



Published in final edited form as:

J Clin Densitom. 2014 ; 17(4): 449–457. doi:10.1016/j.jocd.2013.09.020.

Prediction Models of Prevalent Radiographic Vertebral Fractures Among Older Men

John T. Schousboe, MD, PhD^{1,2}, Harold R. Rosen, MD³, Tamara J. Vokes, MD⁴, Jane A. Cauley, DrPH⁵, Steven R. Cummings, MD⁶, Michael C. Nevitt, PhD⁷, Dennis M. Black, PhD⁷, Eric S. Orwoll, MD⁸, Deborah M. Kado, MD, MS⁹, Kristine E. Ensrud, MD, MPH^{10,11}, and For the Osteoporotic Fractures in Men (MrOS) Study Research Group

¹Park Nicollet Osteoporosis Center and Institute for Research and Education, Minneapolis, MN

²Division of Health Policy and Management, University of Minnesota

³Beth Israel Deaconess Medical Center, Boston, MA

⁴University of Chicago, Chicago, IL

⁵Department of Epidemiology, Graduate school of Public Health, University of Pittsburgh, Pittsburgh, PA

⁶San Francisco Coordinating Center, California Pacific Medical Center, San Francisco, CA

⁷Division of Epidemiology and Biostatistics, University of California at San Francisco, San Francisco, CA

⁸Oregon Health and Science University, Portland, OR

⁹University of California at San Diego, San Diego, CA

¹⁰Division of Epidemiology, University of Minnesota, Minneapolis, MN

¹¹Minneapolis Veterans Administration Medical Center, Minneapolis, MN

Abstract

No studies have compared how well different prediction models discriminate older men who have a radiographic prevalent vertebral fracture (PVFx) from those who do not. We used area under receiver operating characteristic (AUROC) curves and a net reclassification index to compare how well regression-derived prediction models and non-regression prediction tools identify PVFx among men age 65 years with femoral neck T-score -1.0 enrolled in the Osteoporotic Fractures in Men (MrOS) Study. The AUROC for a model with age, bone density (BMD), and historical height loss (HHL) was 0.682 compared to 0.692 for a complex model with age, BMD, HHL, prior non-spine fracture, body mass index, back pain, grip strength, smoking, and

© 2013 International Society for Clinical Densitometry. Published by Elsevier Inc. All rights reserved.

Corresponding Author: John T. Schousboe, MD, PhD, Park Nicollet Institute for Research and Education, 3800 Park Nicollet Blvd., Minneapolis, MN 55416, Phone: 952-993-3877, Fax: 952-993-1312, schouj@parknicollet.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

glucocorticoid use (p-values for difference in five bootstrapped samples 0.14 to 0.92). This complex model, using a cutpoint prevalence of 5%, correctly re-classified only a net 5.7% (p-value 0.13) of men as having or not having a PVFx compared to a simple criteria list (age \geq 80 years, HHL $>$ 4 cm, or glucocorticoid use). In conclusion, simple criteria identify older men with PVFx as well as regression-based models. Future research to identify additional risk factors that more accurately identify older men with PVFx is needed.

Keywords

prevalent vertebral fracture; prediction models; model discrimination; vertebral fracture assessment; bone densitometry

Introduction

Radiographic prevalent vertebral fractures (PVFx) are present in over 10% of older men,(1) are a marker of bone fragility and high fracture risk,(2–7) and yet are frequently unrecognized in clinical practice.(8) Vertebral fracture assessment either with densitometric lateral spine images(9) or standard spine radiographs(10) is a cost-effective method to identify many who have a (PVFx) who otherwise might not be recognized as being at high risk of subsequent fractures and in need of treatment. A barrier to implementation of case finding strategies to identify those with PVFx may be the complexity of existing guidelines. (11) While many risk factors have been shown to be independently associated with PVFx in multivariable-adjusted regression models,(12–22) only a few of these studies included significant numbers of men.(18, 21–23) Moreover, studies to date have not investigated whether or not prediction models incorporating most or all of these risk factors perform better than more parsimonious models that might be easier to implement in clinical practice. Additionally, it is unclear if regression-based prediction models more accurately discriminate men with from men without PVFx than simple lists of indications for lateral spine imaging. Regression-based prediction models are more difficult to implement into clinical practice, because they require a complex calculator to render an estimated risk of PVFx being present from parameter inputs. In contrast, a non-regression based list of indications simply requires the practitioner to know if the patient has one or more indications for lateral spine imaging.

Our objectives were two-fold; a) examine how well nested regression models discriminate men with low bone mass who have radiographic PVFx from those who do not among men with low bone mass enrolled in the Osteoporotic Fractures in Men (MrOS) Study, using area under receiver operating characteristic (AUROC) curves and the net reclassification improvement (NRI) statistic of Pencina;(24) and b) examine how well, again using AUROC analyses and the NRI statistic, these regression models discriminate those with, from those without radiographic PVFx, compared to simple lists of indications for lateral spine imaging.

Materials and Methods

The Osteoporotic Fractures in Men (MrOS) Study enrolled 5,994 men in 2000 to 2002 in six metropolitan areas of the United States (Birmingham AL, Minneapolis MN, Palo Alto CA, Pittsburgh PA, Portland OR, and San Diego CA). Methods of study recruitment have been described previously.(25, 26)

Identification of Prevalent Radiographic Vertebral Fractures

Lateral lumbar and thoracic spine radiographs were obtained for all men at the baseline study visit with an x-ray tube to film distance of 40 inches using a breathing technique, with thoracic x-rays centered at T7 and lumbar x-rays centered at L3. All technically adequate x-rays (for 5,958 men) were scanned to digital format. A triage process that has been validated in other populations was used to identify those (2,745 men) with unequivocally normal x-rays.(27) The lateral spine radiographs of the remaining 3,213 men were evaluated by an expert reader (JTS) for vertebral fracture using the Genant semi-quantitative criteria,(28) with the revision that those vertebrae judged visually to be mild (grade 1) deformities had to have endplate depression or obvious cortical buckling to be considered fractured. Intra-reader reliability was evaluated with kappa statistic on four occasions in this process, and ranged from 0.79 to 0.92.

We excluded those without a prior spine fracture (n=242) from our analyses. While patient self-reports of vertebral fracture are not as accurate as self-reports of prior hip fracture,(29) the positive predictive value of a positive self-report of vertebral fracture may be as high as 85%.(30) We chose to further restrict our analyses to the population with a femoral neck T-score ≥ -1.0 (n = 3, 271) according to young male reference norms because to our knowledge there is no evidence regarding the efficacy of currently available fracture prevention therapies in those with normal BMD.

Measurement of bone mineral density

Bone mineral density was measured at the femoral neck and total hip with QDR-4500 fan-beam densitometers (Hologic, Bedford, MA, USA), at the baseline MrOS visit. Central training of densitometry technologists and cross-calibration of densitometers across study centers with a phantom was done to ensure consistency and quality of bone mass measurement.(36)

Measurement of other covariates

At the baseline visit, all MrOS participants were asked their height at age 25, and height loss was calculated as the difference between recalled young height and height measured at the baseline visit with a Harpenden stadiometer. Current weight was measured at the baseline visit with a balance beam or electronic scale, and body mass index calculated weight (kg) divided by height(meters) squared. Participants were asked if they experienced any fractures since age 50, and if so, the skeletal location of the fracture(s). They were asked whether or not they were currently smoking cigarettes, whether or not they were taking glucocorticoids, and whether or not they were on anti-androgenic drugs or had had an orchiectomy. Grip

strength was measured using Jamar dynamometers.(37) They were asked if they had lower back pain over the past year and to rate how much it limited their activity.

Selection of Covariate Predictors for Regression-Based Prediction Models

Virtually all studies published to date have identified age and bone mineral density (BMD) as predictors of PVFx, and hence we chose age and femoral neck BMD as our first regression model (Model 1) to test in the subset with no self-reported vertebral fracture as of visit 3. Historical height loss (recalled young adult height minus current height) has also been consistently identified as an independent risk factor for PVFx,(15, 17, 19, 31, 32) and is a stand-alone indication for vertebral fracture assessment in the 2007 International Society for Clinical Densitometry (ISCD) Position Statement for densitometric vertebral fracture assessment (VFA) indications.(11) Hence, Model 2 for comparison included age, femoral neck BMD and historical height loss as predictors.

Model 3 included age, femoral neck BMD, historical height loss, prior non-vertebral fracture, body mass index (BMI), grip strength, and self-reported back pain as predictors. Prior fracture is a secondary indication (when combined with age) in the 2007 ISCD *indications for VFA*,(11) and *BMI has been identified in some studies* (13, 15, 17, 20), but not others (18, 21, 22), as a risk factor for vertebral fracture. Other studies have identified back pain to be associated with PVFx in women,(22, 32–35) and two have identified grip strength as to be associated with PVFx.(12, 22) Model 4, the most complex model, included all the covariates of Model 3 plus current glucocorticoid use and current smoking as predictors.

Covariates Included in Different Lists of Indications for Lateral Spine Imaging

The 2007 ISCD criteria for vertebral fracture assessment for men were the following;(11) 1) age ≥ 80 years; 2) historical height loss > 6 cm; 3) current glucocorticoid use; or 4) two of the following (age 70 to 79 years combined with prior non-vertebral fracture, height loss > 3 but ≤ 6 cm, prior orchiectomy, current androgen deprivation therapy). The three sets of simple criteria that we chose to test were the following: a) age ≥ 80 years, historical height loss > 6 cm or current glucocorticoid use [Simple 1]; b) age ≥ 75 years, historical height loss > 6 cm or current glucocorticoid use [Simple 2]; or c) age ≥ 80 years, historical height loss > 4 cm or current glucocorticoid use [Simple 3].

Statistical Analyses

The primary analyses were done using logistic regression models with semi-quantitative (SQ) grade 2 or grade 3 PVFx as the dependent variable in those with a femoral neck T-score (using young male reference data) at the baseline visit of ≥ -1.0 . Five sets of secondary analyses were done; one with fractures of all SQ grades as the dependent variable, one restricted to just those with a femoral neck T-score ≥ -1.0 but > -2.5 , a third set including those within all levels of BMD, and a fourth set substituting spine for femoral neck BMD. A fifth set of secondary analyses were done where we tested if non-linear predictors might improve model discrimination. We did this in two ways; a) we added age squared and interaction terms between age and femoral neck BMD, age and height loss, height loss and BMD, and height loss and prior non-spine fracture to the models; and b)

modeled continuous variables as four-level categorical variables. For all regression models, model fit and calibration was tested with the Hosmer-Lemeshow test, and model specification with Pregibon's linktest.(38)

Because AUROC statistics that are derived in the same samples in which they were produced can be overinflated (unless the sample size is very large),(39) we produced five bootstrapped models for each of the four parent regression models, and compared the AUROC statistic between the nested models for each of five pairs of bootstrapped samples.

While AUROC statistics give an overall assessment of model discrimination across the entire range of pre-test probability of the dependent variable, lateral spine imaging for PVFx is likely to be cost-effective in populations with relatively modest or even low prevalence of vertebral fracture.(9, 10) Net reclassification indexes are a method of testing how well two prediction rules discriminate those who have from those who do not have an outcome at a set prevalence of that outcome. Suppose, for example, that clinicians choose to get spine imaging on anyone who has a pre-test probability $\geq 10\%$ of having a radiographic PVFx. In this instance, true positives would be those with a model predicted probability of having a prevalent vertebral fracture $\geq 10\%$ who truly have one, and true negative would be those who have a model predicted probability of having a radiographic PVFx $< 10\%$ who in fact do not have one. False positives and false negatives are, respectively, those with a $\geq 10\%$ probability of having a radiographic PVFx who do not have one, and those with a $< 10\%$ probability of having a radiographic PVFx who in fact do have one. The proportion of individuals correctly classified by the model is the sum of true positives and true negatives divided by the total sample number. By the Pencina method,(24) the NRI using a second model instead of a first model is the proportion who are shifted from being incorrectly classified to *correctly classified* using Model 2 instead of Model 1, minus the proportion who are shifted *from being correctly classified* to incorrectly classified using Model 2 instead of Model 1. We compared nested models with NRI at pre-test probability cutpoints of 5%, 10%, and 15%.

To better understand the practical impact of using any of the four prediction models to decide who should have lateral spine imaging to detect PVFx, we calculated for each of the four regression models using prevalence cutpoints of 5%, 10% or 15%: a) the proportion who would be chosen to have lateral spine imaging; b) the proportion of men with a PVFx who would be detected; and c) among those who did receive a lateral spine image, the proportion who would have one or more PVFx.

We used this NRI method to compare how well these regression models correctly classified PVFx status compared to a list comprising nearly all of the indications for lateral spine imaging for men in the 2007 ISCD Position Statement. We also compared how well the regression-based prediction models and 2007 ISCD criteria classified those with, and those without PVFx (among those with a femoral T-score of ≥ -1.0), relative to the three simple lists of indications (Simple 1, Simple 2, and Simple 3), using Stata 12.0.

Results

Among all 5,958 men with evaluable lateral spine radiographs, 689 (11.6%) had one or more PVFxs; 448 (7.5%) had a moderate or severe (SQ grade 2 or 3) PVFxs. Two hundred forty two (242, or 4.1%) self-reported a history of vertebral fracture at the baseline visit, and of these, 59% had a PVFxs of any grade, and 50% had a moderate or severe PVFxs. Hence, we reasoned that a self-reported (but undocumented) history of prior vertebral fracture would be a reasonable stand-alone indication for lateral spine imaging, and did not include these men in subsequent analyses, as well as those with a femoral neck T-score > -1.0 .

The characteristics of the remaining 3,271 men are shown in Table 1; those who had a radiographic PVFxs were older, had lower bone mineral density, had more historical height loss, lower grip strength, and slightly higher body mass index. Those with one or more radiographic PVFxs were also more likely to have had a self-reported prior non-spine fracture since age 50, to have had back pain, and to be on glucocorticoid therapy.

Among the 3,271 men with femoral neck BMD T score ≥ -1 and no self-reported prior spine fracture, the associations of potential predictor variables with SQ grade 2 or 3 radiographic PVFxs within the four nested models is shown in Table 2. Lower BMD, greater height loss, prior non-vertebral fracture, higher BMI, lower grip strength, and current glucocorticoid use were all independently associated with prevalent PVFxs. In all bootstrapped model comparisons, the AUROC of Model 2 based on age, femoral neck BMD and historical height loss (AUC range 0.676–0.681) was superior to that of Model 1 based on age and femoral neck BMD alone (AUC range 0.638 – 0.642, χ^2 range 8.5 to 16.5, p-value range <0.001 to 0.003) indicating modestly better discrimination of whether or not a PVFxs was present (Table 3). Moreover, 5.9% and 4.9% of men had a net correct re-classification of PVFxs status using Model 2 instead of Model 1 at, respectively, pre-test probability cutpoints of 5% and 10%. Models 3 and 4 did not significantly discriminate those with compared to those without PVFxs compared to model 2 by either AUROC analyses or NRI at any of the three pre-test probability cutpoints. Repeating these analyses with radiographic PVFxs of all SQ grades, restricting the analyses to men with a femoral neck T-score of -1.0 to -2.4 , or including all 5,712 men regardless of baseline BMD level did not alter these results (data not shown). Adding age squared, interaction terms between age and BMD, age and height loss, height loss and BMD, and height loss and prior fracture did not improve model discrimination for any of the four models by AUROC analyses (data not shown). Similarly, modeling age, femoral neck BMD, height loss, or grip strength as four level categorical variables also did not improve model discrimination (data not shown).

There was a wide range of values for the calculated proportion who would be screened, the proportion of men with one or more moderate to severe radiographic PVFxs who would be detected, and the prevalence of moderate to severe PVFxs among those who would be screened. This was driven primarily by the prevalence cutpoint of a PVFxs being present chosen to decide whether or not lateral spine imaging should be done (Table 4). As the prevalence cutpoint is raised, a lower proportion of men would receive lateral spine imaging, and a lower proportion of men with PVFxs will be detected. As greater numbers of covariates are added to the prediction model, there are slight changes in the proportion of men who

would be screened, the proportion of men with PVFx who are detected, and among those screened the proportion who have a PVFx detected, but these differences were relatively small.

Using a 5% prevalence cutpoint for a grade 2 or 3 PVFx, regression Model 2 and Model 4 showed a net correct classification PVFx status, respectively, of 5.2% and 10.5% compared to the ISCD 2007 criteria (Table 5). However, regression Models 2 and 4 did not show any better ability to discriminate those with from those without PVFx than simple criteria set 1 (age \geq 80 years, historical height loss $>$ 6 cm, or current glucocorticoid use). However, simple criteria set 1 also only detected 51% of men with a prevalent vertebral fracture. If the age criterion was lowered to \geq 75 years (simple criteria set 2; age \geq 75 years, historical height loss $>$ 6 cm, or current glucocorticoid use), the proportion of men with a PVFx detected rose to 65%, but regression models 2 and 4 both discriminated those with from those without PVFx better than simple criteria set 2 (NRI of 8.7% (11.8%) for Model 2 (Model 4) vs. simple criteria set 2). However, if the height loss criterion was lowered to $>$ 4 cm (simple criteria set 3; age \geq 80 years, height loss $>$ 4 cm, or current glucocorticoid use), the proportion of those with PVFx detected rose to 73%, and the NRI of regression Models 2 and 4 was not statistically significantly better compared to simple criteria set. Simple criteria 3 had an NRI of 5.1% compared to the 2007 ISCD criteria, but this comparison did not quite reach significance (p-value 0.06).

Discussion

Previously undiagnosed prevalent radiographic vertebral fractures are present in a significant minority of men age 65 and older with low bone mass, albeit not as commonly as is seen in age-matched women. Although we confirmed findings of multiple other studies that numerous risk factors, including femoral neck BMD, historical height loss, self-reported prior non-vertebral fractures, BMI, and smoking are each associated with PVFx after multivariable adjustment, regression models incorporating all of these risk factors did not discriminate those with from those without PVFx better than more parsimonious regression models. Based on the findings from our model comparisons using AUROC analyses and the NRI statistic, a model that includes only age, femoral neck BMD, and historical height loss appears to perform better than a model with age and BMD alone, and performs as well as more complex models. More importantly, a simple list of three dichotomous criteria based on clinical risk factors alone and easily assessed in the busy clinical practice setting (age \geq 80 years, historical height loss $>$ 4 cm, or current glucocorticoid use) appears to discriminate those with, from those without PVFx, as well as any of the regression-based prediction models we tested.

There is no consensus as to how high the pre-test probability of a clinically unapparent radiographic vertebral fracture should be before lateral spine imaging is worthwhile. However, based on the low cost of lateral spine imaging and generic therapies now available to treat osteoporosis, and the sensitivity analyses of previously published cost-effectiveness studies, lateral spine imaging may be cost effective even when the prevalence of vertebral fracture in the screened population is as low as 5%.⁽⁹⁾ However, this is predicated on: a) the assumption that detection of a moderate to severe prevalent PVFx would lead to

commencement of or change in fracture prevention medication; and b) that the reader of the lateral spine image has the requisite training and experience to accurately discern vertebral fractures from non-fracture vertebral deformities and from normal vertebrae. Several studies have shown that appropriately trained non-radiologists can accurately interpret densitometric lateral spine images.(28, 40–42)

The prevalence of undetected PVFx among men selected for lateral spine imaging using simple criteria set 3 (low bone mass combined with age \geq 80 years, historical height loss $>$ 4 cm, or current glucocorticoid use) is about 10%, and these three simple criteria discriminate men with low bone mass who have a moderate to severe radiographic PVFx from those who do not as well as regression-based prediction models using currently known risk factors for PVFx. Based on the results of our analyses, these criteria (combined with the stand alone indication of a self-reported but undocumented prior vertebral fracture) may constitute at this time a reasonable set of indications to use in the clinical practice setting for performing densitometric lateral spine imaging or standard radiographs (combined with interpretation by appropriately trained readers) in men age 65 and older with low bone mass.

There are important limitations to our analyses. First and foremost, although we performed internal validation of our prediction models by comparing their AUROC's in bootstrapped samples from the parent study population, both the prediction models and these sets of non-regression based prediction tools for PVFx among men should be externally validated in other studies and populations. Second, although men of all ethnic backgrounds residing in the U.S. were enrolled in MrOS, non-Caucasian men represent only 10.5% of the study population, and we did not have adequate power to examine prediction models in the small non-Caucasian subset of the cohort. Third, the overall predictive power of all of these prediction tools for PVFx was modest, with AUROCs all below 0.7. Research to identify risk factors that more accurately identify older men at high risk of having a PVFx is clearly needed. Finally, it is uncertain whether our findings based on expert reading of radiographs using modified Genant criteria for identification of PVFx is translatable to readings by community based readers.

There are many strengths of our study. MrOS is the largest cohort study of men that includes comprehensive assessment of PVFx with lateral spine radiographs, and MrOS study participants were recruited from large groups and registries closely representative of the Caucasian male population of the United States. Assessment of lateral spine radiographs for PVFx was done with careful attention to definitions of fracture and non-fracture vertebral deformities, and with repeated checks of intra-rater reliability.

In conclusion, a simple list of indications for lateral spine imaging to detect radiographic PVFx (age \geq 80 years, historical height loss $>$ 4 cm, or current glucocorticoid use) performs as well as regression-based prediction models discriminating older men with low bone mass with from those without radiographic PVFx. However, future research is needed to identify additional risk factors that better identify older men with PVFx.

Acknowledgments

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 TR000128.

References

- O'Neill TW, Felsenberg D, Varlow J, Cooper C, Kanis JA, Silman AJ. The prevalence of vertebral deformity in European men and women: the European Vertebral Osteoporosis Study. *J Bone Miner Res.* 1996; 11(7):1010–1018. [PubMed: 8797123]
- Black DM, Arden NK, Palermo L, Pearson J, Cummings SR. Prevalent vertebral deformities predict hip fractures and new vertebral deformities but not wrist fractures. Study of Osteoporotic Fractures Research Group. *J Bone Miner Res.* 1999; 14(5):821–828. [PubMed: 10320531]
- McCloskey EV, Vasireddy S, Threlkeld J, et al. Vertebral fracture assessment (VFA) with a densitometer predicts future fractures in elderly women unselected for osteoporosis. *J Bone Miner Res.* 2008; 23(10):1561–1568. [PubMed: 18505372]
- Schousboe JT, Fink HA, Lui LY, Taylor BC, Ensrud KE. Association between prior non-spine non-hip fractures or prevalent radiographic vertebral deformities known to be at least 10 years old and incident hip fracture. *J Bone Miner Res.* 2006; 21(10):1557–1564. [PubMed: 16995810]
- van der Klift M, de Laet CE, McCloskey EV, et al. Risk factors for incident vertebral fractures in men and women: the Rotterdam Study. *J Bone Miner Res.* 2004; 19(7):1172–1180. [PubMed: 15177001]
- Ross PD, Davis JW, Epstein RS, Wasnich RD. Pre-existing fractures and bone mass predict vertebral fracture incidence in women. *Ann Intern Med.* 1991; 114(11):919–923. [PubMed: 2024857]
- Siris ES, Genant HK, Laster AJ, Chen P, Misurski DA, Krege JH. Enhanced prediction of fracture risk combining vertebral fracture status and BMD. *Osteoporos Int.* 2007; 18(6):761–770. [PubMed: 17245546]
- Fink HA, Milavetz DL, Palermo L, et al. What proportion of incident radiographic vertebral deformities is clinically diagnosed and vice versa? *J Bone Miner Res.* 2005; 20(7):1216–1222. [PubMed: 15940375]
- Schousboe JT, Ensrud KE, Nyman JA, Kane RL, Melton LJ 3rd. Cost-effectiveness of vertebral fracture assessment to detect prevalent vertebral deformity and select postmenopausal women with a femoral neck T-score > -2.5 for alendronate therapy: a modeling study. *J Clin Densitom.* 2006; 9(2):133–143. [PubMed: 16785071]
- Schousboe JT, Ensrud KE, Nyman JA, Kane RL, Melton LJ 3rd. Potential cost-effective use of spine radiographs to detect vertebral deformity and select osteopenic post-menopausal women for amino-bisphosphonate therapy. *Osteoporos Int.* 2005; 16(12):1883–1893. [PubMed: 16133649]
- Schousboe JT, Vokes T, Broy SB, et al. Vertebral Fracture Assessment: the 2007 ISCD Official Positions. *J Clin Densitom.* 2008; 11(1):92–108. [PubMed: 18442755]
- Cauley JA, Palermo L, Vogt M, et al. Prevalent vertebral fractures in black women and white women. *J Bone Miner Res.* 2008; 23(9):1458–1467. [PubMed: 18442309]
- El Maghraoui A, Rezqi A, Mounach A, Achemlal L, Bezza A, Ghozlani I. Systematic vertebral fracture assessment in asymptomatic postmenopausal women. *Bone.* 2013; 52(1):176–180. [PubMed: 23017663]
- Gallacher SJ, Gallagher AP, McQuillan C, Mitchell PJ, Dixon T. The prevalence of vertebral fracture amongst patients presenting with non-vertebral fractures. *Osteoporos Int.* 2007; 18(2):185–192. [PubMed: 17109062]
- Kaptoge S, Armbrecht G, Felsenberg D, et al. When should the doctor order a spine X-ray? Identifying vertebral fractures for osteoporosis care: results from the European Prospective Osteoporosis Study (EPOS). *J Bone Miner Res.* 2004; 19(12):1982–1993. [PubMed: 15537441]

16. Jacobs-Kosmin D, Sandorfi N, Murray H, Abruzzo JL. Vertebral deformities identified by vertebral fracture assessment: associations with clinical characteristics and bone mineral density. *J Clin Densitom.* 2005; 8(3):267–272. [PubMed: 16055955]
17. Vogt TM, Ross PD, Palermo L, et al. Vertebral fracture prevalence among women screened for the Fracture Intervention Trial and a simple clinical tool to screen for undiagnosed vertebral fractures. Fracture Intervention Trial Research Group. *Mayo Clin Proc.* 2000; 75(9):888–896. [PubMed: 10994823]
18. van der Jagt-Willems HC, van Hengel M, Vis M, et al. Why do geriatric outpatients have so many moderate and severe vertebral fractures? Exploring prevalence and risk factors. *Age Ageing.* 2012; 41(2):200–206. [PubMed: 22217460]
19. Vokes TJ, Gillen DL. Using clinical risk factors and bone mineral density to determine who among patients undergoing bone densitometry should have vertebral fracture assessment. *Osteoporos Int.* 2010; 21(12):2083–2091. [PubMed: 20306020]
20. Sanfelix-Gimeno G, Sanfelix-Genoves J, Hurtado I, Reig-Molla B, Peiro S. Vertebral fracture risk factors in postmenopausal women over 50 in Valencia, Spain. A population-based cross-sectional study. *Bone.* 2013; 52(1):393–399. [PubMed: 23103928]
21. Waterloo S, Nguyen T, Ahmed LA, et al. Important risk factors and attributable risk of vertebral fractures in the population-based Tromso study. *BMC Musculoskelet Disord.* 2012; 13:163. [PubMed: 22935050]
22. Kwok AW, Gong JS, Wang YX, et al. Prevalence and risk factors of radiographic vertebral fractures in elderly Chinese men and women: results of Mr. OS (Hong Kong) and Ms. OS (Hong Kong) studies. *Osteoporos Int.* 2013; 24(3):877–885. [PubMed: 22707064]
23. El Maghraoui A, Mounach A, Gassim S, Ghazi M. Vertebral fracture assessment in healthy men: prevalence and risk factors. *Bone.* 2008; 43(3):544–548. [PubMed: 18585994]
24. 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):157–72. discussion 207–12. [PubMed: 17569110]
25. Blank JB, Cawthon PM, Carrion-Petersen ML, et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials.* 2005; 26(5):557–568. [PubMed: 16085466]
26. Orwoll E, Blank JB, Barrett-Connor E, et al. 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]
27. Genant HK, Jergas M, Palermo L, et al. Comparison of semiquantitative visual and quantitative morphometric assessment of prevalent and incident vertebral fractures in osteoporosis The Study of Osteoporotic Fractures Research Group. *J Bone Miner Res.* 1996; 11(7):984–996. [PubMed: 8797120]
28. Genant HK, Wu CY, van Kuijk C, Nevitt MC. Vertebral fracture assessment using a semiquantitative technique. *J Bone Miner Res.* 1993; 8(9):1137–1148. [PubMed: 8237484]
29. Beard CM, Melton LJ 3rd, Cedel SL, Richelson LS, Riggs BL. Ascertainment of risk factors for osteoporosis: comparison of interview data with medical record review. *J Bone Miner Res.* 1990; 5(7):691–699. [PubMed: 2396496]
30. Nakamura K, Inoue M, Kaneko Y, Tsugane S. Positive predictive values for self-reported fractures in an adult Japanese population. *Environ Health Prev Med.* 2011; 16(2):129–132. [PubMed: 21432228]
31. Siminoski K, Warshawski RS, Jen H, Lee K. The accuracy of historical height loss for the detection of vertebral fractures in postmenopausal women. *Osteoporos Int.* 2006; 17(2):290–296. [PubMed: 16143833]
32. Roux C, Priol G, Fechtenbaum J, Cortet B, Liu-Leage S, Audran M. A clinical tool to determine the necessity of spine radiography in postmenopausal women with osteoporosis presenting with back pain. *Ann Rheum Dis.* 2007; 66(1):81–85. [PubMed: 16793842]
33. Silverman SL, Minshall ME, Shen W, Harper KD, Xie S. The relationship of health-related quality of life to prevalent and incident vertebral fractures in postmenopausal women with osteoporosis: results from the Multiple Outcomes of Raloxifene Evaluation Study. *Arthritis Rheum.* 2001; 44(11):2611–2619. [PubMed: 11710717]

34. Silverman SL, Piziak VK, Chen P, Misurski DA, Wagman RB. Relationship of health related quality of life to prevalent and new or worsening back pain in postmenopausal women with osteoporosis. *Journal of Rheumatology*. 2005; 32(12):2405–2409. [PubMed: 16331772]
35. Doo TH, Shin DA, Kim HI, et al. Clinical Relevance of Pain Patterns in Osteoporotic Vertebral Compression Fractures. *Journal of Korean Medical Science*. 2008; 23(6):1005–1010. [PubMed: 19119444]
36. Cawthon PM, Ewing SK, McCulloch CE, et al. Loss of hip BMD in older men: the osteoporotic fractures in men (MrOS) study. *J Bone Miner Res*. 2009; 24(10):1728–1735. [PubMed: 19419308]
37. Harkonen R, Harju R, Alaranta H. Accuracy of the Jamar dynamometer. *J Hand Ther*. 1993; 6(4): 259–262. [PubMed: 8124439]
38. Pregibon D. Goodness of link tests for generalized linear models. *Applied Statistics*. 1980; 29:15–24.
39. Steyerberg, EW. *Clinical Prediction Models: a Practical Approach to Development, Validation, and Updating*. In: M, G.; K, K.; A, T.; J, S.; W, W., editors. *Statistics for Biology and Health*. Springer: LLC; 2009.
40. Fuerst T, Wu C, Genant HK, et al. Evaluation of vertebral fracture assessment by dual X-ray absorptiometry in a multicenter setting. *Osteoporos Int*. 2009; 20(7):1199–1205. [PubMed: 19083074]
41. Binkley N, Krueger D, Gangnon R, Genant HK, Drezner MK. Lateral vertebral assessment: a valuable technique to detect clinically significant vertebral fractures. *Osteoporos Int*. 2005; 16(12): 1513–1518. [PubMed: 15834512]
42. Vokes TJ, Dixon LB, Favus MJ. Clinical utility of dual-energy vertebral assessment (DVA). *Osteoporos Int*. 2003; 14(11):871–878. [PubMed: 13680099]

Table 1

Baseline Characteristics of Participants without a self-reported history of prior vertebral fracture and with femoral neck T-score -1.0^*

Parameter	Radiographic PVFx Absent (n = 2,870)	SQ1 Only Fracture Present (n=148)	SQ2 or SQ3 Fracture Present (n=238)	P-Value
Age, years (SD)	74.3 (6.0)	74.2 (6.1)	75.6 (6.5)	0.006**
Femoral Neck BMD, gm/cm ² (SD)	0.702 (0.066)	0.693 (0.070)	0.665 (0.080)	<0.001**
Height Loss, cm (SD)	3.8 (3.0)	4.0 (2.9)	5.2 (3.2)	<0.001**
Grip Streng, kg (SD)	37.8 (7.9)	37.1 (8.3)	35.8 (8.1)	<0.001**
Back Pain: None	32.5%	32.5%	28.0%	0.021 [^]
Yes, No Limitation	44.4%	41.2%	40.4%	
Yes, Minor Limitation	14.9%	15.0%	18.5%	
Yes, Major Limitation	8.2%	11.2%	13.0%	
Percent with Prior Non-Spine Fracture Since Age 50	18.9%	21.2%	39.5%	<0.001 [^]
BMI, kg/m ² (SD)	26.4 (3.5)	27.1 (4.0)	26.7 (3.8)	0.048*
Percent Current Smokers	3.3%	3.1%	4.3%	0.23 [^]
Percent Currently Using Oral Glucocorticoids ^{^^}	2.8%	3.0%	6.0%	0.014 [^]

* Prior self-reported spine fracture variable is missing on 1,404; these individuals are included in these analyses

** one-way analysis of variance

[^] chi-square statistic

^{^^} data available on only 2,686

Table 2

Comparison of Nested Models Predicting Moderate to Severe Prevalent Radiographic Vertebral Fractures (SQ Grade 2 or 3)

Parameter	Odds Ratio (95% C.I.)			
	Model 1	Model 2	Model 3	Model 4
Age (per 5 year increase)	1.12 (1.00 – 1.25)	1.04 (0.93 – 1.17)	1.03 (0.91 – 1.17)	1.00 (0.85 – 1.14)
Femoral Neck BMD (per SD increase)	0.43 (0.34 – 0.54)	0.46 (0.37 – 0.58)	0.44 (0.36 – 0.58)	0.47 (0.36 – 0.62)
Height Loss (per SD increase)		1.43 (1.26 – 1.62)	1.38 (1.21 – 1.58)	1.33 (1.15 – 1.54)
Non-Spine Fx Hx Since Age 50			1.51 (1.11 – 2.07)	1.46 (1.02 – 2.07)
BMI (per SD increase)			1.19 (1.03 – 1.87)	1.20 (1.02 – 1.41)
Grip Strength, kg (SD)			0.91 (0.78 – 1.07)	0.78 (0.66 – 0.94)
Level of Back Pain				
None			Reference	Reference
Yes, No limitation			0.95 (0.69 – 1.30)	1.08 (0.74–1.57)
Yes, Minor limitation			1.17 (0.78 – 1.75)	1.33 (0.84–2.10)
Yes, Major limitation			0.82 (0.47 – 1.42)	0.87 (0.48–1.60)
Glucocorticoid Use (yes vs. no)				1.02 (0.42 – 2.52)
Current Smoking				
Never				Reference
Past				1.07 (0.77–1.50)
Current				1.39 (0.60–3.23)
C-statistic: development	0.642 (0.605 – 0.679)	0.682 (0.646–0.717)	0.690 (0.654 – 0.726)	0.692 (0.652 – 0.732)
C-statistic: Validation datasets*	0.638 – 0.642	0.676–0.681	0.668 –0.687	0.668–0.689

* Range of c-statistics in 5 separate bootstrapped datasets

Table 3

Comparisons of Model Discrimination of Those With from Those Without Prevalent Radiographic Vertebral Fracture

Comparison Measure	Model 2 vs. 1	Model 3 vs. 2	Model 4 vs. 2
^{&} Range of C-stat χ^2 (Range of p-values)	8.55 – 16.5 (<0.001 to 0.003)	0.00 – 1.02 (0.28 to 0.96)	0.01 – 2.13 (0.14 to 0.92)
NRI – 5% [^] (p-value)	0.059 (0.01)	0.026 (0.22)	0.038 (0.13)
NRI – 10% ^{**} (p-value)	0.049 (0.09)	0.003 (0.90)	0.037 (0.26)
NRI – 15% ^{^^} (p-value)	0.024 (0.22)	0.006 (0.78)	0.012 (0.64)

[&]Comparisons across five pairs of bootstrapped models

[^]Net Reclassification Index Score (Pepe Method), with a pre-test probability cutpoint of 5%

^{**}Net Reclassification Index Score (Pepe Method), with a pre-test probability cut point of 10%

^{^^}Net Reclassification Index Score (Pepe Method), with a pre-test probability cut point of 15%

Table 4

Proportion of Men Who Would Receive Lateral Spine Imaging & Proportion With Prevalent Radiographic Vertebral Fracture (PVFx) Detected[%]

	Screening Pre-Test Probability Cutpoint ^{**}	Percent Screened [^]	Percent of Men with PVFx Detected	Prevalence PVFx Among Those Screened
Model 1 [#]	5%	68.1%	81.1%	8.7%
	10%	18.6%	34.5%	13.6%
	15%	4.9%	11.3%	16.9%
Model 2 ^{&}	5%	63.5%	81.9%	9.4%
	10%	18.2%	38.7%	15.5%
	15%	6.4%	15.1%	17.3%
Model 3 ^{##}	5%	60.5%	81.4%	9.7%
	10%	18.6%	39.8%	15.5%
	15%	6.6%	16.0%	17.5%
Model 4 ^{&&}	5%	60.5%	81.2%	9.7%
	10%	19.1%	42.0%	15.9%
	15%	7.1%	18.8%	19.1%

* Prediction model used to determine who has lateral spine imaging to look for prevalent vertebral fracture

** Pre-test probability of PVFx cutpoint at and above which lateral spine imaging would be done

[^] Proportion above the cutpoint according to the prediction model who would have spine imaging

[#] Model 1: Age and Femoral Neck BMD

[&] Model 2: Age, Femoral Neck BMD, and historical height loss

^{##} Model 3: Age, Femoral Neck BMD, historical height loss, prior non-vertebral fracture, body mass index, presence of and limitations from back pain, and grip strength

^{&&} Model 4: Age, Femoral Neck BMD, historical height loss, prior non-vertebral fracture, body mass index, presence of and limitations from back pain, grip strength, smoking, and glucocorticoid use

[%] Analyses limited to men who did not self-report prior spine fracture at the baseline visit and with femoral neck BMD T-score ≥ -1.0

Table 5
 Net Correct Reclassification of Those with SQ2 or SQ3 Fractures with Prediction Models Compared to Simple Indication Lists

	Model					
	Regression Model 2 [*]	Regression Model 4 [#]	ISCD 2007 ^{**} Criteria	Simple Criteria 1 [†]	Simple Criteria 2 [^]	Simple Criteria 3 ^{^^}
% Vert Fx Captured	81.9%	81.4%	76.9%	51.3%	65.5%	73.1%
% Screened	63.5%	60.5%	63.3%	35.3%	55.2%	54.8%
Yield ^{&&}	9.4%	9.7%	8.9%	10.6%	8.7%	9.8%
NRI Model 2 Criteria List			0.052 (p=0.10)	0.026 (p=0.53)	0.087 (p=0.02)	0.001 (p=0.97)
NRI Model 4 Criteria List			0.105 (p=0.003)	0.045 (p=0.32)	0.118 (p=0.004)	0.057 (p=0.13)
NRI vs. ISCD complex				0.026 (p=0.45)	-0.035 (p=0.21)	0.051 (p=0.06)

[&] All analyses done in subset of men with femoral neck BMD T-score ≥ -1.0

^{&&} Proportion of those screened who have a positive spine image for vertebral fracture

^{*} Regression model 2, with a 5% predicted probability cutpoint

[#] Regression model 4, with a 5% predicted probability cutpoint

^{**} ISCD 2007 Position Statement Criteria: age ≥ 80 years or height loss ≥ 6 cm or glucocorticoid use; OR two of following; age 70–79 years plus, height loss 3 to 6 cm, prior non-vertebral fracture, androgen deprivation

[†] Simple: age ≥ 80 years or height loss ≥ 6 cm or glucocorticoid use

[^] Simple2: age ≥ 75 years or height loss ≥ 6 cm or glucocorticoid use

^{##} Simple3: age ≥ 80 years or height loss ≥ 4 cm or glucocorticoid use