, convergent, and predictive validities and responsiveness to ADL changes than the other three tests.

**CONCLUSION:** According to psychometric properties, the most suitable performance measure for evaluating balance in community-dwelling older people was the TB, followed by the TUG. **J Am Geriatr Soc 52:1343–1348, 2004.**

**Key words: balance; community; older people; psychometrics; Taiwan**

From the *Institute of Injury Prevention and Control, College of Public Health and Nutrition, Taipei Medical University, Taipei, Taiwan, Republic of China; †Department of Nursing, National Taipei College of Nursing, Taipei, Taiwan, Republic of China; ‡School and Graduate Institute of Physical Therapy, College of Medicine, National Taiwan University, Taipei, Taiwan, Republic of China; §School of Public Health, China Medical University, Taichung, Taiwan, Republic of China; and ‖Department of Neurosurgery, Mackay Memorial Hospital, Taipei, Taiwan, Republic of China.

Funding received from the Bureau of Health Promotion, Department of Health (DOH 92-PH-1103) and the National Science Council (NSC 91–2320-B-038–011), Taiwan, Republic of China.

Address correspondence to Dr. Mau-Roung Lin, Institute of Injury Prevention and Control, Taipei Medical University, 250 Wu-Hsing St., Taipei 110, Taiwan, ROC. E-mail: mrlin@tmu.edu.tw,

Balance is required for maintaining a static posture, stabilizing dynamic movements, performing daily activities, and moving around in the community.[1–3] Chronic diseases and aging can saliently affect the balance ability of older people;[4] therefore, an assessment of balance ability is valuable in predicting and preventing falls and a reduction in independent living in older people.[5–7] Assessments of balance ability can be categorized into the three types: self-reported, laboratory, and physical performance measures. Performance measures are more reliable than self-reports and can help contravene memory errors from aging, reduce the effect of cognition and hearing impairment on a person's ability to perceive or answer a questionnaire, and show excellent validity in predicting falls and physical function in older people.[8,9] Furthermore, unlike laboratory tests (e.g., dynamic posturography), performance tests do not require sophisticated and costly equipment[10] and can be easily administered to larger community-dwelling elderly populations.

Several performance balance measures, such as the timed up and go (TUG),[11] one-leg stand (OLS),[12] functional reach (FR),[13] Tinetti balance (TB),[14] and Berg balance scale,[15] are available for evaluating community-dwelling older people. However, it is time-consuming to use all of these measures for each individual, and each of them may not be appropriate for all elderly populations. For instance, some investigators have suggested that the TUG test is more appropriate for older people who are frailer or who use walking aids, whereas the OLS is more suitable for healthy

JAGS 52:1343–1348, 2004

0002-8614/04/$15.00

older people.[16] Nevertheless, these suggestions were based on personal experiences or were not justified by sufficient empirical evidence.

To help geriatric professionals and providers better understand the relative advantages and disadvantages of performance balance measures in older people, this study compared the psychometric properties of four performance balance measures in a community-dwelling elderly population in Taiwan.

## METHODS

### Study Subjects

Shin-Sher Township, located in Taichung County in west-central Taiwan, is a rural area with 11.9% of the population aged 65 and older in 1999, compared with 8.6% for Taiwan as a whole.[17] Of 13 villages in the township, six with the largest elderly populations were selected. Based on records in the Shin-Sher Household Registration Office, where demographic information is collated and supplied and personal status and relations are officially recognized, 2,072 eligible people aged 65 and older in the six villages were identified, with information on name, address, birth date, sex, and education.

During a 2-week period designated for the initial assessment, 1,200 eligible subjects agreed to participate in the study. Of the 872 subjects who did not participate, 24 had died, 59 were hospitalized or bed-ridden, 252 had moved out of the area, 323 were not at home during the assessment period, and 214 declined to participate. Participants had similar distributions of sex and educational level as nonparticipants but tended to be younger. Participants had a similar age distribution to that of older people in Taichung County or in Taiwan, but there were more men, and they had lower educational levels.

### Procedures

Subjects were first interviewed at their own residence to collect information on their age, sex, education, marital status, cognition, fall history in the previous year, use of walking aids, and disability in activities of daily living (ADLs). Interview procedures and interviewer attitudes were standardized through participation in a 4-hour training course. Cognitive status was assessed using the Mini-Mental State Examination (MMSE);[18,19] MMSE scores were categorized into three levels (0–17, 18–23, and 24–30, indicating severe, mild, and no impairment, respectively).[20] The level of ADL disability was assessed using the Older Adults Resources and Services (OARS) ADL scale,[21] consisting of seven physical ADLs (self-feeding, self-dressing, grooming, walking, getting in/out of bed, bathing, and controlling the bladder), and seven instrumental ADLs (using a telephone, transporting oneself, shopping, preparing meals, doing housework, taking medication, and managing one's money). These items were scored 1 (inability) or 0 (independence). The score for the OARS ADL scale ranges from 0 to 14 points, with a higher score indicating greater disability.

After completing the interviews, subjects' performances on four balance and two mobility tests were examined. The TUG, OLS, FR, and TB were conducted in sequence. Additionally, two mobility tests, the Tinetti gait and walking speed, were assessed for convergent validity of the four balance measures.

The occurrence of falling and changes (decline or improvement) in ADL score over a 1-year follow-up period were used as indicators to assess the predictive validity and responsiveness of the balance measures. Each participant was asked to report by postcard when a fall occurred; participants were also called every 3 months to ascertain whether they had had a fall over the previous 3-month period. Of those participants who had no history of falls, 91 experienced a fall during the follow-up period. Furthermore, the four balance measures and the OARS ADL scale were subsequently reassessed at the end of the follow-up year. Of 798 subjects who completed the follow-up assessment, 252 had a decline and 128 had an improvement in ADLs. Eighty subjects who had died or were hospitalized or bed-ridden at follow-up were considered to have experienced a decline in ADL.

### Balance and Mobility Performance Measures

In the TUG test, subjects were asked to stand up from a standard chair with a seat height of between 40 and 50 cm, walk a 3-m distance at a normal pace, turn, walk back to the chair, and sit down. Timing measured in seconds began at the word "go" and ended when the subject's back touched the backrest of the chair, with a shorter time taken indicating better balance ability.

In the OLS test, subjects were instructed to start in a position with a comfortable base of support, with eyes open and arms by the side of the trunk and then stand unassisted on any one leg. The OLS was timed in seconds from the time one foot was lifted from the floor to when it touched the ground or the standing leg, with a longer time indicating better balance ability.

In the FR test, each subject was positioned next to the wall with one arm raised 90° with the fingers extended, and a yardstick was mounted on the wall at shoulder height. The distance in centimeters that a subject was able to reach forward from an initial upright posture to the maximal anterior leaning posture without moving or lifting the feet was measured by visual observation of the position of the third finger tip against the mounted yardstick. The distances of two trials were averaged as the FR score, with a greater distance indicating better balance ability.

The TB test is one part of the performance-oriented assessment of mobility problems and consists of 13 maneuvers such as sitting balance, sit to stand, immediate standing balance (first 3–5 seconds), standing balance, balance with eyes closed, turning 360°, nudging the sternum, turning the neck, unilateral stance, extending the back, bending down and picking up an object, and sitting down.[14] Each maneuver was graded as two points (normal), one point (adaptive), or zero points (abnormal). The TB score ranges from 0 to 26, with a higher score indicating better balance ability.

The Tinetti gait test is the other part of the performance-oriented assessment of mobility problems and consists of the nine components of initiation of gait, step height and length, step symmetry and continuity, path deviation, trunk stability, walking stance, and turning while walking.[14] Each component was scored as 1 (normal) or 0 (abnormal). The score of the Tinetti gait ranges from 0 to 9, with a higher score indicating better functional mobility.

Each subject was asked to walk at a normal pace back and forth on a 3-m walkway for 60 seconds. Walking speed was obtained by dividing the walking distance in meters by 60 seconds.

### Practicality

The length of time required to complete each balance test (including its instruction time) and the proportions of subjects who refused and were unable to perform a balance test across groups with respect to cognitive status, fall history in the previous year, use of a walking aid, and ADL disability were used to evaluate respondent burden and difficulty of completion, respectively. For those unable to perform a balance test, a minimum possible value of the test was imputed. A balance measure completed in a shorter time length with a lower nonparticipation rate would supposedly be considered more practical.

### Test-Retest Reliability

To assess the original performance of the four balance tests, a sample of 60 subjects randomly stratified by village was selected. Within 2 weeks, half of the sample was retested for estimating intrarater reliability and the other for interrater reliability. Intraclass correlation coefficients ($r$)[22] were calculated for both types of test-retest reliability.

### Discriminant Validity

For each balance measure, the known-groups validity[23] in distinguishing age groups (65–74, $\geq$75), fall history in the previous year (yes, no), use of a walking aid (yes, no), and number of disabilities of ADL (0, 1, or more) was evaluated using the Student $t$ test. Lower performance scores were expected for subjects who were older, had had a fall in the previous year, used a walking aid, and suffered from an ADL disability. Furthermore, to facilitate quantitative comparisons of the discriminant abilities of the balance measures, the area under the receiver operating characteristic curve (i.e., a plot of the true-positive rate against the false-positive rate) was computed.[24] The area under the curve (AUC) ranges from 0.5 for a noninformative instrument to 1.0 for perfect discrimination.

### Convergent Validity

Measures that conceptually converge should be strongly correlated, whereas those measures with less in common should show weaker correlations. Therefore, it was hypothesized that moderate or strong correlations ($r = 0.4$) between the four balance measures, as well as functional mobility[25] (evaluated using the Tinetti gait and walking speed) and functional activities[26] (evaluated using the OARS ADL scale). Furthermore, factor analysis was applied to further understand whether these measures were in the same dimension. The three criteria of factor eigenvalues ($>1$), proportion of total variance ($>5\%$), and Scree test[27] can indicate how many common factors should be adequate to represent these balance measures. The measurements of these tests were expected to converge into one single common factor.

### Predictive Validity

The ability of these balance measures to predict the occurrences of a fall and the decline and improvement in ADLs over the study period was evaluated using the logistic regression model. A logistic regression model was considered to support the predictive validity of predicting future events if it was statistically significant after adding the balance measure under study. Furthermore, the AUC for each balance measure was also computed, with a larger AUC indicating better predictive ability.

### Responsiveness

The responsive statistic for each balance measure was calculated by evaluating the mean change in scores for that measure over the study period of the fallers who had no previous history of falls divided by the standard deviation of the score changes of the subjects who had never fallen.[28] Declines and improvements in ADLs over the period were also used as external indicators to evaluate the responsiveness of these balance measures. An effect size of 0.2 to 0.5 was considered clinically meaningful as small effects, 0.5 to 0.8 as moderate effects, and greater than 0.8 as large effects.[29]

SAS version 6.12 (SAS Institute, Cary, NC) was used for all statistical analyses.

## RESULTS

Of 1,200 subjects, the mean age was 73.4; 709 (59%) were men, 387 (37%) had no formal education, and 766 (66%) lived with a spouse. Two hundred twenty (18%) had severe cognitive impairment, 423 (37.8%) had mild cognitive impairment, 127 (11%) reported having experienced at least one fall in the previous year, 128 (11%) needed a walking aid, and 435 (36%) suffered one or more ADL disabilities. The mean values of the TUG, OLS, FR, and TB measures were 13.3 seconds, 8.3 seconds, 14.8 centimeters, and 20.3 points, respectively. Furthermore, the intraclass $r$ for intra- and interrater reliability within 2 weeks were excellent for all balance measures, with a range of 0.93 to 0.99, and they did not differ among the three levels of cognitive status.

The average time to complete the TUG, OLS, FR, and TB tests (including instruction time) was 86, 58, 76, and 160 seconds, respectively. Of the four balance measures, as shown in Table 1, the TUG and TB had lower refusal rates than did the OLS and FR, and a substantial proportion of subjects were unable to perform the OLS and FR tests. Nonparticipation was higher in subjects who were cognitively impaired, had fallen in the previous year, used a walking aid, or had an ADL disability.

As shown in Table 2, the discriminant ability of the four balance measures was excellent. Specifically, subjects who were older, had experienced a fall in the previous year, used a walking aid, and suffered more ADL disabilities required a longer time to complete the TUG, stood a shorter time on the OLS, reached a shorter distance for the FR, and obtained lower scores on the TB. Furthermore, the TB exhibited the largest AUC (i.e., the largest discrimination power) for the four characteristics (age, falls in the past year, use of a walking aid, and ADL disability).

As for convergent validity, the TB scores were moderately or strongly correlated with the TUG ($r = -0.55$), FR

**Table 1. Rates of Refusal and Inability to Perform a Balance Test with Respect to Cognitive Impairment, Fall History, Use of a Walking Aid, and Activity of Daily Living (ADL) Disability**

| | Timed Up and Go | | One-Leg Standing | | Functional Reach | | Tinetti Balance | |
|---|---|---|---|---|---|---|---|---|
| Characteristic | R rate | I rate | R rate | I rate | R rate | I rate | R rate | I rate |
| Cognitive impairment | | | | | | | | |
| Severe | 5.9 | 0 | 6.8 | 30.5 | 5.5 | 11.4 | 5.9 | 0 |
| Mild | 0.2 | 0 | 3.3 | 10.2 | 1.4 | 2.1 | 1.9 | 0 |
| None | 0.2 | 0 | 1.6 | 5.4 | 0.7 | 1.1 | 2.0 | 0 |
| Falls in the past year | | | | | | | | |
| Yes | 5.5 | 0 | 7.9 | 17.3 | 5.5 | 5.5 | 5.5 | 0 |
| No | 0.7 | 0 | 2.6 | 11.0 | 3.1 | 0.7 | 2.3 | 0 |
| Use of a walking aid | | | | | | | | |
| Yes | 4.7 | 0 | 10.9 | 46.9 | 6.3 | 19.5 | 5.5 | 0 |
| No | 0.7 | 0 | 2.1 | 7.6 | 1.1 | 1.4 | 2.4 | 0 |
| ADL disability | | | | | | | | |
| Yes | 3.0 | 0 | 6.0 | 22.3 | 4.4 | 7.4 | 4.6 | 0 |
| No | 0.3 | 0 | 1.6 | 5.6 | 0.4 | 1.0 | 1.6 | 0 |

R = refusal; I = inability.

($r = 0.48$), Tinetti gait ($r = 0.81$), walking speed ($r = -0.54$), and ADL scale ($r = 0.60$), whereas the TUG was also moderately or strongly correlated with the Tinetti gait ($r = -0.53$), walking speed ($r = 0.66$), and ADL scale ($r = -0.45$). Conversely, the OLS was not moderately or strongly correlated with any other measures. In the factor analysis, one single common factor that explained 51% of the total variance in the four balance measures was selected, indicating that the four measures were unidimensional. The factor loadings of the four balance measures onto the factor were $-0.74$, 0.54, 0.72, and 0.83, respectively.

As shown in Table 3, a longer time to perform the TUG and lower scores on the TB significantly predicted the occurrence of falling and ADL decline and improvement over the follow-up period. A shorter time for the OLS and a shorter distance for the FR significantly predicted a decline in ADL, but they did not significantly predict the occurrence of falling or ADL improvement. When including cognitive status, the prediction of each balance measure for the occurrences of falling and ADL decline and improvement did not change (data not shown). Furthermore, of the four balance measures, the TUG had the largest AUC in predicting the occurrence of falling, whereas the TB had the largest AUC in predicting ADL decline and improvement.

The effect sizes of the responsiveness to falls and ADL decline and improvement were 0.12, 0.42, and 0.05 for the

**Table 2. Discriminant Ability of Each Balance Measure with Respect to Age, Fall History, Use of a Walking Aid, and Activity of Daily Living (ADL) Disability**

| Characteristic | Timed Up and Go (seconds) | One-Leg Stand (seconds) | Functional Reach (cm) | Tinetti Balance (points) |
|---|---|---|---|---|
| Age, mean ± SD | | | | |
| 65–74 | 12.4 ± 7.4 | 9.7 ± 13.7 | 15.8 ± 8.1 | 21.4 ± 4.6 |
| ≥75 | 15.2 ± 10.3 | 5.6 ± 7.4 | 12.6 ± 8.8 | 18.0 ± 6.8 |
| AUC | 0.592 | 0.636 | 0.617 | 0.644 |
| Falls in the past year, mean ± SD | | | | |
| Yes | 16.8 ± 12.9 | 4.7 ± 7.6 | 15.1 ± 8.5 | 16.8 ± 7.0 |
| No | 12.9 ± 7.8 | 8.7 ± 12.4 | 11.5 ± 7.7 | 20.7 ± 5.3 |
| AUC | 0.614 | 0.640 | 0.623 | 0.672 |
| Use of a walking aid, mean ± SD | | | | |
| Yes | 23.6 ± 16.7 | 2.7 ± 5.4 | 8.5 ± 8.5 | 11.6 ± 6.9 |
| No | 12.0 ± 5.7 | 9.0 ± 12.6 | 15.5 ± 8.1 | 21.3 ± 4.5 |
| AUC | 0.778 | 0.794 | 0.744 | 0.862 |
| ADL disability, mean ± SD | | | | |
| Yes | 16.5 ± 12.0 | 5.8 ± 10.4 | 11.7 ± 8.7 | 17.1 ± 6.8 |
| No | 11.5 ± 5.0 | 9.7 ± 12.7 | 16.4 ± 7.9 | 22.0 ± 3.9 |
| AUC | 0.648 | 0.665 | 0.670 | 0.727 |

*Note:* $P = .000$, except $P = .001$ for falls in the previous year for One-Leg Standing (Student $t$ test).
AUC = area under the receiver operating characteristic curve.

**Table 3. Predictive Ability of Each Balance Measure for the Occurrence of Falls and ADL Decline and Improvement Over a 1-Year Follow-Up Period, According to the Logistic Regression Model and Area Under the Receiver Operating Characteristic Curve (AUC)**

| Predicted Outcome | Timed Up and Go (seconds) | One-Leg Standing (seconds) | Functional Reach (cm) | Tinetti Balance (points) |
|---|---|---|---|---|
| | Odds Ratio (95% Confidence Interval) AUC | | | |
| Falls | 1.02 (1.01–1.03) 0.610 | 0.99 (0.98–1.01) 0.527 | 1.00 (0.98–1.02) 0.509 | 0.96 (0.93–0.99) 0.559 |
| ADL decline | 1.03 (1.01–1.04) 0.560 | 0.98 (0.97–0.99) 0.577 | 0.96 (0.94–0.97) 0.610 | 0.92 (0.90–0.94) 0.629 |
| ADL improvement | 1.03 (1.01–1.05) 0.543 | 1.00 (0.99–1.01) 0.523 | 0.99 (0.97–1.02) 0.531 | 0.94 (0.91–0.97) 0.613 |

ADL = activity of daily living.

TUG; 0.10, 0.19, and 0.00 for the OLS; 0.04, 0.11, and 0.38 for the FR; and 0.19, 0.94, and 0.39 for the TB, respectively.

## DISCUSSION

In this study, all four performance balance measures showed excellent test-retest reliability, and discriminant, convergent, and predictive validities of the TB and TUG were confirmed. Although the TB and TUG needed longer to complete, they had higher participation rates than the other two measures. The four measures showed poor responses to falls, although the TB showed a large response to ADL decline but a smaller response to ADL improvement. Therefore, according to these psychometric properties, the most suitable performance measure for evaluating balance in community-dwelling older people was the TB, followed by the TUG.

The proportion of subjects unable to complete the balance test supports the TUG and TB tests being more appropriate for older people who are frailer or who use walking aids, whereas the OLS and FR are more suitable for healthy older people.[16] Moreover, the refusal rate in the balance tests may somewhat reflect that older people can accept the balance tests without fear of falling. Despite the refusal rate for these balance tests (except the OLS) remaining acceptable ($<6\%$), caution should be applied when these study results are generalized to persons who are more likely to have cognitive impairment, have experienced falls, use a walking aid, or have an ADL disability, because their nonparticipation rate may be drastically higher. Furthermore, even though the minimum possible value could be imputed to those who were unable to complete the OLS or FR to maintain qualitative information, having a substantial number of those participants would reduce the power of discrimination and responsiveness. Accordingly, modifications to these tests may be required for such populations.[30]

Several factors may explain the poor performance in responsiveness to falls for these balance measures. First, the severity of falls in community-dwelling elderly, for which most cases do not require medical care, would be expected to be milder than for those treated in the hospital, and the effect sizes of the balance measures for milder falls should be smaller than those in the hospital. The fact that the effect sizes of the balance measures somewhat increased (0.25 for the TUG, 0.05 for the OLS, 0.13 for the FR, and 0.21 for the TB) when including only falls that required medical treatment partly supports this reasoning. Second, clinically important changes are less likely to be detected in people who initially have poorer health than in those in better health;[31] thus, changes in these balance measures for older people might be smaller than for younger populations. Third, because the interval between the occurrence of a fall and the follow-up assessment could differ from 1 day to 1 year, the responsiveness of these balance measures, particularly for fallers with longer intervals, may have been underestimated, even though the time lag of assessment did not differ between these measures. Finally, the presence of a substantial number of subjects who were unable to perform the OLS and FR measures (i.e., the floor effect) may also lead to their having a poorer responsiveness to falls.

This study had two limitations. First, the participants were rural, community-dwelling people aged 65 and older, thus these findings are not generalizable without certain qualifications to urban and institutionalized people. For instance, the practicality, discriminant validity, and responsiveness might be reduced because the floor values of the balance measures in those people would be increased to a larger extent. Second, unlike in clinics or laboratories, the circumstances for assessment (e.g., floor conditions and chair heights) varied and were not standardized in subjects' residences, which may have influenced the subjects' performance in the balance tests, even though is has been reported that the TUG can be performed using chairs available in the elderly person's residence.[32]

## REFERENCES


1. Manchester D, Woollacott M, Zederbauer-Hylton N et al. Visual, vestibular, and somatosensory contributions to balance control in the older adult. J Gerontol 1989;44:M118–M127.
2. Ledin T, Kronhed AC, Möller M et al. Effects of balance training in elderly evaluated by clinical tests and dynamic posturography. J Vestibular Res 1990/91;1:129–138.
3. Newton RA. Balance screening of an inner city older adult population. Arch Phys Med Rehabil 1997;78:587–591.
4. Tinetti ME, Williams TF, Mayewski R. A fall risk index for elderly patients based on number of chronic disabilities. Am J Med 1986;80:429–434.
5. Rubenstein LZ, Robbins AS. Falls in the elderly: A clinical perspective. Geriatrics 1984;39:67–78.
6. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. N Engl J Med 1988;319:1701–1707.
7. Cho CY, Alessi CA, Cho M et al. The association between chronic illness and functional change among participants in a comprehensive geriatric assessment program. J Am Geriatr Soc 1998;46:677–682.
8. Guralnik JM, Branch LG, Cummings SR et al. Physical performance measures in aging research. J Gerontol 1989;44:M141–M146.

9. Reuben DB, Valle LA, Hays RD et al. Measuring physical function in community-dwelling older persons: A comparison of self-administered, interviewer-administered, and performance-based measures. J Am Geriatr Soc 1995; 43:17–23.
10. Berg KO, Maki BE, Williams JI et al. Clinical and laboratory measures of postural balance in an elderly population. Arch Phys Med Rehabil 1992;73:1073–1080.
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. Vellas BJ, Wayne SJ, Romero L et al. One-leg balance is an important predictor of injurious falls in older persons. J Am Geriatr Soc 1997;45:735–738.
13. Duncan PW, Weiner DK, Chandler J et al. Functional reach: A new clinical measure of balance. J Gerontol 1990;45:M192–M197.
14. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc 1986;34:119–126.
15. Berg KO, Wood-Dauphinee SL, Williams JI et al. Measuring balance in the elderly: Preliminary development of an instrument. Physiother Can 1989; 41:304–311.
16. Lafont C, Costes-Salon MC, Dupui P et al. Reply to 'The timed "up & go" test is a useful predictor of falls in community-dwelling older people. J Am Geriatr Soc 1998;46:929.
17. Vital Statistics in 1999. Taipei: Department of Health, Executive Yuan, Taiwan, Republic of China, 2000.
18. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.
19. Guo NW, Liu HC, Wong PF et al. Chinese version and norms of the Mini-Mental State Examination. J Rehabil Med Assoc (ROC) 1988;16:52–59.
20. George LK, Landerman R, Blazer DG et al. Cognitive impairment. In: Robins LN, Regier DA, eds. Psychometric Disorders in America: The Epidemiologic Catchment Area Study. New York: Free Press, 1991, pp 291–327.
21. Fillenbaum GG, Smyer MA. The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire. J Gerontol 1981;36:428–434.
22. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979;86:420–428.
23. Kerlinger FN Foundations of Behavioral Research. New York: Holt, Rinehart, and Winston, 1973.
24. Sackett DL, Haynes RB, Guyatt GH et al. Clinical Epidemiology—A Basic Science for Clinical Medicine, 2nd Ed. Boston: Little Brown, 1985.
25. Shinkai S, Watanabe S, Kumagai S et al. Walking speed as a good predictor for the onset of functional dependence in Japanese rural community population. Age Aging 2000;29:441–446.
26. Salbach NM, Mayo NE, Higgins J et al. Responsiveness and predictability of gait speed and other disability measures in acute stroke. Arch Phys Med Rehabil 2001;82:1204–1212.
27. Stevens J Applied Multivariate Statistics for the Social Sciences, 2nd Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1992.
28. Guyatt GH, Walter S, Norman G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40:171–178.
29. Cohen J Statistical Power Analysis for the Behavioral Sciences, 2nd Ed. New York: Lawrence Erlbaum Associates, 1998.
30. Tappen RM, Roach KE, Buchner D et al. Reliability of physical performance measures in nursing home residents with Alzheimer's disease. J Gerontol A Biol Sci Med Sci 1997;52A:M52–M55.
31. Baker DW, Hays RD, Brook RH. Understanding changes in health status: Is the floor phenomenon merely the last step of the staircase? Med Care 1997;35:1–15.
32. Eekhof JAH, De Bock GH, Schaapveld K et al. Functional mobility assessment at home—timed up and go test using three different chairs. Can Fam Physician 2001;47:1205–1207.