An Approach to Identifying Osteopenic Women at Increased Short-term Risk of Fracture

Paul D. Miller, MD; Suna Barlas, PhD; Susan K. Brenneman, PT, PhD; Thomas A. Abbott, PhD; Ya-Ting Chen, PhD; Elizabeth Barrett-Connor, MD; Ethel S. Siris, MD

Background: Identification and management of women to reduce fractures is often limited to T scores less than −2.5, although many fractures occur with higher T scores. We developed a classification algorithm that identifies women with osteopenia (T scores of −2.5 to −1.0) who are at increased risk of fracture within 12 months of peripheral bone density testing.

Methods: A total of 57,421 postmenopausal white women with baseline peripheral T scores of −2.5 to −1.0 and 1-year information on new fractures were included. Thirty-two risk factors for fracture were entered into a classification and regression tree analysis to build an algorithm that best predicted future fracture events.

Results: A total of 1130 women had new fractures in 1 year. Previous fracture, T score at a peripheral site of −1.8 or less, self-rated poor health status, and poor mobility were identified as the most important determinants of short-term fracture. Fifty-five percent of the women were identified as being at increased fracture risk. Women with previous fracture, regardless of T score, had a risk of 4.1%, followed by 2.2% in women with T scores of −1.8 or less or with poor health status, and 1.9% for women with poor mobility. The algorithm correctly classified 74% of the women who experienced a fracture.

Conclusions: This classification tool accurately identified postmenopausal women with peripheral T scores of −2.5 to −1.0 who are at increased risk of fracture within 12 months. It can be used in clinical practice to guide assessment and treatment decisions.

Arch Intern Med. 2004;164:1113-1120

The association between bone mineral density (BMD) and fracture risk is continuous, with an approximate doubling of fracture risk for each standard deviation decline in BMD T score.1 Despite this continuous relationship, efforts at fracture risk reduction are often limited to women whose central BMD measurement has been classified as “osteoporotic,” that is, T scores of −2.5 or less based on the 1994 World Health Organization (WHO).2

For editorial comment see page 1047

CME course available at www.archinternmed.com

diagnostic classification. Postmenopausal women with a diagnosis of “osteopenia,” that is, T scores between −2.5 and −1.0 based on the WHO classification, may also be at risk of fracture.3,4

In practice, the physician is often faced with making treatment decisions for women with T scores of −2.5 to −1.0 but is provided with little evidence-based guidance regarding who is at highest risk of fracture. Using a T score threshold of −2.5 for diagnostic and treatment decisions has limitations in effectively managing postmenopausal women with low BMD because more than 50% of fractures occur in women whose BMD levels are in the osteopenic range.1,3-8 Therefore, effective strategies to reduce patients’ risk of fracture must include identification and management of individuals who are osteopenic and at high risk of near-term and lifetime fractures.2 The ability to identify osteopenic women at greatest risk of fracture could maximize effective use of therapies and minimize treatment of women who are at lower fracture risk.

A few published guidelines9-11 attempt to provide strategies incorporating BMD measures and additional risk factors for identification of women at increased risk of fracture with T scores greater than −2.5. However, the selection of risk factors for these guidelines does not indicate their relative importance, and the

From the Colorado Center for Bone Research, Lakewood, and the University of Colorado Health Sciences Center, Denver (Dr Miller); the Department of Outcomes Research and Management, Merck & Co Inc, West Point, Pa (Drs Barlas, Brenneman, Abbott, and Chen); the Department of Family and Preventive Medicine, University of California, San Diego, La Jolla (Dr Barrett-Connor); and the College of Physicians and Surgeons, Columbia University, New York, NY (Dr Siris). Drs Miller, Barrett-Connor, and Siris receive consulting fees from Merck & Co Inc to compensate them for time spent working on the NORA project.
T score cutoff value for intervention is somewhat arbitrary. Other studies have identified demographic and clinical factors, in addition to BMD, that play important roles as predictors of fracture in postmenopausal women primarily older than 65 years or have attempted to provide guidance for clarification of salient risk factors. The ability to use the results of the studies is hampered because more risk factors have been identified than would be practical to assess, and they do not specifically address fracture risk within the T score range of −2.5 to −1.0 or in women younger than 65 years, where guidance for clinical decision making is often needed.

To address these issues, we attempted to identify osteopenic women at high risk of fracture within 12 months of peripheral BMD testing. Although the WHO diagnostic criteria for osteoporosis and osteopenia are primarily based on central dual-energy x-ray absorptiometry (DXA) measurements, we use the terminology based on peripheral measurements in this article for purposes of convenience.

The large cohort of postmenopausal women in the National Osteoporosis Risk Assessment (NORA) provides a unique opportunity to assess short-term fracture risk in women with a diagnosis of osteopenia; 39% of the NORA women had peripheral T scores of −2.5 to −1.0 and experienced approximately 50% of the osteoporotic fractures. Using a tree-based approach with noninvasive and easily ascertained data, we aimed to develop a simple algorithm that allows clinicians to classify women with a peripheral BMD-based diagnosis of osteopenia into varying levels of risk for appropriate management.

METHODS

DATA SOURCE

We used data from a longitudinal observational study of osteoporosis among postmenopausal women in the United States that began in 1997. In brief, postmenopausal women who were 50 years and older, without an osteoporosis diagnosis, and who had not had a BMD measurement within the preceding 12 months were eligible for participation. Women currently being treated with a bisphosphonate, calcitonin, or raloxifene hydrochloride were ineligible for participation, as were women who were participants in any other clinical trial related to osteoporosis. At baseline, each participant completed a common core questionnaire and 4 of 8 supplemental questionnaires. The supplemental questionnaires were randomly assigned to each participant. Measurement of BMD at 1 of 3 skeletal sites (heel, hand, or forearm) was performed at the physician's office. Approximately 12 months after enrollment, follow-up questionnaires inquired about fractures that had occurred since enrollment in NORA. All study protocols and consent documents were approved by a national institutional review board (Essex Institutional Review Board Inc, Lebanon, NJ).

STUDY POPULATION

This analysis is restricted to 57,421 white women with baseline peripheral BMD T scores of −2.5 to −1.0 who participated in the 1-year follow-up and for whom fracture incidence data were available. Women with T scores less than −2.5 were excluded from the analysis because treatments are considered essential in these women. Women with T scores greater than −1.0 were not included because these values are not considered clinically relevant for treatment decisions.

DEFINITION OF RISK FACTORS

The risk factors used as potential predictors were compiled from the literature and self-reported in the baseline questionnaires, including the main survey and 3 supplements, as indicated in Table 1.

Each participant received BMD testing at one of the following peripheral sites: forearm using peripheral DXA (pDEXA; Norland Medical Systems Inc, White Plains, NY), finger using peripheral DXA (AccuDEXA; Schick Technologies Inc, Long Island City, NY), or heel using either single-energy x-ray absorptiometry (OsteoAnalyzer; Norland Medical Systems Inc) or ultrasound (Sahara; Hologic Inc, Bedford, Mass). Bone mineral density measures from the 4 devices were pooled in this analysis because they have been shown to predict fracture equally well.

INCIDENT FRACTURES

The outcome variable was incident osteoporotic fractures self-reported at 1-year follow-up. Osteoporotic fractures were defined as clinical fractures of the hip, wrist, or forearm, rib, and vertebrae. These fractures have been shown to be associated with low bone mass.

STATISTICAL ANALYSIS

Tree Derivation

Multivariate analysis was performed using classification trees introduced by Breiman et al; CART (Classification and Regression Trees) 4.0 software (Salford Systems, San Diego, Calif) was used for the analysis. Briefly, trees provide a hierarchical classification process that is represented by a series of yes-no decision points similar to the way clinicians make prognostic and diagnostic decisions. The goal is to place each patient into a class in which the incidence of the outcome is either very high or very low.

To ensure high sensitivity of the prediction rule, the loss due to incorrectly classifying a fracture event was set higher than the loss due to misclassification of a patient without fracture. Assumption of equal previous probabilities of fracture was not appropriate owing to the low fracture rate (2%); therefore, this asymmetry was incorporated into the tree-building process. Terminal subgroups resulting from any given split were required to have at least 10 patients.

To obtain a set of reliable estimates of the independent predictive accuracy of the tree, we used 10-fold cross validation that split the data into approximately 10 parts. After the maximal tree was built on the entire sample, the sample was divided into 10 equal parts, each containing a similar distribution of the outcome variable. The first 9 parts of the data were used to construct the largest possible tree, and the remaining 1 part was used to obtain initial estimates. The process was repeated on another 9 of 10 data parts while using a different part as the test sample until each part of the data had been held in reserve 1 time as a test sample. The results of the 10 mini-test samples were then combined and applied to the tree based on the entire sample.

To assess the importance of variables that were not incorporated into the final tree, we examined the surrogate and competitor splits at each node of the tree. A surrogate split uses an additional variable to complete a class in which the incidence of the outcome is either very high or very low.

The outcome variable was incident osteoporotic fractures self-reported at 1-year follow-up. Osteoporotic fractures were defined as clinical fractures of the hip, wrist, or forearm, rib, and vertebrae. These fractures have been shown to be associated with low bone mass.

STATISTICAL ANALYSIS

Tree Derivation

Multivariate analysis was performed using classification trees introduced by Breiman et al; CART (Classification and Regression Trees) 4.0 software (Salford Systems, San Diego, Calif) was used for the analysis. Briefly, trees provide a hierarchical classification process that is represented by a series of yes-no decision points similar to the way clinicians make prognostic and diagnostic decisions. The goal is to place each patient into a class in which the incidence of the outcome is either very high or very low.

To ensure high sensitivity of the prediction rule, the loss due to incorrectly classifying a fracture event was set higher than the loss due to misclassification of a patient without fracture. Assumption of equal previous probabilities of fracture was not appropriate owing to the low fracture rate (2%); therefore, this asymmetry was incorporated into the tree-building process. Terminal subgroups resulting from any given split were required to have at least 10 patients.

To obtain a set of reliable estimates of the independent predictive accuracy of the tree, we used 10-fold cross validation that split the data into approximately 10 parts. After the maximal tree was built on the entire sample, the sample was divided into 10 equal parts, each containing a similar distribution of the outcome variable. The first 9 parts of the data were used to construct the largest possible tree, and the remaining 1 part was used to obtain initial estimates. The process was repeated on another 9 of 10 data parts while using a different part as the test sample until each part of the data had been held in reserve 1 time as a test sample. The results of the 10 mini-test samples were then combined and applied to the tree based on the entire sample.

To assess the importance of variables that were not incorporated into the final tree, we examined the surrogate and competitor splits at each node of the tree. A surrogate split uses another predictor but results in similar classification of cases. Competitor splitters are variables that can be used instead of primary splitters, resulting in a tree with performance similar
Table 1. Demographic Characteristics and Risk Factor Profiles of Participants at Baseline by Osteoporotic Fracture Status Within 1 Year

<table>
<thead>
<tr>
<th>Response, %</th>
<th>Fracture (n = 1130)</th>
<th>No Fracture (n = 56291)</th>
<th>Total (n = 57421)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>99</td>
<td>68.7 (9.2)</td>
<td>66.7 (8.8)</td>
<td>66.8 (8.8)</td>
</tr>
<tr>
<td>Age category, No. (% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-59 y</td>
<td>224 (1.6)</td>
<td>13,555 (98.4)</td>
<td>13,779 (24)</td>
<td></td>
</tr>
<tr>
<td>60-69 y</td>
<td>345 (1.7)</td>
<td>20,418 (98.3)</td>
<td>20,763 (36)</td>
<td></td>
</tr>
<tr>
<td>70-79 y</td>
<td>418 (2.2)</td>
<td>18,209 (97.8)</td>
<td>18,627 (32)</td>
<td></td>
</tr>
<tr>
<td>≥80 y</td>
<td>142 (3.4)</td>
<td>4,103 (96.6)</td>
<td>4,245 (7)</td>
<td></td>
</tr>
<tr>
<td>T score at baseline, mean (SD)</td>
<td>100</td>
<td>−1.72 (0.41)</td>
<td>−1.61 (0.40)</td>
<td>−1.62 (0.40)</td>
</tr>
<tr>
<td>Health status</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>10.2</td>
<td>12.2</td>
<td>12.1</td>
<td></td>
</tr>
<tr>
<td>Very good</td>
<td>29.5</td>
<td>34.8</td>
<td>34.7</td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>37.2</td>
<td>38.8</td>
<td>38.8</td>
<td></td>
</tr>
<tr>
<td>Fair/poor</td>
<td>23.2</td>
<td>14.2</td>
<td>14.4</td>
<td></td>
</tr>
<tr>
<td>History of fracture after age 45 y</td>
<td>99</td>
<td>30.3</td>
<td>14.3</td>
<td>14.6</td>
</tr>
<tr>
<td>Maternal history of fracture</td>
<td>94</td>
<td>28.0</td>
<td>26.5</td>
<td>26.5</td>
</tr>
<tr>
<td>Fracture in first-degree relative</td>
<td>79</td>
<td>31.5</td>
<td>26.0</td>
<td>26.1</td>
</tr>
<tr>
<td>Maternal history of osteoporosis</td>
<td>72</td>
<td>15.2</td>
<td>16.7</td>
<td>16.6</td>
</tr>
<tr>
<td>Dowager hump b</td>
<td>79</td>
<td>21.4</td>
<td>14.1</td>
<td>14.2</td>
</tr>
<tr>
<td>Poor mobility c</td>
<td>100</td>
<td>48.9</td>
<td>34.9</td>
<td>35.2</td>
</tr>
<tr>
<td>Regular exercise, times/wk</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2</td>
<td>49.0</td>
<td>47.0</td>
<td>47.1</td>
<td></td>
</tr>
<tr>
<td>&gt;3</td>
<td>51.0</td>
<td>53.0</td>
<td>52.0</td>
<td></td>
</tr>
<tr>
<td>Height loss, mean (SD), cm d</td>
<td>95</td>
<td>1.97 (2.8)</td>
<td>1.64 (2.5)</td>
<td>1.64 (2.5)</td>
</tr>
<tr>
<td>Weight loss since 20 s</td>
<td>99</td>
<td>10.9</td>
<td>7.0</td>
<td>7.1</td>
</tr>
<tr>
<td>Low body mass index (&lt;18.5 kg/m²)</td>
<td>97</td>
<td>2.7</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Low weight (&lt;127 lb)</td>
<td>100</td>
<td>18.3</td>
<td>20.0</td>
<td>19.9</td>
</tr>
<tr>
<td>Hormone therapy</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>43.1</td>
<td>39.1</td>
<td>39.1</td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>24.8</td>
<td>22.1</td>
<td>22.1</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>32.1</td>
<td>38.9</td>
<td>38.8</td>
<td></td>
</tr>
<tr>
<td>Menopause at age &lt;40 y</td>
<td>89</td>
<td>22.4</td>
<td>17.9</td>
<td>17.9</td>
</tr>
<tr>
<td>Age at menopause, mean (SD), y</td>
<td>89</td>
<td>45.0 (7.6)</td>
<td>45.9 (7.2)</td>
<td>45.9 (7.2)</td>
</tr>
<tr>
<td>Nulliparity e</td>
<td>96</td>
<td>10.7</td>
<td>10.7</td>
<td>10.7</td>
</tr>
<tr>
<td>Poor vision e</td>
<td>94</td>
<td>56.7</td>
<td>50.4</td>
<td>50.4</td>
</tr>
<tr>
<td>Adequate lifetime calcium intake g</td>
<td>100</td>
<td>7.2</td>
<td>7.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Adequate vitamin D exposure h</td>
<td>100</td>
<td>23.8</td>
<td>24.2</td>
<td>24.2</td>
</tr>
<tr>
<td>Thyroid medication use</td>
<td>95</td>
<td>21.0</td>
<td>19.1</td>
<td>19.1</td>
</tr>
<tr>
<td>Sedative use i</td>
<td>83</td>
<td>17.2</td>
<td>10.1</td>
<td>10.2</td>
</tr>
<tr>
<td>Epilepsy medication use i</td>
<td>85</td>
<td>3.2</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>Diuretic use</td>
<td>92</td>
<td>20.1</td>
<td>16.8</td>
<td>16.9</td>
</tr>
<tr>
<td>Cortisone use</td>
<td>91</td>
<td>4.6</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Insulin-dependent diabetes mellitus i</td>
<td>97</td>
<td>3.9</td>
<td>2.4</td>
<td>2.5</td>
</tr>
<tr>
<td>Hyperthyroidism i</td>
<td>87</td>
<td>5.7</td>
<td>4.7</td>
<td>4.7</td>
</tr>
<tr>
<td>Presence of &gt;1 chronic illness j</td>
<td>100</td>
<td>26.8</td>
<td>21.9</td>
<td>22.0</td>
</tr>
<tr>
<td>Cigarette use, mean (SD), pack-years k</td>
<td>88</td>
<td>12.6 (24.0)</td>
<td>10.8 (20.0)</td>
<td>10.8 (20.1)</td>
</tr>
<tr>
<td>Smoking status</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>50.6</td>
<td>53.4</td>
<td>53.3</td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>37.0</td>
<td>35.2</td>
<td>35.3</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>12.4</td>
<td>11.4</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>Alcohol use, drinks/wk</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>76.6</td>
<td>72.7</td>
<td>72.8</td>
<td></td>
</tr>
<tr>
<td>1-6</td>
<td>14.4</td>
<td>18.2</td>
<td>16.1</td>
<td></td>
</tr>
<tr>
<td>7-13</td>
<td>5.8</td>
<td>6.1</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>≥14</td>
<td>3.3</td>
<td>3.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Caffeine use, drinks/wk</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>21.4</td>
<td>20.4</td>
<td>20.4</td>
<td></td>
</tr>
<tr>
<td>1-6</td>
<td>16.0</td>
<td>13.8</td>
<td>13.8</td>
<td></td>
</tr>
<tr>
<td>7-13</td>
<td>21.6</td>
<td>21.6</td>
<td>21.6</td>
<td></td>
</tr>
<tr>
<td>≥14</td>
<td>41.1</td>
<td>44.3</td>
<td>44.2</td>
<td></td>
</tr>
</tbody>
</table>

aData are given as percentages, except where indicated otherwise.

bSupplement 5 (n = 28,443).

Supplement 7 (n = 27,524). Poor mobility is positive responses to 2 of the 4 questions related to physical functioning found in the 12-Item Short-Form Health Survey.25

dHeight loss is the difference between the tallest ever height and the current height.

Supplement 6 (n = 28,124).

Supplement 3 (n = 27,915).

Adequate lifetime calcium intake is lifetime calcium intake greater than 800 mg at all phases (child, teen, adult, and current).

Adequate lifetime vitamin D exposure is total vitamin D exposure (from milk, fish, sunlight, and vitamin D supplement) exceeding 600 U.

Supplement 4 (n = 22,219).

Multiple chronic illnesses is the reporting of at least 2 of the following medical conditions: hypothyroidism, hyperthyroidism, parathyroidectomy, type 1 or type 2 diabetes mellitus, epilepsy, kidney disease, liver disease, osteoarthritis, coronary heart disease, breast cancer, and other cancer.

Pack-years smoked is calculated as (number of cigarettes smoked per day/20) × (number of years smoked).
Of the 57,421 NORA women included in the present analysis, 1130 (2.0%) reported an incident clinical osteoporotic fracture within 12 months of BMD testing, including 196 hip fractures, 126 vertebral fractures, 319 rib fractures, and 535 wrist or forearm fractures. The participant age range was 50 to 99 years (mean [SD] age, 66.8 [8.8] years), and the mean (SD) T score was −1.62.

The algorithm identified 55% of the women as being at increased risk, with an overall 1-year risk of fracture of 2.4% (range, 1.9%-4.1%). Women with previous fracture, regardless of T score, had a 1-year fracture risk of 4.1%, followed by 2.2% in women without previous fracture with T scores of −1.8 or less or with poor health. Women with none of these conditions but with poor mobility had a fracture risk of 1.9%. Among 57,421 women in this analysis, 26,037 (45.3%) were identified as not being at increased risk, with 1.1% experiencing a fracture. The algorithm correctly classified 74.1% of the women in this cohort who experienced an incident fracture within 1 year.

As noted in the Figure, knowledge of previous fracture alone identified 339 (30.0%) of 1130 fractures, suggesting that the simplest tree, containing only previous fracture, can be used to identify one third of the women who subsequently had a fracture. Combining previous fracture and BMD T score, 58.8% of women with new fractures (n=665) occurring within 1 year can be identified.

Table 2 summarizes the classification rules indicated by the decision tree in the Figure into 4 simple steps. In general, the algorithm retained the ability to identify women at various ages who were at increased fracture risk (Table 3). In particular, women identified in age groups 50 to 59, 60 to 69, and 70 to 79 years had overall fracture risks of 2.6%, 2.2%, and 2.8%, respec-

---

CHARACTERISTICS OF THE STUDY SAMPLE

As shown in order of importance in the Figure, 4 of the 32 predictors were identified as the most important determinants for short-term fracture prediction. Previous fracture was the strongest predictor of fractures within 1 year. The BMD T score cutoff point that best differentiated women at increased risk of fracture among those without previous fracture was found to be −1.8. Self-reported fair or poor general health status and poor mobility were also associated with increased risk of short-term fractures. Previous fracture information was obtained from responses to the question, “Since the age of 45, have you broken any of the following bones: hip, rib, wrist, or spine (backbone)?” Self-reported health status was obtained from responses to the question, “In general, would you say your health is excellent, very good, good, fair, or poor?” Mobility was determined by the average response to 4 questions related to physical functioning in the 12-item Short-Form Health Survey: (1) Does your health limit you in moderate activities, such as pushing a vacuum cleaner, bowling, or playing golf? (2) Does your health limit you in climbing several flights of stairs? (3) As a result of your physical health, have you accomplished less than you would like? (4) As a result of your physical health, are you limited in the kind of work you do or other activities?30 Poor mobility was defined as 2 or more positive responses to these 4 questions.

The algorithm identified 55% of the women as being at increased risk, with an overall 1-year risk of fracture of 2.4% (range, 1.9%-4.1%). Women with previous fracture, regardless of T score, had a 1-year fracture risk of 4.1%, followed by 2.2% in women without previous fracture with T scores of −1.8 or less or with poor health. Women with none of these conditions but with poor mobility had a fracture risk of 1.9%. Among 57,421 women in this analysis, 26,037 (45.3%) were identified as not being at increased risk, with 1.1% experiencing a fracture. The algorithm correctly classified 74.1% of the women in this cohort who experienced an incident fracture within 1 year.

As noted in the Figure, knowledge of previous fracture alone identified 339 (30.0%) of 1130 fractures, suggesting that the simplest tree, containing only previous fracture, can be used to identify one third of the women who subsequently had a fracture. Combining previous fracture and BMD T score, 58.8% of women with new fractures (n=665) occurring within 1 year can be identified. Table 2 summarizes the classification rules indicated by the decision tree in the Figure into 4 simple steps. In general, the algorithm retained the ability to identify women at various ages who were at increased fracture risk (Table 3). In particular, women identified in age groups 50 to 59, 60 to 69, and 70 to 79 years had overall fracture risks of 2.6%, 2.2%, and 2.8%, respec-

---

TREE STRUCTURE VALIDATION

We confirmed the accuracy of the classification tree by performing a logistic regression model and a hybrid of CART and a logistic regression model to eliminate the possibility of overrepresentation of simple structures in the data. In the logistic regression analysis, all risk factors were included as main effects in a forward stepwise fashion without higher-order terms or interactions. In the hybrid of CART and logistic regression analysis, the CART terminal nodes were entered as a single, categorical predictor (cartnode) into logistic regression along with all 32 predictors. All logistic regression analyses were performed using SAS version 6.12 (SAS Institute Inc, Cary, NC).

RESULTS

CHARACTERISTICS OF THE STUDY SAMPLE

As shown in order of importance in the Figure, 4 of the 32 predictors were identified as the most important determinants for short-term fracture prediction. Previous fracture was the strongest predictor of fractures within 1 year. The BMD T score cutoff point that best differentiated women at increased risk of fracture among those without previous fracture was found to be −1.8. Self-reported fair or poor general health status and poor mobility were also associated with increased risk of short-term fractures. Previous fracture information was obtained from responses to the question, “Since the age of 45, have you broken any of the following bones: hip, rib, wrist, or spine (backbone)?” Self-reported health status was obtained from responses to the question, “In general, would you say your health is excellent, very good, good, fair, or poor?” Mobility was determined by the average response to 4 questions related to physical functioning in the 12-item Short-Form Health Survey: (1) Does your health limit you in moderate activities, such as pushing a vacuum cleaner, bowling, or playing golf? (2) Does your health limit you in climbing several flights of stairs? (3) As a result of your physical health, have you accomplished less than you would like? (4) As a result of your physical health, are you limited in the kind of work you do or other activities?30 Poor mobility was defined as 2 or more positive responses to these 4 questions.

The algorithm identified 55% of the women as being at increased risk, with an overall 1-year risk of fracture of 2.4% (range, 1.9%-4.1%). Women with previous fracture, regardless of T score, had a 1-year fracture risk of 4.1%, followed by 2.2% in women without previous fracture with T scores of −1.8 or less or with poor health. Women with none of these conditions but with poor mobility had a fracture risk of 1.9%. Among 57,421 women in this analysis, 26,037 (45.3%) were identified as not being at increased risk, with 1.1% experiencing a fracture. The algorithm correctly classified 74.1% of the women in this cohort who experienced an incident fracture within 1 year.

As noted in the Figure, knowledge of previous fracture alone identified 339 (30.0%) of 1130 fractures, suggesting that the simplest tree, containing only previous fracture, can be used to identify one third of the women who subsequently had a fracture. Combining previous fracture and BMD T score, 58.8% of women with new fractures (n=665) occurring within 1 year can be identified. Table 2 summarizes the classification rules indicated by the decision tree in the Figure into 4 simple steps. In general, the algorithm retained the ability to identify women at various ages who were at increased fracture risk (Table 3). In particular, women identified in age groups 50 to 59, 60 to 69, and 70 to 79 years had overall fracture risks of 2.6%, 2.2%, and 2.8%, respec-
Women 80 years and older identified by the algorithm had a greater risk of 3.9%. Likewise, the absolute risk for fracture within each age group according to individual predictors was similar to that of the overall cohort that was identified as being at increased risk.

No surrogate splits were found by CART at any of the 4 steps of the algorithm, indicating that none of the remaining variables would result in similar classification at the node. Evaluation of competitor splits, however, showed several risk factors that were important in their relation to fractures. Age (split at 71.5 years) and years since menopause (split at 3.5 years) were competitors at every split. Age, although not as strong as the primary splitters, was the next important predictor. In cases in which the primary predictor is not available or is difficult to obtain, the competitor variable, age, can be used for identification. The use of sedatives was a competitor for the T score and health status primary splits. Low body weight (<127 lb) was a competitor split for the health status and mobility primary splits. Inclusion of additional variables did not add substantial improvement to the performance of the tree beyond the 4 selected variables shown in the Figure.

**TREE STRUCTURE VALIDATION**

A logistic regression model entering all 32 risk factors in a forward stepwise selection resulted in selection of the 4 variables used in the CART classification tree as the first variables in the final parsimonious model. In the hybrid of CART and logistic regression analysis, CART terminal nodes (cartnode) was the first variable selected in a forward stepwise procedure. Additional variables selected were years since menopause and health status. Logistic regression and the hybrid of CART and logistic regression results confirm the accuracy of the simple classification tree derived from CART in identifying individuals at increased risk of fractures.

**COMMENT**

Using prospective data from a large cohort of postmenopausal osteopenic women, we showed that a simple algorithm derived from easily assessed risk factors predicted the risk of osteoporotic fracture within 1 year of BMD testing. Among 57,421 osteopenic women, the proposed algorithm of previous fracture, T score of −1.8 or less, poor self-reported health status, and poor self-reported mobility identified 55% as being at increased risk and correctly identified 74.1% of the women who had a fracture within 1 year. Osteopenic women with a previous fracture were found to have a risk similar to that of women in the NORA cohort with T scores of −2.5 or less (4.1% and 4.3%, respectively). On the other hand, osteopenic women not identified as being at increased risk had a fracture risk of 1.1%, similar to that of women in the NORA cohort with T scores greater than −1.0. If the goal of osteoporosis-specific therapy is to reduce a woman’s risk of fracture, an algorithm should be used that correctly identifies the most women who have the highest risk of fractures by the most practical means. The NORA-based algorithm provides the clinician with a valuable and practical tool to accomplish this risk assessment and to design appropriate management strategies.

To our knowledge, this is the only study to date to develop a risk prediction tool for 1-year osteoporotic fracture specifically for postmenopausal women of all ages (50-99 years) with osteopenia (T score of −2.5 to −1.0). The FRACTURE Index developed by Black and colleagues included 7782 women 65 years and older, with no restriction on BMD T scores, to predict 5-year fracture risk. McGrother et al developed a risk score to predict 3-year hip fracture risk in elderly women based on a sample of 1864 women older than 70 years with no limitation on BMD T scores. The Rotterdam study developed a risk score for prediction of 4-year hip fracture risk in 5208 men and women 55 years and older at all BMD levels.

Focusing the analysis on women with osteopenia provides relevant clinical information for women with what is considered to be low bone mass and for whom guidance regarding additional risk factors for further classification of fracture risk is lacking. The National Osteoporosis Foundation recommendations provide widely accepted guidance for the management of women with T scores higher than −2.5. These guidelines recommend that pharmacologic treatment should be initiated in women who have T scores less than −2.0 or less than −1.5 with at least 1 of 15 risk factors. Clinicians’ ability to use the National Osteoporosis Foundation guidelines is somewhat limited since the guidelines do not indicate the relative importance of the risk factors, leaving the clinician to rely on judgment or to revert to using only T score values. The NORA-based risk algorithm is consistent with the National Osteoporosis Foundation treatment recommendations but is further enhanced with an empirically derived T score threshold and a well-defined small subset of risk factors that can be easily implemented in clinical practice.

A concern may be that peripheral devices were used to determine BMD T scores and to define this cohort. However, the WHO diagnostic criteria were established based on central (hip and spine) and peripheral (wrist) BMD measurement devices. T scores obtained using peripheral devices may not always be as low as T scores determined using central DXA devices, resulting in a prevalence of WHO osteoporosis using peripheral devices—specific databases of 3% to 14% compared with a prevalence based on hip measurements for white women of 16% to 20%. The discrepancies among T score calculations across various BMD devices are well recognized and exist among different central DXA skeletal sites and devices as well. Thus, even among the 3 available central DXA devices, the same patient could be clas-

---

**Table 2. Classification Tree Rules for Predicting Fracture**

<table>
<thead>
<tr>
<th>Decision rules for high risk of short-term fracture:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. History of fracture or</td>
</tr>
<tr>
<td>2. Bone mineral density T score ≤−1.8 or</td>
</tr>
<tr>
<td>3. Fair or poor self-reported health status or</td>
</tr>
<tr>
<td>4. Poor self-reported mobility</td>
</tr>
</tbody>
</table>

---

©2004 American Medical Association. All rights reserved.
A macologic intervention has been shown to reduce sub-
T scores between –2.5 and –1.0, a previous fracture risk were diagnostic of osteoporosis. In NORA women with 
ments became widely available, low-traumatic fractures assessed by any BMD device, whether central or periph-
BMD T scores less than −1.0 in postmenopausal women measured by a central device. Despite the discrepancies, low 
differences among peripheral devices, if the T score was 
less than −1.0, women were at increased risk of fracture, 
dictor of future fracture. In fact, before BMD measure-
tral BMD measurements is not possible.

The risk factors used in the NORA-based algo-

The NORA-based algorithm can assist clinicians in 

This study has several limitations. First, women who 

The NORA-based algorithm can assist clinicians in 

The predictive relationships found in this study are 

The risk factors used in the NORA-based algo-

The risk factors used in the NORA-based algo-

The predictive relationships found in this study are 

The NORA-based algorithm is well established for their relationship with osteoporosis or fractures, especially history of previous fracture and low BMD. Poor health status and poor mobility have also been shown to affect fracture risk, and they may function as surrogates for propensity to fall. Although age, height, weight, and maternal fracture were identified to be important predictors in risk models based on women 65 years and older or with T scores less than −2.5, our analysis did not find that these factors added incremental value to the predic-
tive ability of the final model in this cohort that included women younger than 65 years and was re-
stricted to T scores between −2.5 and −1.0. Although, as a competitor variable, a cutoff age of 71.5 years can be used as a predictor if any of the primary predictors are unavailable or are difficult to obtain.

The NORA-based algorithm can assist clinicians in identifying a group of women with osteopenia who are at increased short-term risk of new osteoporotic fractures, women for whom interventions to reduce risk should be considered. The strongest risk factor in our algorithm was previous fracture, a known powerful pre-
dictor of future fracture. In fact, before BMD measurements became widely available, low-traumatic fractures were diagnostic of osteoporosis. In NORA women with T scores between −2.5 and −1.0, a previous fracture risk level was similar to a BMD diagnosis of osteoporosis. Pharmacologic intervention has been shown to reduce sub-
sequent fracture in women with previous vertebral frac-
ture.

In women without previous fracture, a peripheral T score of −1.8 or lower indicated increased risk of short-
term fracture. Further testing or treatment needs to be considered in these women. In women who do not have either previous fracture or low T scores, poor reported health status and poor mobility may function as surrogates for predictors of future fracture. Each of the latter factors is potentially modifiable through nonpharmaco-
logic interventions, such as diet and exercise.

The predictive relationships found in this study are further supported by the ability to apply the algorithm to specific age groups and obtain similar risk profiles for fracture prediction. Of particular interest are younger women aged 50 to 59 years who had an overall 1-year risk of fracture similar to that of women aged 60 to 69 years (1.6% and 1.7%, respectively). Women aged 50 to 59 years who were identified by the algorithm to be at increased risk had an absolute fracture risk of 2.6%, similar to that of the total identified cohort (2.2%). In addition, women of this age with a history of fracture after age 45 years also had a 4.5% fracture risk, similar to the risk of women in NORA with T scores less than −2.5 (4.3%).

This study has several limitations. First, women who responded to the follow-up survey may differ from the 18% of nonresponders in that women who had a fracture may have been more (or less) likely to respond to follow-up than women who did not. Second, fracture information was collected by self-report. However, other researchers have found self-report of fractures to be generally reliable. Because most of the spine fractures are asymptomatic or at least unrecognized, NORA cannot add-

NORA cannot address the value of risk factors or peripheral BMD to predict nonclinical spine fractures. In the long term, clinical and subclinical vertebral fractures are associated with increased morbidity and mortality. Third, information on some risk factors for fracture, such as muscle strength and propensity to fall, were not available; these factors may exert an effect on fracture risk through health status or mobility status. Finally, the algorithm was derived based on 1-year fracture data. Its utility for long-
term fracture prediction is unknown. However, unless

<table>
<thead>
<tr>
<th>Age Group, y</th>
<th>Total fractures, No. (%)</th>
<th>Predictor variables, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-59 (n = 13 779)</td>
<td>224 (1.6)</td>
<td>2.6</td>
</tr>
<tr>
<td>60-69 (n = 20 763)</td>
<td>345 (1.7)</td>
<td>2.2</td>
</tr>
<tr>
<td>70-79 (n = 18 627)</td>
<td>418 (2.2)</td>
<td>2.8</td>
</tr>
<tr>
<td>≥80 (n = 4245)</td>
<td>142 (3.4)</td>
<td>3.9</td>
</tr>
<tr>
<td>Total Cohort (N = 57 421)</td>
<td>1130 (2.0)</td>
<td>2.4</td>
</tr>
</tbody>
</table>

*Mutually exclusive categories of women identified by each predictor variable.
the level of a given short-term predictor is unstable over time relative to other predictors, it should be a risk indicator of longer-term fracture risk.

In conclusion, we developed a simple classification tool to identify postmenopausal women with osteopenia (T scores of −2.5 to −1.0) at the highest risk of fracture within 12 months. This classification tool identified more than 70% of those who experience a fracture based on information that can easily be obtained during a routine clinical visit. These results are based on the NORA study population. Although validation in a separate cohort is necessary to further substantiate the algorithm, our intent is to present on option for managing women with T scores greater than −2.5 who are at increased risk for fracture. This NORA-based algorithm can be useful in clinical practice to guide further assessment and management decisions in a large group of women.

Accepted for publication June 30, 2003.

The NORA project was funded and managed by Merck & Co. Inc in collaboration with the International Society for Clinical Densitometry, West Hartford, Conn.

Presented in part as posters at the annual meeting of the American Society for Bone and Mineral Research; September 21, 2002; San Antonio, Tex.

We thank Kenneth Faulkner, PhD, for his direction in study design, data collection, and data analysis and our colleagues at Merck & Co, Inc, Parlex International (Waltham, Mass), and Abt Associates Inc (Cambridge, Mass), who were involved in the implementation and data collection efforts undertaken on behalf of NORA.

Corresponding author and reprints: Paul D. Miller, MD, Colorado Center for Bone Research, 3190 S Wadsworth Blvd, Suite 250, Lakewood, CO 80227 (e-mail: millercbr@aol.com).

REFERENCES