Published in final edited form as:

Osteoporos Int. 2013 April ; 24(4): 1185–1193. doi:10.1007/s00198-012-2215-3.

Performance of FRAX in a Cohort of Community Dwelling, Ambulatory Older Men: The Osteoporotic Fractures in Men (MrOS) Study

Bruce Ettinger, MD¹, Kristine E. Ensrud, MD, MPH², Terri Blackwell, MA³, Jeffrey R. Curtis, MD, MS, MPH⁴, Jodi A. Lapidus, PhD⁵, and Eric S. Orwoll, MD⁶ For the Osteoporotic Fracture in Men (MrOS) Study Research Group

¹Department of Medicine, University of California, San Francisco

²Department of Medicine, Center for Chronic Disease Outcomes Research, VA Medical Center, Minneapolis, MN

³California Pacific Medical Center Research Institute, San Francisco, CA

⁴Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, AL

⁵Department of Public Health and Preventive Medicine, Division of Biostatistics, Oregon Health & Sciences University, Portland, OR

⁶Bone and Mineral Unit, Oregon Health & Science University, Portland, OR

Abstract

FRAX has been extensively studied in women, but there are few studies of its performance in men. We evaluated its performance in older men who participated in the Osteoporotic Fractures in Men Study (MrOS).

FRAX estimates for 10-year hip fracture and major osteoporotic fracture (MOF; hip, clinical spine, forearm, and shoulder) were calculated from data obtained from MrOS participants and compared to observed 10-year fracture cumulative incidence calculated using product limit estimate methods, accounting for competing mortality risk.

5,891 men were followed for an average of 8.4 years. Without BMD in the FRAX model, the mean 10-year predicted fracture probabilities for hip and MOF were 3.5% and 8.9%, respectively; addition of BMD to the calculations reduced these estimates to 2.3% and 7.6%. Using FRAX without BMD, predicted quintile probabilities closely estimated cumulative incidence of hip fracture (range of observed to predicted ratios 0.9-1.1). However, with BMD in the FRAX calculation, observed to predicted hip fracture probabilities were not close to unity and varied markedly across quintiles of predicted probability. For MOF, FRAX without BMD overestimated observed cumulative incidence (range of observed to predicted ratios 0.7-0.9) and addition of BMD did not improve this discrepancy (range of observed to predicted ratios 0.7-1.1). Addition of BMD to the calculation had mixed effects on the discriminatory performance of FRAX, depending on the analysis tool applied.

Correspondence: Bruce Ettinger, MD, 156 Lombard Street #13, San Francisco, CA 94111, Telephone: 1-415-433-6836, FAX: 1-415-433-6836, doc.ettinger@gmail.com.

Conflict of Interest: With the exception of Dr. Ettinger, none of the other co-authors report potential conflicts of interest. Dr. Ettinger has received consulting fees from Foundation for Osteoporosis Research and Education (FORE) and currently serves on their Medical Science Advisory Board.

Among this cohort of community dwelling older men, the FRAX risk calculator without BMD was well calibrated to hip fracture but less well to MOF.

Keywords

fracture; prediction; men; risk assessment

INTRODUCTION

FRAX is the world-wide fracture risk assessment tool [1, 2] and has been incorporated into osteoporosis management and treatment guidelines in the US [3] and abroad [4]. The International Osteoporosis Foundation (IOF) has recently offered FRAX estimates in multi-patient entry mode [5]; this could allow data from electronic medical records or large patient cohorts and registries to be batch processed for research studies and population management programs.

Two published studies of FRAX's predictive performance in men are currently available [6,7]. Yun et al. used area under the receiver operating curve (AUC) analyses to assess US-FRAX performance for hip fracture among Medicare beneficiaries 65 years and older; the C-statistic of 0.67 was the same as that obtained in older women in the same cohort [6], and the performance of FRAX (calculated without BMD) was significantly worse than fracture risk estimated from a simple model using only administrative claims data. Leslie et al. performed more extensive analyses to assess calibration of Canadian-FRAX against observed hip and major osteoporotic fracture probabilities among men and women in a Manitoba, Canada database [7]. Although limited by small numbers of fractures in men (43 hip and 163 MOF), they concluded there was "reasonable agreement between observed and predicted hip fracture rates"; the FRAX tool, with BMD included, minimally overestimated the observed hip fracture probability. In contrast, they reported that agreement between observed and predicted MOF was "more variable" with an overall 24% underestimation. This group was also able to assess the possible contribution of BMD to the FRAX estimates. Compared to results obtained without BMD, addition of BMD to the calculation had minimal (<10% reduction) effects on FRAX-estimated hip fracture and MOF probabilities but significantly increased the AUC.

Performance of an alternative fracture risk tool (FRC) in the MrOS cohort was recently published [8]. Based on measures of calibration and discrimination, the tool showed promising results; further, addition of BMD to the FRC estimation minimally altered the risk estimates.

The MrOS cohort of older men has been closely followed for fracture up to 10 years, [9,10] and thus provides an excellent opportunity to evaluate FRAX's calibration and discrimination. Our primary analysis goal was to determine how closely FRAX-predicted estimates would mirror observed fracture probabilities in the MrOS cohort. Our secondary analysis goal was to determine the ability of FRAX scores to discriminate individuals who would subsequently fracture from those remaining fracture free. In both analyses, we hypothesized that addition of BMD would enhance FRAX's predictive performance.

METHODS

PARTICIPANTS

The Osteoporotic Fractures in Men (MrOS) Study is a prospective cohort study designed to examine the extent to which fracture risk is related to various factors. Design and

recruitment have been previously described. [9,10] From March 2000 to April 2002, 5,994 community-dwelling men 65 years or older were enrolled at 6 clinical centers in the United States: Birmingham, Alabama; Minneapolis, Minnesota; Palo Alto, California; the Monongahela Valley near Pittsburgh, Pennsylvania; Portland, Oregon; and San Diego, California. We excluded 101 men who had used a bisphosphonate within 30 days prior to the baseline visit.

The institutional review board at each clinical site approved the study protocol and written informed consent was obtained for all participants.

FRAX AND 10-YEAR FRACTURE RISK ESTIMATES

The US-FRAX (version 3.3) 10-year risk estimates of both hip fracture and any one of four major osteoporotic fractures (MOF; hip, clinical spine, forearm, shoulder) were calculated at the WHO Collaborating Centre for Metabolic Bone Disease, University of Sheffield, UK. Specific patient characteristics (body mass index [BMI], history of fracture, parental history of hip fracture, smoking and alcohol consumption, use of corticosteroids, prevalence of rheumatoid arthritis, secondary osteoporosis) are included in the calculation. If data was missing on any of the other clinical characteristics other than age, race, or BMI, the input value for that characteristic is set to null, as recommended by the World Health Organization. The 10-year risk estimates can be calculated with and without femoral neck bone mineral density (BMD) data.

BONE MINERAL DENSITY

BMD (g/cm^2) of the total hip and hip sub-regions was measured using dual-energy x-ray absorptiometry (DXA) (QDR 4500W, Hologic, Inc., Waltham, MA) ¹⁰; femoral neck BMD in gm/cm^2 was used as the FRAX input variable. T-scores and Z-scores for the femoral neck were calculated based on the race-specific means and standard deviations for men obtained from the National Health and Nutrition Examination Survey III (NHANES III). [11]

OTHER RISK FACTORS

All participants completed questionnaires, which included items about demographics, medical history including rheumatoid arthritis, fracture history, parental hip fracture history, smoking status, and alcohol consumption. Ethnicity was self-reported. The history of any fracture after age 50 was gathered by self-report. Prescription and nonprescription medications used within the preceding 30 days were identified, recorded by the clinics, and stored in an electronic medications inventory database (San Francisco Coordinating Center, San Francisco, CA); each medication was matched to its ingredient(s) based on the Iowa Drug Information Service (IDIS) Drug Vocabulary (College of Pharmacy, University of Iowa, Iowa City, IA). [12] Input values for secondary osteoporosis were set to null when passed to the risk calculator due to lack of clinical information on the conditions associated with bone loss; the FRAX calculation algorithm ignores this variable when BMD is included in the model.

A comprehensive examination included measurements of body weight and height; BMI was calculated as weight in kilograms divided by the square of height in meters.

Bisphosphonate use occurring after the baseline visit was also gathered at the Interim visit (2.0 ± 0.2 years from baseline), the Sleep visit (3.4 ± 0.5 years from baseline), Visit 2 (4.6 ± 0.4 years from baseline), Visit 3 (6.9 ± 0.4 years from baseline) and Interim Visit 2 (8.9 ± 0.4 years from baseline). Use was defined as any bisphosphonate reported to be taken within 30 days prior to a visit.

INCIDENT FRACTURE

After the baseline exam, surviving participants completed and returned a questionnaire about incident fractures every four months; response rates exceeded 99%. Incident fractures were confirmed by centralized physician review of radiology reports or radiologist review of clinical x-rays. The fracture outcomes of interest in this analysis were hip fracture and fracture of either the hip, wrist, shoulder, or clinical spine (major osteoporotic fracture; MOF), regardless of the degree of trauma.

Follow-up time for the fracture analyses was truncated to 10 years to correspond with the 10-year fracture risk estimates from FRAX. Thus, follow-up time ended at the time of the first fracture of that type, time of death, termination from MrOS, or after 10 years of follow-up. For those 418 (7.1%) men who started use of a bisphosphonate after the baseline visit the follow-up time was truncated to the midpoint between the last known visit when they were nonusers and the first visit bisphosphonate use was recorded.

STATISTICAL ANALYSIS

Risk factors of participants were compared across 10-year fracture risk categories established by the NOF cost-effectiveness analyses (3% for hip fracture and 20% for MOF) [13] using chi-square tests for categorical variables, t-tests for normally distributed continuous variables, and Wilcoxon rank-sum tests for continuous variables with skewed distributions.

Calibration (comparison of observed cumulative incidence to predicted probabilities) of the FRAX tool was evaluated using five risk sub-categories. The Kaplan Meier product limit method was used to calculate 10-year cumulative incidence probabilities of fracture in the presence of the competing risk of mortality, with observations censored at 10 years or end of follow-up or the start of bisphosphonate use (whichever came sooner). These observed cumulative incidences were determined for men in each quintile of 10-year fracture probability estimates determined from FRAX and ratios of observed to predicted probabilities were calculated within and across these quintiles.

Logistic regression models with fracture as the outcome and the corresponding FRAX 10-year fracture risk estimate as the predictor were performed. From these models, receiver operating characteristic (ROC) curves were created to compare sensitivity and specificity of the FRAX risk estimates, and the area under the curves (AUC, c-statistic) were calculated. The ROC curves for the fracture risk estimates with and without BMD used in the risk calculation were compared to examine if the addition of BMD to the risk calculator improved the AUC; a c-statistic value of 0.5 is no better than chance alone and 1.0 is perfect prediction.

Reclassification of risk after the addition of BMD to the risk calculator was also assessed using Net Reclassification Improvement (NRI).[14-16] For these analyses, we used clinical cut points based on the NOF cost-effective analyses [13] as well as a no-category approach [14]. Risk stratification tables were created to evaluate the incremental value of adding BMD to the risk prediction calculator.[15,16] In the categorical NRI approach, the NRI sums the proportions of individuals who are appropriately (e.g., moving from a lower- to a higher-risk category for those who did go on to fracture) moved between risk categories and subtracts those moving inappropriately (e.g., moving from a higher- to a lower-risk category for those who did go on to fracture); these calculations are first performed separately by incident fracture status and then combined. The range of the NRI using the categorical approach is -100% to 100%, with positive values indicating a net reclassification in the appropriate direction, negative values indicating inappropriate reassignment; the higher the positive value, the better the reclassification. Using the no-category approach, we

determined the proportions among those with incident fractures whose risk estimate was increased with addition of BMD and subtracted the proportion of those fracturing who moved inappropriately to lower risk estimates with BMD addition. Similar calculations were performed among those men who did not sustain an incident fracture. The overall NRI using the no-category approach is the sum of the fracturing and non-fracturing NRIs and could range from -200% to +200%.

A number of sensitivity analyses were performed, including removal of those men from the analyses who were missing data on any of the characteristics used by the risk calculator, removal of those both alive and fracture free but having less than the full 10 years of follow-up, and using all available follow-up information for those 418 men who began bisphosphonate therapy. All significance levels reported were two-sided and all analyses were conducted using SAS version 9.2 (SAS Institute Inc, Cary, NC) or Stata version 12.0 (StataCorp, College Station, TX).

RESULTS

Of the 5,891 men included in this analysis, 161 (2.7%) had an incident hip fracture and 374 (6.4%) had an incident MOF. The men were followed for an average of 8.4 ± 2.3 years. Of those 4671 men with less than 10 years of follow-up, 337 had an incident MOF before end of follow-up, 1471 died without having an incident MOF, and 2863 men simply had less than full follow-up time. Among the latter group, 221 had terminated from MrOS and the remainder were still participating in the study with an average ($\pm$ SD) follow-up time of 9.1 ± 1.2 years, ranging from 7.3-9.99 years. As previously shown, the cohort was predominately white (89.6%) with an average age at enrollment of 73.5 ± 5.8 year and a femoral neck BMD Z-score $+0.2$. [8].

Data was complete for all risk factors passed to the risk calculator for 4,291 (72.9%) men. [8] The factors most commonly missing data were information on parental history of hip fracture (23.9%) and information on medications used for the definition of corticosteroid use and secondary osteoporosis (4.1% missing). Other risk factors had missing data for nine or fewer men.

The mean ($\pm$ SD) FRAX 10-year predicted fracture probabilities for hip and MOF without BMD in the model were 3.5% (3.6) and 8.9% (4.6), respectively. Addition of BMD to the calculation reduced these estimates to 2.3% (3.1) and 7.6% (4.3), respectively.

Table 1 shows baseline characteristics of the cohort by NOF treatment cut point thresholds. In contrast to the 22% of the cohort who met or surpassed the hip fracture threshold, only 2% met or surpassed the MOF threshold. As expected, many of the fracture risk factors were statistically significantly different between those at or above versus those below these thresholds; these included age, BMI, prior fracture, parent with hip fracture, corticosteroid use, rheumatoid arthritis, and BMD.

Table 2 compares fracture incidence rates per 1000 patient years in the study cohort to published rates for men employed by US-FRAX. [17]. Of all hip fractures, 93% occurred in men 75 years and older; of all MOFs 86% occurred in this age group. Among men in this older age group, the rate of hip fracture was approximately 35-38% lower among men in MrOS compared to the National Inpatient Survey (NIS). The rates of MOF in MrOS were quite close to the calculated rates offered in the same publication.

Overall, the median predicted FRAX 10-year fracture probabilities (using BMD) were 1.4% and 6.5% for hip fracture and MOF, respectively, while observed median cumulative incidences were 3.0% and 6.9%. The tool's predictive accuracy is further shown in Figures

1A and 1B, where relationships between observed cumulative incidences and predicted probabilities of fracture at 10 years are examined within each quintile of estimated risk. For hip fracture without BMD, the tool estimated risk fairly accurately (the average of observed to predicted ratios across the 5 quintiles was 1.0; range of these 5 ratios was 0.9-1.1), whereas this same calculation with BMD yielded an average ratio of 1.3 due to the tool's underestimation in 4 of the 5 risk quintiles (range of these 5 ratios was 0.4-2.0). In contrast, for MOF, FRAX generally overestimated risk yielding an average of observed to predicted ratios across 5 quintiles of 0.8 without BMD and 0.9 with BMD in the calculation (range of these 5 ratios was 0.7-0.9 and 0.7-1.1, for BMD excluded and included, respectively).

We next examined the discrimination performance of FRAX. Figures 2A and 2B show the ROCs for hip and major osteoporotic fracture; the AUC c-statistics were significantly higher both for hip fracture (0.77 versus 0.69; $p<.001$), and for major osteoporotic fracture (0.67 versus 0.63, $p<.001$), for calculations with BMD versus no BMD, respectively. The hip fracture c-statistics, both with and without BMD inclusion, were lower among those 75 and older (approximately half the cohort) versus 65-74 years (data not shown).

Superimposed on the ROC curves are the points where for the NOF treatment thresholds fall (3% for hip and 20% for MOF); without including BMD in the calculations, the NOF cut points correspond to sensitivities/1-specificities of 0.68/0.39 for hip fracture and 0.06/0.03 for MOF, respectively.

Using NOF cut points to calculate the contribution of BMD to FRAX by the Net Reclassification Index (NRI), we found appropriate movement (from higher to lower NOF risk category) among those who did not go on to have a hip fracture ($\text{NRI}=18.3\%$, $P<.001$). However, adding BMD to the model resulted in a tendency for inappropriate movement (from high to low risk category) for those who subsequently had an incident hip fracture ($\text{NRI}= -6.9\%$, $P=.09$). The overall NRI for hip fracture was thus 11.4% ($P=.006$) (Table 3A). In contrast, for MOF, adding BMD resulted in a statistically significant but minor appropriate reclassification of those without an incident fracture ($\text{NRI}=0.9\%$, $P<.001$) and no improvement in overall NRI or NRI specific to those who fractured. (Table 3B).

Using the no category NRI approach, BMD addition significantly ($P=.03$; $P<.001$) reduced the NRI among those going on to suffer either an incident hip fracture or a MOF. (Table 4) This was especially noted among older men with subsequent hip fracture; among those 75 years and older, addition of BMD to the FRAX calculation resulted in a negative NRI (-33.6%) whereas among those less than 75 years, BMD addition produced a positive NRI (27.3%). In contrast, among those subsequently found to be free of incident fracture, addition of BMD significantly ($P<.001$; $P<.001$) increased the NRI by appropriately lowering risk levels for both hip fracture and MOF. Thus, BMD addition significantly ($P<.001$; $P<.001$) increased the overall NRI (45.6% and 24.3% for hip fracture and MOF, respectively) but only through appropriate movement of risk downward for those free of incident fracture.

Our findings were not altered in the two sensitivity analyses removing men with any missing risk factor data and allowing for full follow-up of those men who began bisphosphonate use. After removing those 2863 men who had less than 10 years of follow up but had neither died nor suffered an incident fracture, NRI scores and area under the ROC curves were also similar to the primary analyses; as expected, removing from analyses a large number of men who were fracture free, artificially increased the calculated proportions with fracture.

CONCLUSIONS

We found that the FRAX risk calculator without BMD calibrates well to hip fractures but less well to major osteoporotic fractures observed among older men in the MrOS cohort. However, addition of BMD to the FRAX calculation lowers risk estimates (by 34% for hip and by 15% for MOF) and has inconsistent effects on the tool's fracture prediction performance. On one hand BMD inclusion increases the AUC c-statistic and improves discrimination among those who do not subsequently fracture, while on the other hand it both negatively affects calibration (observed versus predicted probabilities) and assignment of appropriate higher risk to those who ultimately go on to fracture.

Assessing the performance of a risk tool involves several assumptions. Tests of calibration involve grouping individuals into risk categories (e.g. quintiles or clinical value categories) and observing relationships between the median of each predicted risk category and the observed cumulative incidence of fracture after accounting for competing mortality within that category. Conversely, tests of discrimination examine individuals and the correctness of risk categorization based on subsequent fracture versus no fracture status; this approach can use cut points for categorization, if available. If not, no-category analyses are appropriate. Finally, area under the ROC curve examines the relationships between test sensitivity and specificity across all levels of risk.

Statistical methods used to evaluate risk tools have become more sophisticated in recent years and we have learned of the limitations of previous methods [14-16, 18]. Whereas AUC was the most widely used method to assess the performance of risk tools, its limitations have been recognized [18]. Reclassification analyses have been suggested as the preferred method to assess the impact of added input variables on risk model performance. Additionally, limitations of comparative fracture assessment studies have been summarized [19]. In designing our study, we were cognizant of these issues and aware of the newer analytic approaches. First, we avoided making comparisons across studies or tools--- focusing on one risk tool and one cohort. Second, we did not attempt to create new, simpler tools for comparison, but instead examined FRAX performance with and without inclusion of its BMD parameter. Third, we calculated cohort fracture probabilities over many years rather than extrapolating from short term incidence rates; having a large cohort with most men followed up to 10 years provided hundreds of fractures and robustness to our analyses. Recognizing the insensitivity of area under ROC curves [18] to the inclusion of any additional risk factor, we performed NRI studies to assess BMD's contribution to risk levels of individuals. Finally, when calculating fracture probabilities, we adjusted for the known relationship between FRAX score and death hazard. [19]

Our study finding confirms other reports, both in men [20] and women [21, 22], that adding BMD to FRAX yields lower predicted fracture risk estimates. How does this contribute to the tool's performance? We [8] and others [23] have found that addition of BMD increases AUC, yet in the current study we found that calibration and other measures of discrimination (e.g. NRI) appeared no better (or even worse) when BMD was included in the calculation. By reducing risk estimates, BMD reduced the likelihood of an inappropriately high risk estimate in the large proportion of men who did not fracture; however, this also reduced its ability to identify those going on to fracture. The net effect of lower risk estimates was "improvement" because the proportion not fracturing was many times greater than those fracturing--- indeed >90% of the MrOS cohort did not sustain an outcome fracture during follow-up. Further, the better discrimination of hip fracture over MOF that we observed could be driven by hip fracture incidence being less than half that of major fractures. Thus, correct reassignment performance in this cohort with BMD appears to be a trade-off

between inappropriate lower risk estimates among the few destined to fracture offset by appropriate lower risk estimates among the many not going on to fracture.

Our study has several strengths. The baseline input data are fairly complete on a large number (nearly 4300) of men. The follow-up for incident fracture is 99% complete and self-reported fractures were validated by review of radiograph reports; fractures were recorded up to 10 years and the majority of subjects had length of follow-up close to the 10-year value used by in the risk calculation, thus obviating the need to extrapolate. We performed several sensitivity analyses, confirming the robustness of our results despite missing data or truncated follow-up.

Limitations include drug exposure information based only on the 30 days prior to baseline and no information on dosage of glucocorticoids. We did not exclude fractures on the basis of degree of trauma; there are several reasons for this. First, determining degree of trauma is challenging. Second, FRAX 10-year estimates are for fractures of any degree of trauma, thus making comparisons of observed and predicted fracture events more valid. Finally, Mackey et al [24] showed that low and high trauma fractures observed in both older men (MrOS) and older women (SOF) were similarly associated with low BMD.

MrOS is an epidemiological observational study and a healthy volunteer bias could be present. The fact that the group's mean BMD Z-score was +0.2 suggests that these men are at substantially less risk for fracture (assuming RR 2.0 per SD BMD, the +0.2 Z-score translates into 40% lower fracture risk). Our finding that hip fracture incidences among the MrOS cohort are somewhat lower than those obtained from the US NIS (rates used by FRAX) supports this contention. The MOF rates used by FRAX are the product of cobbling together NIS hip fracture rates with wrist and humerus fracture rates from Rochester, MN, to which estimates of clinical spine fracture rates are added based on ratios of clinical spine to hip fracture incidence in Swedish data, and finally adjusting the MOF rate downward to account for overlap of multiple fractures in the same individuals. [17] Despite the multiple assumptions and opportunities for miscalculation in these data, the MOF incidences among the older men in the MrOS cohort are remarkably close to the base fracture rates used by FRAX. We excluded 1.7% of the cohort at baseline because they reported taking bisphosphonates – thus making the remaining cohort less prone to fracture than the original cohort, but those remaining also had lower fracture risk scores as a result of the winnowing.

We believe that FRAX performance or the performance of other fracture risk prediction tools are best evaluated using several measures of both calibration and discrimination, as well as newer measures of risk reclassification. Furthermore, these evaluations should be performed in a population similar to the target population in which the tool will be applied. The results of our study, taken in the context of other work in this area, indicate that age and sex greatly affect a fracture risk tool's performance. This effect could be explained by marked differences in observed fracture incidence rates in various populations.

In conclusion, among community-dwelling, ambulatory, older men in the MrOS cohort, the FRAX calculator without BMD calibrates reasonably well with both hip and major osteoporotic fractures observed. Also without BMD, the tool's ability to discriminate between men who will and will not experience these fracture outcomes appears adequate only for population management, i.e. categorizing groups by risk level. However, addition of BMD to the tool's calculations shows differential effects on fracture risk estimates and does little to improve its predictive performance.

These issues require explanation and call for care in the way we evaluate fracture risk tool performance.

Acknowledgments

Funding/Support: The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Institute on Aging (NIA), the National Center for Research Resources (NCRR), and NIH Roadmap for Medical Research under the following grant numbers: U01 AR45580, U01 AR45614, U01 AR45632, U01 AR45647, U01 AR45654, U01 AR45583, U01 AG18197, U01-AG027810, and UL1 RR024140. Dr. Curtis receives support from National Institutes of Health AR053351.

References

1. [April 30, 2012] FRAX. www.shef.ac.uk/FRAX/
2. Kanis JA, Oden A, Johnell O, Johansson H, De Laet C, Brown J, Burckhardt P, Cooper C, Christiansen C, Cummings S, Eisman JA, Fujiwara S, Glüer C, Goltzman D, Hans D, Krieg MA, La Croix A, McCloskey E, Mellstrom D, Melton LJ 3rd, Pols H, Reeve J, Sanders K, Schott AM, Silman A, Torgerson D, van Staa T, Watts NB, Yoshimura N. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. *Osteoporos Int*. 2007; 18(8):1033–1046. [PubMed: 17323110]
3. Dawson-Hughes B. A revised clinician's guide to the prevention and treatment of osteoporosis. *J Clin Endocrinol Metab*. 2008; 93(7):2463–2465. [PubMed: 18544615]
4. Osteoporosis - Clinical guideline for prevention and treatment. Executive Summary. National Osteoporosis Guideline Group (NOGG) on behalf of the Bone Research Society, British Geriatrics Society, British Orthopaedic Association, British Society of Rheumatology, National Osteoporosis Society, Osteoporosis 2000, Osteoporosis Dorset, Primary Care Rheumatology Society, Royal College of Physicians and Society for Endocrinology; Updated July 2010. www.sheffield.ac.uk/NOGG/downloads.html [April 30 2012]
5. [April 30, 2012] www.WHO-FRAX.org
6. Yun H, Delzell E, Ensrud KE, Kilgore ML, Becker D, Morrissey MA, Curtis JR. Predicting hip and major osteoporotic fractures using administrative data. 2010; 170:1940–1941.
7. Leslie WD, Lix LM, Johansson H, Oden A, McCloskey E, Kanis JA. Manitoba Bone Density Program. Independent clinical validation of a Canadian FRAX tool: fracture prediction and model calibration. *J Bone Miner Res*. 2010; 25(11):2350–2358. [PubMed: 20499367]
8. Ettinger B, Liu H, Blackwell T, Hoffman AR, Ensrud KE, Orwoll ES. for the Osteoporotic Fracture in Men (MrOS) Research Group. Validation of FRC, a fracture risk assessment tool, in a cohort of older men: The Osteoporotic Fracture in Men (MrOS) Study. *J Clin Densitom*. 2012 epub ahead of print.
9. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, Lewis C, Cawthon PM, Marcus R, Marshall LM, McGowan J, Phipps K, Sherman S, Stefanick ML, Stone K. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. *Contemp Clin Trials*. 2005; 26(5):569–585. [PubMed: 16084776]
10. Blank JB, Cawthon PM, Carrion-Petersen ML, Harper L, Johnson JP, Mitson E, Delay RR. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials*. 2005; 26:557–568. [PubMed: 16085466]
11. Looker AC, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP, Johnston CC Jr, Lindsay R. Updated data on proximal femur bone mineral levels of US adults. *Osteoporos Int*. 1998; 8(5): 468–489. [PubMed: 9850356]
12. Pahor M, Chrischilles EA, Guralnik JM, Brown SL, Wallace RB, Carbonin P. Drug data coding and analysis in epidemiologic studies. *Eur J Epidemiol*. 1994; 10(4):405–411. [PubMed: 7843344]
13. Tosteson AN, Melton LJ 3rd, Dawson-Hughes B, Baim S, Favus MJ, Khosla S, Lindsay RL. National Osteoporosis Foundation Guide Committee. Cost-effective osteoporosis treatment thresholds: the United States perspective. *Osteoporos Int*. 2008; 19(4):437–447. [PubMed: 18292976]

14. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011; 30(1):11–21. [PubMed: 21204120]
15. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med*. 2009; 150(11):795–802. [PubMed: 19487714]
16. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008; 27(2):157–172. [PubMed: 17569110]
17. Ettinger B, Black DM, Dawson-Hughes B, Pressman AR, Melton LJ 3rd. Updated fracture incidence rates for the US version of FRAX. *Osteoporos Int*. 2010; 21(1):25–33. [PubMed: 19705048]
18. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007; 115(7):928–935. [PubMed: 17309939]
19. Kanis JA, Oden A, Johansson H, McCloskey E. Pitfalls in the external validation of FRAX. *Osteoporos Int*. 2012; 23:423–431. [PubMed: 22120907]
20. Schwartz AV, Vittinghoff E, Bauer DC, Hillier TA, Strotmeyer ES, Ensrud KE, Donaldson MG, Cauley JA, Harris TB, Koster A, Womack CR, Palermo L, Black DM. Study of Osteoporotic Fractures (SOF) Research Group; Osteoporotic Fractures in Men (MrOS) Research Group; Health, Aging, and Body Composition (Health ABC) Research Group. Association of BMD and FRAX score with risk of fracture in older adults with type 2 diabetes. *JAMA*. 2011; 305(21):2184–2192. [PubMed: 21632482]
21. Donaldson MG, Cawthon PM, Schousboe JT, Ensrud KE, Lui LY, Cauley JA, Hillier TA, Taylor BC, Hochberg MC, Bauer DC, Cummings SR. Study of Osteoporotic Fractures (SOF). Novel methods to evaluate fracture risk models. *J Bone Miner Res*. 2011; 26(8):1767–1773. [PubMed: 21351143]
22. Pressman AR, Lo JC, Chandra M, Ettinger B. Methods for Assessing Fracture Risk Prediction Models: Experience with FRAX in a large integrated healthcare delivery system. *J Clin Densitometry*. 2011; 14(4):407–415.
23. Adler RA, Hastings FW, Petkov VI. Treatment thresholds for osteoporosis in men on androgen deprivation therapy: T-score versus FRAX. *Osteoporos Int*. 2010; 21:647–653. [PubMed: 19533207]
24. Mackey DC, Lui LY, Cawthon PM, Bauer DC, Nevitt MC, Cauley JA, Hillier TA, Lewis CE, Barrett-Connor E, Cummings SR. Study of Osteoporotic Fractures (SOF) and Osteoporotic Fractures in Men Study (MrOS) Research Groups. High-trauma fractures and low bone mineral density in older women and men. *JAMA*. 2007; 298(20):2381–2388. [PubMed: 18042915]

A.

B.

Fig 1.

A. Cumulative (%) incidence of hip fracture (accounting for competing risk of mortality) versus predicted (%) FRAX probability of hip fracture, with and without BMD, by quintiles of FRAX 10-year fracture probability estimates.

B. Cumulative (%) incidence of major osteoporotic fracture (MOF) (accounting for competing risk of mortality) versus predicted (%) FRAX probability of MOF, with and without BMD, by quintiles of FRAX 10-year fracture probability estimates.

B. **Fig 2.**

A. Receiver Operating Characteristic (ROC) curves for FRAX-predicted hip fracture risk.

The solid line represents the FRAX estimates without BMD used in the risk calculator,

dashed lines with BMD used in the calculator. The thick solid line is the line of no

discrimination. The circles represent the NOF treatment cut points.

B. Receiver Operating Characteristic (ROC) curves for FRAX-predicted major osteoporotic

fracture (MOF) risk. The solid line represents the FRAX estimates without BMD used in the

risk calculator, dashed lines with BMD used in the calculator. The thick solid line is the line

of no discrimination. The circles represent the NOF treatment cut points.

Baseline characteristics of 5891 MrOS participants, overall and by National Osteoporosis Foundation cost-effectiveness treatment thresholds [ref 13].

Table 1

Variable	FRAX Risk Score for Hip Fracture*				FRAX Risk Score for Major Osteoporotic Fracture*	
	Overall (n=5891)	<3% (n=4589)	3% (n=1301)	<20% (n=5768)	20% (n=122)	
Age, yrs, mean (SD)	73.6 (5.9)	72.5 (5.5)	77.8 (5.2)**	73.6 (5.9)	77.0 (5.0)**	
Caucasian, n (%)	5268 (89.4)	4016 (87.5)	1251 (96.2)**	5150 (89.3)	117 (95.9)	
Body mass index, kg/m ² , mean (SD)	27.4 (3.8)	27.8 (3.9)	25.9 (3.2)**	27.4 (3.8)	26.0 (3.1)**	
Fracture after age 50, n (%)	1302 (22.1)	781 (17.0)	520 (40.0)**	1221 (21.2)	80 (65.6)**	
Parental hip fracture, n (%)	743 (16.6)	325 (9.4)	418 (41.1)**	640 (14.7)	103 (88.0)**	
Rheumatoid arthritis, n (%)	304 (5.2)	201 (4.4)	103 (7.9)**	289 (5.0)	15 (12.3)**	
Current smoker, n (%)	205 (3.5)	105 (3.3)	55 (4.2)	197 (3.4)	8 (6.6)	
Oral corticosteroid use past 30 days, n (%)	104 (1.8)	55 (1.3)	49 (3.9)**	95 (1.7)	9 (7.6)**	
3 or more alcoholic drinks/day, n (%)	237 (4.0)	167 (3.6)	70 (5.4)**	233 (4.1)	4 (3.3)	
Femoral neck BMD, g/cm ² , mean (SD)	0.79 (0.13)	0.82 (0.11)	0.66 (0.09)**	0.79 (0.13)	0.60 (0.08)**	
Femoral neck BMD T-score, n (%)						
Normal (−1.0 or greater)	2458 (41.7)	2366 (51.6)	92 (7.1)**	2457 (42.6)	1 (0.8)**	
Osteopenia (−2.5 to −1.0)	3150 (53.5)	2202 (48.0)	948 (72.9)	3078 (53.4)	72 (59.0)	
Osteoporosis (less than −2.5)	282 (4.8)	21 (0.5)	261 (20.1)	233 (4.0)	49 (40.2)	
Femoral neck BMD Z-score, mean (SD)	0.21 (0.93)	0.43 (0.87)	−0.56 (−0.70)**	0.24 (0.92)	−1.05 (−0.68)**	

* FRAX fracture probability calculation includes BMD.

** P-value <0.05

Table 2

Incident hip and major osteoporotic fracture (MOF) rates per 1000 patient years by 5 year age categories for white men in MrOS study cohort compared to US FRAX reference rates.

Age (years)	Hip Fractures			Major Osteoporotic Fractures (MOF)		
	No. Observed	MrOS Rate Per 1000 pt yrs	US FRAX Reference Rate* Per 1000 pt yrs	No. Observed	MrOS Rate Per 1000 pt yrs	US FRAX Reference Rate* Per 1000 pt yrs
65-69	1	0.3	1.2	9	2.6	4.0
70-74	11	1.0	2.1	49	4.5	5.5
75-79	34	2.5	4.0	94	6.9	7.5
80-85	53	5.0	8.1	117	11.4	11.2
85-89	76	10.6	16.3	144	21.0	25.2

* ref 17 (Osteoporos Int 2009)

Table 3**A. Assignments to risk categories based on FRAX 10-year hip fracture probability**

Fracture Probability Without BMD	Fracture Probability With BMD		
	<3%	3%	Total
<3%			
Men, n	3351	181	3532
Fractured, n	35	16	51
Not fractured, n	3316	165	3481
Percentage of men with fracture	1.04	8.84	1.44
3%			
Men, n	1238	1120	2358
Fractured, n	27	82	109
Not fractured, n	1211	1038	2249
Percentage of men with fracture	2.18	7.32	4.62
Total			
Men, n	4859	1301	5890
Fractured, n	62	98	160
Not fractured, n	4527	1203	5730
Percentage of men with fracture	1.35	7.53	2.72

B. Assignments to risk categories based on FRAX 10-Year major osteoporotic fracture (MOF) probability.

Fracture Probability Without BMD	Fracture Probability With BMD		
	<20%	20%	Total
<20%			
Men, n	5654	60	5714
Fractured, n	347	5	352
Not fractured, n	5307	55	5362
Percentage of men with fracture	6.14	8.33	6.16
20%			
Men, n	114	62	176
Fractured, n	11	10	21
Not fractured, n	103	52	155
Percentage of men with fracture	9.65	16.13	11.93
Total			
Men, n	5768	122	5890
Fractured, n	358	15	373
Not fractured, n	5410	107	5517
Percentage of men with fracture	6.21	12.30	6.33

A. No-Category NRI for Adding Bone Mineral Density (BMD) to FRAX Calculations of Hip and of Major Osteoporotic Fracture.

Table 4

Adding BMD to Risk Estimate Equation:						
	Increased risk score	Decreased risk score	No change in risk score	Total	100*NRI Score	P-value
Hip						
Men, n	1129	4736	25	5890	45.6	<0.001
Fractured, n	66	93	1	160	-16.9	0.03
Not fractured, n	1063	4643	24	5730	62.5	<0.001
Major Osteoporotic Fracture						
Men, n	1526	4353	11	5890	24.3	<0.001
Fractured, n	139	233	1	373	- 25.2	<0.001
Not fractured, n	1387	4120	10	5517	49.5	<0.001