“Timed Up and Go” Test and Bone Mineral Density Measurement for Fracture Prediction

Kun Zhu, PhD; Amanda Devine, PhD; Joshua R. Lewis, PhD; Satvinder S. Dhaliwal, PhD; Richard L. Prince, MD

Background: Two major factors associated with skeletal fracture in older persons are intrinsic bone strength and risk of falling. This study examined the role of Timed Up and Go (TUG) test performance, a validated predictor of falling, and hip areal bone mineral density (BMD), a validated predictor of bone strength in fracture prediction in a 10-year longitudinal study.

Methods: The study participants were 1126 women (mean [SD] age at baseline, 75.0 [2.6] years) living in Perth, Western Australia. Assessments included TUG test at baseline and dual-energy x-ray absorptiometry total hip areal BMD measurement at year 1. Incident clinical osteoporotic fracture over 10 years was confirmed by radiographic records. Complete incident hip fracture data were obtained from a hospital morbidity database.

Results: One-third (32.7%) of participants had slow TUG test performance (>10.2 seconds), and 54.2% of participants had low hip areal BMD (T-score of less than –1). Relative to risks among participants having normal TUG test performance and normal BMD, risks of nonvertebral fracture and hip fracture were significantly higher among participants who had slow TUG test performance and normal hip BMD (nonvertebral fracture hazard ratio [HR], 1.84; hip fracture HR, 2.48) or both slow TUG test performance and low hip BMD (nonvertebral fracture HR, 2.51; hip fracture HR, 4.68). For nonvertebral fracture and hip fracture, the population-attributable risks of slow TUG test performance with normal hip BMD were 19.3% and 32.3%, of normal TUG test performance with low hip BMD were 31.3% and 50.3%, and of both slow TUG test performance and low hip BMD were 30.1% and 55.9%, respectively.

Conclusion: TUG test performance is an independent risk factor for incident nonvertebral fracture and a feasible inexpensive physical performance assessment for use in clinical practice to screen patients with increased risk of fracture.

Arch Intern Med. 2011;171(18):1655-1661

Fracture represents a major public health problem and is a leading cause of morbidity, mortality, and hospitalization among older persons. Since the development of dual-energy x-ray absorptiometry (DXA) measurement of areal bone mineral density (BMD) in the 1980s, great progress has been made in our understanding of the important role of bone structure in resistance to fracture. However, less attention has been paid to the role of risk factors for falling, which in addition to bone mass are important determinants of the occurrence of most appendicular skeletal fracture.

The Timed Up and Go (TUG) test, in which individuals are timed when rising from a chair, walking 3 m, and turning to return to sit on the chair, is a commonly used method of assessing functional mobility among older adults in geriatric clinics to evaluate physical performance. In community-dwelling adults older than 65 years, the TUG test has been shown to be a sensitive and specific measure for identifying older persons who are prone to fall. A meta-analysis that summarized the findings of 21 studies reported a reference TUG value of 9.2 (95% confidence interval [CI], 8.2-10.2) seconds for individuals aged 70 to 79 years, and the authors suggested that patients with test results above the upper limit of the reference CI could be regarded as having performance worse than the average. To our knowledge, the association of slow TUG test performance and incident fracture risk among older women has not been evaluated.

See Invited Commentary at end of article

We hypothesized that slow TUG test performance would predict fracture in older women, independent of BMD and other risk factors. The objective of this study was to examine slow TUG test performance relative to hip BMD in predicting fracture among 1126 women (mean [SD] age, 75.0 [2.6] years) in a 10-year longitudinal study.

Author Affiliations:
Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, School of Medicine and Pharmacology, University of Western Australia (Drs Zhu, Lewis, and Prince), School of Exercise, Biomedical and Health Science, Edith Cowan University (Dr Devine), and School of Public Health, Curtin University (Dr Dhaliwal), Perth, Western Australia.
BONE MEASUREMENTS

Because of the restricted availability of human resources, total hip areal BMD of participants was measured by DXA using a fan beam densitometer (Hologic Acclaim 4500A; Hologic Inc, Waltham, Massachusetts) at year 1. The total hip BMD coefficient of variation was 1.2% in our laboratory. The total hip BMD T-score was determined using the US Third National Health and Nutrition Examination Survey reference database. The World Health Organization defined low bone density as a BMD T-score between −1 and −2.5 and osteoporosis as a T-score of −2.5 or less. Therefore, we categorized participants with hip areal BMD T-scores of less than −1 as having low hip BMD.

TUG TEST

The TUG test, in which the patient is timed when rising from a chair, walking 3 m, and turning to return to sit on the chair, was performed at baseline. Participants were allowed to practice once and were then timed. The interobserver coefficient of variation error was 7.0% in our laboratory, as assessed among a random sample of 20 individuals.

FRACTURE ASCERTAINMENT

Prevalent fracture was determined at baseline by obtaining a fracture history from each participant, which included age at the time of fracture, fracture location, and fracture cause. A prevalent fracture was present if the fracture happened after age 50 years, occurred with minimal trauma (falling from a height of ≤1 m), and was not a fracture of the face, skull, fingers, or toes. Incident atraumatic clinical fracture and atraumatic symptomatic vertebral fracture were recorded in an adverse events diary, which was collected every 4 months during the first 5 years and every 6 months during the second 5 years. Diagnoses of clinical vertebral and nonvertebral fracture were confirmed by radiographic records. In addition, incident hip fracture data were retrieved from the Western Australia Hospital Morbidity Database for each study participant from 1998, when she entered the study, until 10 years after her baseline visit. Because this database captures coded diagnosis data pertaining to all public and private inpatient contacts in Western Australia, its use allows complete ascertainment of verified hip fracture, despite patient events such as loss to follow-up. Because hip BMD measurement was performed at year 1, fracture between year 1 and year 10 was regarded as incident fracture in the data analysis, whereas fracture during the first year was considered prevalent fracture.

OTHER ASSESSMENTS

Height and weight were measured at baseline with the participants in light clothing and without shoes. Body mass index was calculated as weight in kilograms divided by height in meters squared. Nutrient intakes were determined from a self-administered semiquantitative food frequency questionnaire. Physical activity was assessed by a questionnaire and activity levels were calculated (in kilocalories per day) using a validated method that included body weight, questions about the number of hours and type of physical activity, and energy costs of such activities.

DATA ANALYSIS

Descriptive statistics are reported as the mean (SD) for all variables unless otherwise stated. The associations of baseline TUG test performance and hip BMD at year 1 with incident fracture were examined using logistic regression and Cox proportional hazards regression, adjusting for baseline age, prevalent fracture, calcium treatment, current smoking, rheumatoid arthritis, and consumption of alcohol. Other potential covariates considered in the analyses included baseline weight, height, calcium intake, and physical activity. The proportional hazards assumption was tested for each covariate, and no violations were detected. Nine-year fracture risk was calculated using odds ratios obtained from logistic regression models.

©2011 American Medical Association. All rights reserved.
Table 1. Baseline Characteristics of 1126 Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Cohort</th>
<th>BMD T-Score of at Least −1</th>
<th>BMD T-Score of Less Than −1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=1126)</td>
<td>TUG &lt;10.2 s</td>
<td>TUG &gt;10.2 s</td>
</tr>
<tr>
<td></td>
<td>(n=343)</td>
<td>(n=173)</td>
<td>(n=415)</td>
</tr>
<tr>
<td></td>
<td>(n=195)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>75.0 (2.6)</td>
<td>74.5 (2.5)</td>
<td>75.2 (2.7)</td>
</tr>
<tr>
<td>Weight, mean (SD), kg</td>
<td>68.3 (11.9)</td>
<td>72.4 (11.2)</td>
<td>75.7 (12.8)</td>
</tr>
<tr>
<td>Height, mean (SD), cm</td>
<td>159.0 (5.8)</td>
<td>159.9 (5.6)</td>
<td>159.4 (6.1)</td>
</tr>
<tr>
<td>BMI, mean (SD)</td>
<td>27.0 (4.5)</td>
<td>28.4 (4.4)</td>
<td>29.7 (4.4)</td>
</tr>
<tr>
<td>Calcium intake, mean (SD), mg/d</td>
<td>95.0 (347)</td>
<td>994 (340)</td>
<td>949 (362)</td>
</tr>
<tr>
<td>Calcium treatment, %</td>
<td>50.4</td>
<td>48.4</td>
<td>53.2</td>
</tr>
<tr>
<td>Physical activity, median (IQR), kcal/d</td>
<td>116 (45-209)</td>
<td>138 (60-222)</td>
<td>118 (0-201)</td>
</tr>
<tr>
<td>Prevalent fracture, including incident fracture during year 1, %</td>
<td>30.3</td>
<td>26.2</td>
<td>30.1</td>
</tr>
<tr>
<td>Rheumatoid arthritis, %</td>
<td>0.6</td>
<td>0.6</td>
<td>0.0</td>
</tr>
<tr>
<td>Current smoker, %</td>
<td>0.3</td>
<td>0.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Consumption of ≥3 U/d of alcohol, %</td>
<td>3.5</td>
<td>4.4</td>
<td>2.3</td>
</tr>
<tr>
<td>Fall in the past 3 mo, %</td>
<td>10.8</td>
<td>10.8</td>
<td>13.9</td>
</tr>
<tr>
<td>TUG test performance, median (IQR), s</td>
<td>9.2 (8.0-10.8)</td>
<td>8.4 (7.6-9.3)</td>
<td>11.5 (10.8-12.8)</td>
</tr>
<tr>
<td>Hip BMD at year 1, mean (SD)</td>
<td>Measurement, mg/cm²</td>
<td>811 (125)</td>
<td>920 (81)</td>
</tr>
<tr>
<td>T-score</td>
<td>−1.1 (1.0)</td>
<td>−0.2 (0.7)</td>
<td>−0.2 (0.7)</td>
</tr>
</tbody>
</table>

Abbreviations: BMD, bone mineral density; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); IQR, interquartile range; TUG, Timed Up and Go.

a Significantly different from that of the BMD T-score of less than −1 and TUG ≤10.2 s group, P < .05 by analysis of variance with Tukey test.
b Significantly different from that of the BMD T-score of at least −1 and TUG ≤10.2 s group, P < .05 by analysis of variance with Tukey test.
c Significantly different from that of the BMD T-score of less than −1 and TUG >10.2 s group, P < .05 by analysis of variance with Tukey test.
d Significantly different from that of the BMD T-score of at least −1 and TUG >10.2 s group, P < .008 by nonparametric test.

t-Score, there were significant differences in anthropometric measurements, calcium intake, and physical activity (Table 1). Body weight and body mass index were higher in those with normal hip BMD and in those with slower TUG times. Participants having a hip BMD T-score of at least −1 and a TUG test performance of 10.2 seconds or less had significantly higher calcium intake and physical activity compared with those having a hip BMD T-score of −1 or less and a TUG test performance of 10.2 seconds or longer. There were no significant differences among the 4 groups in Table 1 in the percentages of participants who had prevalent fracture or rheumatoid arthritis, received calcium treatment during the intervention phase of the study, consumed 3 U or more of alcohol per day, were currently smoking, or had fallen in the past 3 months.

CHARACTERISTICS OF STUDY PARTICIPANTS

The mean age of 1126 study participants at baseline was 75.0 (2.6) years; 93.7% of participants were between 70 and 79 years of age, and 30.3% had prevalent fracture at year 1 (Table 1). The median baseline TUG test performance was 9.2 seconds; 368 participants (32.7%) took longer than 10.2 seconds to complete the test and were regarded as having slow TUG test performance. At 1 year, the mean hip BMD T-score was −1.1 (1.0); 534 participants (47.4%) had low hip BMD (T-score between −1 and −2.5), and 76 participants (6.7%) had osteoporosis (T-score of less than −2.5).

As expected, when participants were grouped according to baseline TUG test performance and hip BMD T-score, there were significant differences in anthropometric measurements, calcium intake, and physical activity (Table 1). Body weight and body mass index were higher in those with normal hip BMD and in those with

FRACTURE RATES AND POPULATION-ATTRIBUTABLE RISK

From year 1 to year 10, the self-reported incident fracture rates were 17.5% for nonvertebral fracture and 6.0% for clinical vertebral fracture, with 1.6% of participants having at least 1 incident fracture of each type. Based on data obtained from the Western Australia Hospital Morbidity Database, 74 participants (6.6%) had incident hip fracture.

Compared with those who had a normal TUG test performance, participants with a slow TUG test performance had significantly higher rates of incident nonvertebral fracture (21.2% vs 15.7%, P=.02) and hip fracture (9.2% vs 5.3%, P=.02) but not clinical vertebral fracture (5.7% vs 6.1%, P=.89). Participants having low hip BMD had significantly higher incidences of nonvertebral fracture (21.0% vs 13.4%), hip fracture (8.9% vs 3.9%), and clinical vertebral fracture (7.7% vs 3.9%) com-
pared with those having normal hip BMD (P < .01 for all). Slow TUG test performance increased the population-attributable risk of nonvertebral fracture (10.3%) and hip fracture (19.7%) but not clinical spine fracture. Low hip BMD substantially increased the risk of all fracture types (range, 23.6%-41.0%). Figure 2 shows the effect of baseline TUG test performance, dichotomized into normal (≤10.2 seconds) and slow (>10.2 seconds), on the relationship between the 9-year nonvertebral fracture risk and hip BMD T-score. At all hip BMD T-scores, participants with a slow TUG test performance had higher fracture risk.

The combined effects of a slow TUG test performance and a low hip BMD T-score on incidence of nonvertebral fracture and hip fracture and on population-attributable risk are summarized in Table 2. Participants with a normal TUG test performance and a normal hip BMD had significantly lower rates of nonvertebral fracture and hip fracture than those in any other category.

**TIME-TO-EVENT ANALYSIS**

To further validate the ability of the TUG test performance in relation to hip BMD to predict fracture, a time-to-event analysis with censoring for patient death or loss to follow-up was undertaken using Cox proportional hazards regression. A slow baseline TUG test performance was associated with a 54% higher risk of 9-year incident nonvertebral fractures (Figure 3) compared with a normal TUG test performance, after adjusting for baseline age, prevalent fracture, calcium treatment, current smoking, rheumatoid arthritis, alcohol consumption, and total hip BMD T-score at year 1. Similar effects were observed for hip fracture (hazard ratio [HR], 1.86; 95% CI, 1.16-2.97) but not for clinical vertebral fracture (0.92; 0.54-1.56).

In an additional time-to-event analysis, the magnitude of the combined effects of a slow TUG test performance and a low hip BMD on fracture rates were examined, adjusting for age, prevalent fracture, calcium treatment, current smoking, rheumatoid arthritis, and alcohol consumption (Figure 4). Participants who had a slow TUG test performance and a normal hip BMD had a higher HR for nonvertebral fracture compared with participants who had a normal TUG test performance and a normal hip BMD. Although the HR was smaller than that for participants who had both a slow TUG test performance and a low hip BMD, it was similar to the HR for those with a normal TUG test performance and a low hip BMD.

Hip fracture was analyzed using the same approach. Compared with those having a normal TUG test performance and a normal hip BMD, the other 3 groups had significantly higher risk (HR, 2.48; 95% CI, 1.02-6.02 for those with a slow TUG test performance and a normal hip BMD; HR, 2.91; 95% CI, 1.38-6.13 for those with a normal TUG test performance and a low hip BMD; and HR, 4.68; 95% CI, 2.14-10.22 for those with both a slow TUG test performance and a low hip BMD). For all Cox proportional hazards regression models, further analyses that included baseline weight, height, calcium intake, and physical activity as covariates had little influence on the results.

**SENSITIVITY ANALYSIS**

During the 10-year follow-up period, 195 participants (32.0%) with a low hip BMD and 50 participants (9.7%) with a normal hip BMD began taking osteoporosis medication (P < .001). In the sensitivity analyses that excluded these women, the association with nonvertebral fracture was weaker for hip BMD but not for TUG test performance (HR, 2.07; 95% CI, 1.20-3.57 for those with a slow TUG test performance and a normal hip BMD; HR, 1.59; 95% CI, 0.97-2.60 for those with a normal TUG test performance and a low hip BMD; and HR, 2.45; 95% CI, 1.40-4.28 for those with a slow TUG test performance and a low hip BMD). The HRs for hip fracture were similar to those obtained in the entire cohort (data not shown).

**NET RECLASSIFICATION IMPROVEMENT**

Table 3 gives the change in nonvertebral fracture risk category for models with hip BMD T-score and with both TUG test performance and hip BMD T-score. The net reclassification improvement was 8.1% (P = .01).

**COMMENT**

Decreased bone strength, as detected by DXA low hip BMD, is a well-recognized predictor of fracture and is a target for interventions to reduce osteoporotic fracture risk. In addition to low bone mass, other clinical risk factors are related to fracture risk, including past falls. Shorter-term studies using other physical performance tests have shown an association between physical performance and fracture risk in older persons. This study demonstrates that a slow TUG test performance is a predictor of incident nonvertebral fracture and hip frac-
ture risk in older women. The effect is independent of age, hip BMD, prevalent fracture, and lifestyle factors. The TUG test is an effective method to assess functional mobility in older adults and has high reliability. An important finding of the present study is that in women with a normal hip BMD (T-score of at least −1), a slow TUG test performance is associated with an 84% higher risk of nonvertebral fracture and with a 148% higher risk of hip fracture, after adjusting for other known risk factors. The significance of this is shown herein by the high population-attributable risks of nonvertebral fracture (19.3%) and hip fracture (32.3%) associated with a slow TUG test performance in participants with a normal hip BMD. This finding is consistent with limited data from previous cross-sectional and shorter-term studies on the association between physical performance and fracture in older persons. In a cross-sectional study of 484 women (mean age, 55.1 years), a slow TUG test performance among those who were postmenopausal was related to previous peripheral fracture. In a study of 3851 men and women older than 60 years, quadriceps strength and postural sway were independent predictors of fracture over 3 years. Handgrip strength was shown to be a predictor of fracture in a 5-year follow-up study among a cohort of healthy postmenopausal women (mean [SD] age, 59.1 [9.8] years). The Study of Osteoporotic Fractures followed up 9516 women 65 years or older for 4.1 years and found that those who were unable to rise from a chair 5 consecutive times had a 70% higher risk of hip fracture, after adjusting for calcaneal BMD and prevalent fracture. In the Osteoporotic Fractures in Men Study, comprising 5902 men 65 years or older, individuals having the worst performance on at least 3 of 5 physical performance tests had a 214% higher risk of incident hip fracture during the 5.3-year follow-up period compared with men having high performance on all examinations.

Although some previous studies have proposed other cutoffs for poor TUG test performance, the cutoff used in the present study (10.2 seconds) was derived from a recent meta-analysis that summarized the findings of 21 studies, including studies by Shumway-Cook et al and by Bischoff et al. Compared with the model using the hip BMD T-score alone, the net reclassification improvement herein was 8.1% for the model using both TUG test performance and hip BMD T-score. Therefore, if the find-

<table>
<thead>
<tr>
<th>Variable</th>
<th>BMD T-Score of at Least −1</th>
<th>BMD T-Score of Less Than −1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TUG ≤10.2 s (n=343)</td>
<td>TUG &gt;10.2 s (n=173)</td>
</tr>
<tr>
<td>Nonvertebral fracture</td>
<td>10.8</td>
<td>18.5(^a)</td>
</tr>
<tr>
<td>Incidence, %</td>
<td>(Reference)</td>
<td>(13.2-47.1)</td>
</tr>
<tr>
<td>Population-attributable risk, % (95% CI)</td>
<td>19.3 (3.5-35.7)</td>
<td>30.1 (14.6-44.9)</td>
</tr>
<tr>
<td>Hip fracture</td>
<td>2.6</td>
<td>6.4(^a)</td>
</tr>
<tr>
<td>Incidence, %</td>
<td>(Reference)</td>
<td>(0.8-61.4)</td>
</tr>
<tr>
<td>Population-attributable risk, % (95% CI)</td>
<td>32.3 (0.8-61.4)</td>
<td>50.3 (17.0-72.8)</td>
</tr>
</tbody>
</table>

Abbreviations: BMD, bone mineral density; CI, confidence interval; TUG, Timed Up and Go.

\(^a\)Significantly different from that of the BMD T-score of at least −1 and TUG ≤10.2 s group, P < .05.

\(^b\)Significantly different from that of the BMD T-score of at least −1 and TUG ≤10.2 s group, P < .01.
ings of the present study are replicated in other cohort studies of fracture, it may be concluded that fracture prediction should include assessment of both physical performance and skeletal structural risk as assessed by the TUG test performance and DXA hip BMD.

Strengths of this study are the population-based sample and the prospective design. All incident fracture during the study was confirmed by radiographic records, and complete ascertainment of verified hip fracture was obtained from the Western Australia Hospital Morbidity Database. Limitations of the study are that the participants were community-dwelling older women and that 93.7% of them were aged 70 to 79 years. Therefore, application of the findings is limited to this population. The predictive value of the TUG test performance in men and in other age groups deserves further study. The TUG test was performed only once in our study after a single practice, rendering it more easily applicable in practice than performance of the test 3 times.

In conclusion, slow TUG test performance is an independent predictor of nonvertebral fracture and hip fracture. The TUG test is a feasible inexpensive physical performance assessment for use in clinical practice to screen patients with increased risk of fracture.

Accepted for Publication: July 6, 2011.

Correspondence: Kun Zhu, PhD, Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, School of Medicine and Pharmacology, University of Western Australia, 1 Floor C Block, Hospital Avenue, Nedlands, Western Australia 6009, Australia (kzhu@meddent.uwa.edu.au).

Author Contributions: All authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Zhu, Devine, Dhaliwal, and Prince. Acquisition of data: Zhu, Devine, Lewis, and Prince. Analysis and interpretation of data: Zhu, Dhaliwal, and

Table 3. Nine-Year Nonvertebral Fracture Risk Predicted by Models With Hip BMD T-Score and With Both Hip BMD T-Score and TUG Test Performance

<table>
<thead>
<tr>
<th>Model With BMD</th>
<th>Model With Both BMD and TUG, No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;10%</td>
</tr>
<tr>
<td>With nonvertebral fracture (n = 197)</td>
<td>&lt;10%</td>
</tr>
<tr>
<td>10%-15%</td>
<td>4</td>
</tr>
<tr>
<td>&gt;15%</td>
<td>0</td>
</tr>
<tr>
<td>Without nonvertebral fracture (n = 929)</td>
<td>&lt;10%</td>
</tr>
<tr>
<td>10%-15%</td>
<td>45</td>
</tr>
<tr>
<td>&gt;15%</td>
<td>0</td>
</tr>
</tbody>
</table>

Abbreviations: BMD, bone mineral density; TUG, Timed Up and Go.

In conclusion, slow TUG test performance is an independent predictor of nonvertebral fracture and hip fracture. The TUG test is a feasible inexpensive physical performance assessment for use in clinical practice to screen patients with increased risk of fracture.
FRAX, Falls, and Fracture Prediction

Predicting the Future

These are interesting times for a Clinician concerned about the prevention of fractures. Compared with even 10 years ago, the number of effective and Food and Drug Administration–approved treatments has increased several fold. Equally important, our ability to identify those at highest risk of fracture has improved so we can better target treatment to those most likely to benefit and avoid treatment in those who are at lower risk. The recent introduction of the fracture risk assessment tool (FRAX) is particularly noteworthy in this regard.

FRAX is fracture risk assessment algorithm developed by the World Health Organization in 2008 that uses simple clinical risk factors (such as age, sex, body mass index, and previous fracture) and expected mortality to estimate the 10-year probability of hip and major osteo-

REFERENCES