Title
Prediction models of prevalent radiographic vertebral fractures among older women.

Journal
Journal of clinical densitometry : the official journal of the International Society for Clinical Densitometry, 17(3)

ISSN
1094-6950

Authors
Schousboe, John T
Rosen, Harold R
Vokes, Tamara J
et al.

Publication Date
2014-07-01

DOI
10.1016/j.jocd.2013.09.021

Peer reviewed

Published in final edited form as:
J Clin Densitom. 2014 July ; 17(3): 378–385. doi:10.1016/j.jocd.2013.09.021.

Prediction Models of Prevalent Radiographic Vertebral Fractures Among Older Women

John T. Schousboe, MD, PhD[1,2], **Harold R. Rosen, MD**[3], **Tamara J. Vokes, MD**[4], **Jane A. Cauley, DrPH**[5], **Steven R. Cummings, MD**[6], **Michael Nevitt, PhD**[7], **Dennis M. Black, PhD**[7], **Eric S. Orwoll, MD**[8], **Deborah M. Kado, MD, MS**[9], **Kristine E. Ensrud, MD, MPH**[10,11], and **For the Study of Osteoporotic Fractures (SOF) Research Group**

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

[2]Division of Health Policy and Management, University of Minnesota

[3]Beth Israel Deaconess Medical Center, Boston, MA

[4]University of Chicago, Chicago, IL

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

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

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

[8]Oregon Health and Science University, Portland, OR

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

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

[11]Minneapolis Veterans Administration Medical Center, Minneapolis, MN

Abstract

It is unknown how well prediction models incorporating multiple risk factors identify women with radiographic prevalent vertebral fracture (PVFx) compared to simpler models, and what their value might be in clinical practice to select older women for lateral spine imaging. We compared four regression models for predicting PVFx in women age 68 and older enrolled in the Study of Osteoporotic Fractures with a femoral neck T-score of ≤ −1.0, using area under receiving operator characteristics curves (AUROC) and a net reclassification index. The AUROC for a model with age, femoral neck bone mineral density (BMD), historical height loss (HHL), prior non-spine fracture, body mass index, back pain, and grip strength was only minimally better than that of a

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.

more parsimonious model with age, femoral neck BMD, and HHL (AUROC 0.689 vs. 0.679, p-values for difference in five bootstrapped samples <0.001 to 0.35). The prevalence of PVFx among this older population of Caucasian women remained over 20% even when women with low probability of PVFx, as estimated by the prediction models, were included in the screened population. These results suggest that lateral spine imaging is appropriate to consider for all Caucasian women age 70 and older with low bone mass to identify those with PVFx.

Keywords

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

Introduction

Prevalent vertebral fractures (PVFx) are common among older persons,(1, 2) are a marker of bone fragility and high fracture risk,(3–5), but frequently remain unrecognized in clinical practice.(6) Vertebral fracture assessment either with densitometric lateral spine images(7) or standard spine radiographs(8) is a cost-effective method to identify those with PVFx. A barrier to identifying identify those with PVFx may be the complexity of existing guidelines.(9) While many risk factors have been shown to be independently associated with PVFx in multivariable-adjusted regression models,(1, 10–19) it is unknown if prediction models incorporating these additional risk factors identify those with radiographic PVFx better than more parsimonious models. We had two objectives; a) to examine how well simple regression-based versus complex models discriminate those who have radiographic PVFx from those who do not in women age ≥ 68 years enrolled in the Study of Osteoporotic Fractures (SOF), using area under receiving operator characteristics (AUROC) curves and the net reclassification improvement (NRI) method of Pencina;(20) and b) to establish the simplest, most parsimonious model that might be used in clinical practice to detect previously undiagnosed PVFx in older women.

Materials and Methods

The Study of Osteoporotic Fractures (SOF) enrolled 9,704 Caucasian women in 1986 to 1988 in four metropolitan areas of the United States (Baltimore MD, Minneapolis, MN, Pittsburgh PA, and Portland OR). Methods of study recruitment have been described previously.(21) Lateral lumbar and thoracic spine radiographs were obtained at the first (1986 to 1988) and third (1990 to 1991) SOF study visits. Bone mineral density (BMD) was measured at the hip at the second and subsequent SOF visits, but only calcaneal BMD was measured at the first study visit. Since BMD is more often measured at the hip in clinical practice, we used data from the second and third SOF visits for our analyses.

Identification of Prevalent Radiographic Vertebral Fractures

The parent study population consisted of 7,233 women who attended the third SOF visit and had technically adequate lateral lumbar and thoracic spine x-rays. As previously described, six point digitations of each vertebra from T4 through L4 inclusive were done, so anterior

(H_a), middle (H_m), and posterior (H_p) vertebral heights could be accurately measured for quantitative morphometry.(3) A vertebral body was considered deformed if either of two height ratios within the vertebra (Ha/Hp, Hm/Hp) was >3 SD below the mean for that vertebral level or if both the anterior and posterior heights relative to the vertebra immediately inferior (Ha/Ha+1 and Hp/Hp+1) or superior (Ha/Ha−1 and Hp/Hp−1) were >3 SD below the mean for that level. Mild vertebral deformities and moderate to severe vertebral deformities, respectively, were defined as those with height ratios > 3 SD but ≤ 4 SD, and > 4 SD below the expected value for that vertebral level based on normative SOF data.(22)

Detection of a previously undiagnosed vertebral fracture is based upon the supposition that such identification would alter therapy. Therefore, we restricted our analyses to the population with a femoral neck T-score ≤ −1.0 (n = 5,560) because there is no published evidence regarding the efficacy of currently available fracture prevention therapies in those with normal BMD.

Measurement of Bone Mineral Density

BMD was measured at the femoral neck and total hip with QDR-1000 scanners (Hologic, Bedford, MA, USA), at each study site for every fifth person (a total of 1,506) who attended the third SOF study visit, whereas 94% of the 7223 third visit attendees had femoral neck and total hip BMD measured at the second SOF visit. *In vivo* coefficient of variation was 1.2% at the femoral neck. Further details of densitometry quality control methods in SOF have been published previously.(23)

One thousand two hundred and sixty four women (1,264) had hip BMD measured at both visits. We imputed missing femoral neck and total hip bone density values among the 5,531 women with hip BMD only measured at visit 2 in two steps using a validated statistical method,(24, 25) as detailed in the appendix.

Measurement of other covariates

At the baseline visit, all SOF participants were asked their height at age 25 and if they had had any fractures since age 50. Participants were subsequently mailed postcards every 4 months and asked if they had had any fractures and their skeletal locations. They were asked whether or not they were currently smoking cigarettes, taking estrogen replacement therapy, and/or systemic glucocorticoid therapy at the baseline and all subsequent visits. Current height and weight were measured at each study visit, respectively, using a Harpenden stadiometer and a balance beam scale. Historical height loss (HHL) was defined as the difference between recalled height at age 25 minus measured height at the third SOF visit. Body mass index (BMI) was defined as weight (kg) divided by height (meters) squared.

Selection of Covariate Predictors

The positive predictive value of a positive self-report of vertebral fracture has been reported to be as high as 85%.(26) If our analyses confirmed this estimate, we planned to develop models in the subset of the SOF population who had neither a self-reported prior vertebral

fracture at the baseline visit nor an incident clinical vertebral fracture between the first and third visits.

We chose age and femoral neck BMD as our simplest model. HHL is an independent risk factor for PVFx (12, 14, 16) and a stand-alone indication for vertebral fracture assessment in the 2007 ISCD Position Statement for VFA indications.(9) Hence, our second model for comparison included age, femoral neck BMD and HHL as predictors. Prior non-vertebral fracture, BMI, grip strength, and self-reported back pain were included in a third, more complex model. Prior fracture is a secondary indication (when combined with age) in the 2007 ISCD indications for VFA,(9) and BMI has been identified in some studies(10, 12, 14, 17), but not others(15, 18, 19) as a risk factor for vertebral fracture. Other studies have identified back pain to be associated with prevalent vertebral fractures in women,(19, 27, 28) and two have identified grip strength as to be associated with PVFx.(1, 19) The fourth, most complex model included the covariates of the third model, glucocorticoid use, estrogen replacement therapy, and current smoking.

Statistical Analyses

The primary analyses used logistic regression models with all prevalent vertebral fractures (height ratio > 3SD below mean) as the dependent variable in women with a third visit femoral neck T-score of ≤ −1.0. Four sets of secondary analyses were done; one with only moderate to severe fractures (vertebral height ratio > 4 SD below mean) as the dependent variable, restricting the analysis to those with osteopenia (femoral neck T-score between −1.0 and −2.4), 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 to test whether or not including of non-linear predictors might improve model discrimination, and included adding age-squared and interaction terms between age and femoral neck BMD, age and HHL, HHL and BMD, and HHL and prior non-spine fracture. Finally, we tested whether or not modeling age, femoral neck BMD, HHL, BMI, and grip strength as four level categorical rather than continuous variables improved model discrimination. For all regression models, model fit and calibration was tested with the Hosmer-Lemeshow test, and model specification with Pregibon's linktest.(29)

Because AUROC statistics derived in the same samples in which they were produced can be overinflated, we produced five bootstrapped models for each of the four parent models, and compared the areas under the curve (AUROC) statistic (C-statistic) between the nested models for each of five pairs of bootstrapped samples. Models with an AUC of 0.5 have no value predicting the outcome, whereas models with an AUC of 1.0 are able to perfectly discriminate who have from those who do not have the outcome.

While AUROC statistics assess 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 prevalence (e.g. 10%) of vertebral fracture. (7, 8) Net reclassification indices are a method of testing how well two prediction rules correctly classify individuals who have an outcome and those who do not have that outcome, using a specific prevalence cutpoint of that outcome. For example, if we chose a prevalence cutpoint for PVFx of 10%, a good prediction model would be one where the far majority of

those with a model predicted probability of ≥ 10% for PVFx being present truly have one, and the far majority of those with a model predicted probability of < 10% for PVFx being present do not have one. By the Pencina method,(20) the *net reclassification improvement* (NRI) statistic using a second model instead of a first model is calculated as the proportion of individuals that are shifted from being *incorrectly classified to correctly classified* using Model 2 instead of Model 1, minus the proportion of those shifted from being *correctly classified to incorrectly classified* using model 2 instead of model 1. The statistical significance of the NRI for each model comparison was calculated by the method of Pencina.(20) We compared nested models with net reclassification indices at a prevalence cutpoints of 5%, 10%, and 15%.

Finally, 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, using pre-test probability cutpoints of 5%, 10% or 15%, we calculated for each model; a) the proportion who would be chosen to have lateral spine imaging; b) the proportion of women with 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.

Results

Among all 7,233 women who attended visit 3 and had lateral spine films, 1,721 (24%) had one or more PVFx; 1,162 (16%) had a moderate or severe PVFx. Four hundred seventy one (471, or 6.5%) had self-reported a vertebral fracture as of SOF visit 3, and 70.4% of them had one or more PVFx identified on the visit 3 film; 59% had a moderate or severe PVFx. Hence, we reasoned that a self-reported (but undocumented) prior vertebral fracture would be a reasonable stand-alone indication for lateral spine imaging, and did not include these women in subsequent analyses. We further restricted our analyses to those with a femoral neck T-score ≤ −1.0.

The characteristics of the remaining 5,166 women are shown in table 1; those who had a prevalent vertebral fracture were older, had lower BMD, had more HHL, lower grip strength, were more likely to have had a self-reported prior non-spine fracture since age 50, and to have had back pain over the prior two years. Importantly, the youngest woman in our sample was 68 years of age.

The associations of potential predictors included in the 4 nested models with radiographic PVFx (defined as a height ratio > 3 SD below expected) is shown in table 2. Older age, lower BMD, greater HHL, prior non-vertebral fracture, higher BMI, and current smoking were all independently associated with prevalent vertebral fracture. In all bootstrapped model comparisons, the AUROC of model 2 (AUC range 0.678 – 0.679) was superior to that of model 1 (AUC range 0.643 to 0.644, chi^2 range 35.2 to 42.2, p-values all <0.001). Moreover, 5.0% and 5.3% of women had a net correct reclassification of PVFx status using Model 2 instead of Model 1 at pre-test probability cutpoints of 10% and 15%, respectively. Because virtually all participants in this study were women age 68 years and older, nearly all had a pre-test probability of 5% or more of having a PVFx, and the models did not differ in

their ability to discriminate those with from those without PVFx using a pre-test probability cutpoint of 5%.

The difference in the AUROC for model 3 (AUC range 0.684 to 0.689) compared to model 2 (AUC range 0.678 to 0.679) across bootstrapped samples was of marginal significance (chi^2 range 0.89 to 11.5, p-value range <0.001 to 0.35, table 3). Using prevalence cutpoints of 10% and 15% respectively, only a net 2.6% and 2.5% of women were correctly re-classified with respect to their PVFx status. No meaningful differences in discrimination of those with from those without PVFx were when model 4 was compared to model 3 (table 3).

In secondary analyses with prevalent vertebral fracture defined as > 4SD below expected, findings were similar to those of the primary analyses (data not shown). Similarly, results were unchanged when analyses were restricted to the 3,644 women with a femoral neck T-score <−1.0 and > −2.5, or when spine BMD was substituted for femoral neck BMD as a predictor. Moreover, adding age squared, interaction terms between age and BMD, age and height loss, height loss and BMD, and prior non-spine fracture and height loss did not improve discrimination of any of the four models (data not shown). Finally, modeling age, femoral neck BMD, height loss, body mass index, and/or grip strength as 4-level categorical rather than as continuous variables did not improve model performance.

There was a wide range of the calculated proportions of women who would be screened, of women with PVFx who would be identified, and of the prevalence of PVFx among women who are screened, driven primarily by the pre-test probability of a PVFx chosen to decide whether or not lateral spine imaging should be performed (table 4). As the pre-test probability cutpoint was raised, a lower proportion of women would receive lateral spine imaging, but a rising proportion of women with prevalent vertebral fracture would not be detected. At any given pre-test probability cutpoint, the proportions who would be screened and the proportion of women with a PVFx detected were similar regardless of which model was chosen. Among all of these scenarios, the lowest prevalence of radiographic PVFx among those who would have spine imaging was 20.3% (table 4). Using prevalence cutpoints of 10% and 15%, respectively, the prevalence of PVFx in women with low bone mass who would *not* be selected for screening using either model 2 or model 3 were 8% and 11%.

Discussion

In this cohort of women enrolled in the Study of Osteoporotic Fractures (mean age 75 years, minimum age 68 years), radiographic vertebral fractures were quite common, being present in 20.4% among those with a femoral neck T-score of −1.0 or lower, after excluding those with prior self-reported vertebral fractures. Although we confirmed findings of multiple other studies that numerous risk factors, including age, BMD, HHL, self-reported prior non-vertebral fractures, body mass index, and smoking were associated with prevalent vertebral fractures after multivariable adjustment, we could not confirm that more complex models incorporating all of these risk factors could discriminate those with from those without prevalent PVFx substantially better than more parsimonious models. Based on the *statistical* significance of the model comparisons using AUROC analyses and a net reclassification

index, 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 nearly as well as more complex models.

However, the discrimination of all of these models of older women with PVFx from those without PVFx was at best modest, and the *clinical relevance* of the differences of model performance was minimal. Using the most liberal model predicted probability of a PVFx being present to select women for lateral spine imaging, virtually all women with the characteristics of this subset of the SOF population would be screened, and the prevalence of PVFx in the screened population would still be higher than 20%. In essence, these prediction models could not isolate a subset of older women with low bone mass with a sufficiently low enough probability of PVFx being present to warrant not screening them, regardless of how many predictor covariates were included. These data lend strong support to the ISCD Position Statement that lateral spine imaging is appropriate in all Caucasian women with low bone mass age 70 and older. However, this position is predicated on; a) that the results of the lateral imaging study might alter management of the patient; and b) that the reader of the lateral spine image has the requisite training to discern fractured vertebrae from normal vertebrae and non-fracture deformities.

Our findings leave open the possibility that prediction models might have value selecting individuals for lateral spine imaging to detect PVFx in populations where the prevalence of PVFx is lower, such as in younger age women. Our results do suggest that simpler models with a few major risk factors may perform as well as complex models with multiple risk factors in discriminating individuals with from those without PVFx, but additional studies of the performance of such prediction models for PVFx in younger women are needed to test this hypothesis. Based on current knowledge, the clinical utility of *any* prediction model with low to modest discrimination of PVFx in any population may be limited, and further research to identify additional risk factors that are more strongly associated with PVFx is needed.

A major limitation of our study is that women younger than age 68 and non-Caucasian women were not included. Second, radiographic vertebral fractures in this study were adjudicated with full quantitative morphometric criteria that are not practical for use in clinical practice, and that do not differentiate vertebral fractures from non-fracture deformities.(9, 30, 31) The semi-quantitative (SQ) method of Genant, used very commonly in clinical practice, has been compared to the quantitative morphometric method specifically used in SOF in a subset of 503 SOF participants. The prevalences of vertebral fractures, when defined as one or more vertebrae with an SQ grade of ≥ 1, an SQ grade of ≥ 2, or a height ratio > 3 SD below expected were, respectively, 33.4%, 13.9%, and 19.9%.(32) This suggests that the criterion of one or more height ratios > 3 SD below the mean for that vertebral level is capturing vertebral fractures of a severity somewhere in between SQ grade 1 and 2. The validity of our results is also supported in that our results were similar when we used the height ratio >4SD below the mean to define vertebral fractures. A third limitation is that these models have not been externally validated in other studies and populations. Finally, estimates of femoral neck BMD at the third visit were imputed from values obtained at the second visit for 80% of our study sample, but this is mitigated by very high correlation

coefficient (0.93) between femoral neck BMD measurements obtained at the second and third visits, and that the proportion of variance of visit 3 femoral neck BMD explained by the imputation model (R^2 statistic) was 0.93.

Study strengths are that SOF is the largest cohort study of postmenopausal women that includes comprehensive assessment of PVFx with lateral spine radiographs, and that SOF study participants were recruited from large groups and registries closely representative of the Caucasian female population of the United States.

In conclusion, a simple prediction model based on age, femoral neck BMD, and historical height loss discriminates older Caucasian women who have from those who do not have prevalent radiographic vertebral fracture as well as more complex models. However, the prevalence of radiographic vertebral fracture is sufficiently high in this population such that lateral spine imaging to detect PVFx for all women age 70 years or older with low bone mass in whom such detection would alter clinical management.

Acknowledgments

The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576

Appendix

We first regressed visit 3 femoral neck BMD on visit 2 femoral neck BMD, visit 2 spine BMD, current smoking at visit 2, estrogen use, glucocorticoid use, and the time between the dates of visit 2 and visit 3 femoral neck BMD measurements for 1, 212 individuals who had femoral neck BMD at measured at both visits. The model was flawed due to heteroscedasticity. This appeared to be due to 55 observations with undue influence (Cook's distance > 4/1211), and the regression was repeated with these 55 individuals excluded. The associations of predictors with visit 3 femoral neck BMD (appendix table 5) showed that visit 3 femoral neck BMD was by and large determined by visit 2 femoral neck BMD; a regression with femoral neck BMD by itself had an R^2 statistic of 0.90. The addition of the other covariates listed above raised the R^2 statistic further to 0.93. This *ice* command in Stata was used to impute missing values for the remaining 5,531 SOF participants with known visit 2 femoral neck BMD.

Appendix Table 5: Predictors of Femoral Neck BMD at SOF Visit 3

Predictor	Coefficient (95% C.I)	t-statistic
Visit 2 Femoral Neck BMD	0.951 (0.927 to 0.974)	79.06
Visit 2 Spine BMD	0.037 (0.022 to 0.053)	4.76
Weight gain between visits 2 and 3 (per 5 kg increase)	0.0065 (0.0035 to 0.0095)	4.03
Current smoking	−0.0072 (−0.1517 to 0.0007)	−1.78
Number of years between visits 2 and 3*	0.0046 (−0.0023 to 0.0114)	1.30

Predictor	Coefficient (95% C.I)	t-statistic
Glucocorticoid use	−0.0034 (−0.0147 to 0.0078)	−0.59
Estrogen Use	0.0034 (−0.0024 to 0.0092)	1.14

* Mean time between visits 2 and 3: 1.48 years (sd 0.33)

References

1. 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]
2. 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]
3. 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]
4. 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]
5. 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]
6. 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]
7. 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]
8. 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]
9. 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]
10. 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]
11. Gallacher SJ, Gallagher AP, McQuillian 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]
12. 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]
13. 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]
14. 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]

15. 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]
16. 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]
17. 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]
18. 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]
19. 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]
20. 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. discussion 207–12. [PubMed: 17569110]
21. Cummings SR, Black DM, Nevitt MC, et al. Appendicular bone density and age predict hip fracture in women. The Study of Osteoporotic Fractures Research Group. Jama. 1990; 263(5): 665–668. [PubMed: 2404146]
22. Black DM, Cummings SR, Stone K, Hudes E, Palermo L, Steiger P. A new approach to defining normal vertebral dimensions. J Bone Miner Res. 1991; 6(8):883–892. [PubMed: 1785377]
23. Steiger P, Cummings SR, Black DM, Spencer NE, Genant HK. Age-related decrements in bone mineral density in women over 65. J Bone Miner Res. 1992; 7(6):625–632. [PubMed: 1414480]
24. Royston P. Multiple Imputation of Missing Values: Update of ICE. The Stata Journal. 2005; 5(4): 527–536.
25. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009; 338:b2393. [PubMed: 19564179]
26. 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]
27. 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]
28. 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]
29. Pregibon D. Goodness of link tests for generalized linear models. Applied Statistics. 1980; 29:15–24.
30. Ferrar L, Jiang G, Adams J, Eastell R. Identification of vertebral fractures: an update. Osteoporos Int. 2005; 16(7):717–728. [PubMed: 15868071]
31. 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]
32. 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]

Table 1

Characteristics of SOF Population attending Visit 3, EXCLUDING those with self-reported prior spine fracture at baseline visit or incident vertebral fracture between visits 1 and 3 and/or femoral neck T-score> −1.0

Parameter	Prev Vert Fx Absent (n = 4,039)	SD3 Only Fracture Present (n=412)	SD4 Fracture Present (n=715)	P-Value
Age, years (SD)	74.6 (4.8)	75.3 (5.1)	76.8 (5.7)	<0.001*
Femoral Neck BMD, gm/cm^2 (SD)	0.616 (0.072)	0.595 (0.073)	0.574 (0.078)	<0.001*
Height Loss, cm (SD)	3.5 (2.7)	4.5 (2.8)	5.4 (3.4)	<0.001*
Percent with Non-Spine Fx Since Age 50	41.3%	49.9%	57.0%	<0.001^
BMI, kg/m^2 (SD)	25.8 (4.4)	26.0 (4.7)	25.7 (4.2)	0.38*
Grip Strength (kg)	19.1 (4.3)	18.7 (4.1)	17.9 (4.6)	<0.001*
Percent with Back Pain Over Prior 2 Years	60.8%	64.6%	67.2%	0.003^
Percent Current Smokers at Visit 2:	7.5%	9.3%	8.6%	0.33^
Estrogen Use				
Never	57.0%	62.9%	60.0%	0.14^
Past	29.4%	25.0%	26.8%	
Current	13.6%	12.1%	13.1%	
Gluco Use				
Neither V2 nor V3	94.2%	95.1%	92.9%	0.03^
Either V2 or V3	3.9%	2.4%	3.5%	
Both V2 and V3	1.9%	2.4%	3.5%	

*one-way analysis of variance

^chi-square statistic

Table 2

Comparison of Nested Models' (from Table 2) Discrimination of Prevalent Radiographic Vertebral Fractures (Height Ratio > 3 SD below expected)

Parameter	Model 1	Model 2	Model 3	Model 4
Age (per 5 year increase)	1.26 (1.18 – 1.34)	1.11 (1.04 – 1.19)	1.13 (1.04 – 1.21)	1.14 (1.06 – 1.23)
Femoral Neck BMD (per SD increase)	0.67 (0.62 – 0.71)	0.70 (0.65 – 0.75)	0.68 (0.63 – 0.73)	0.68 (0.62 – 0.73)
Height Loss (per SD increase)		1.53 (1.42 – 1.65)	1.47 (1.36 – 1.59)	1.47 (1.36 – 1.59)
Non-Spine Fx Hx Since Age 50			1.34 (1.15 – 1.55)	1.37 (1.18 – 1.58)
BMI (per SD increase)			1.14 (1.06 – 1.23)	1.17 (1.08 – 1.26)
Grip Strength (per SD increase)			0.99 (0.91 – 01.06)	0.99 (0.92 – 1.07)
Back Pain over past 2 years (Yes vs. No)			1.05 (0.90 – 1.23)	1.10 (0.95 – 1.50)
Estrogen Use				
Never				Reference
Past				0.98 (0.83–1.16)
Current				1.19 (0.95–1.50)
Gluco Use*				
Neither V2 nor V3				Reference
Either V2 or V3				0.77 (0.52–1.16)
Both V2 and V3				1.30 (0.84–2.02)
Current Smoking (yes / no)				1.35 (1.04 – 1.75)
C-statistic: development	0.644 (0.625–0.662)	0.679 (0.661–0.697)	0.689 (0.671–0.707)	0.690 (0.672–0.708)
C-statistic: Validation datasets**	0.643–0.644	0.678 – 0.679	0.684 – 0.689	0.686 – 0.689

*Current glucocorticoid use at SOF visit 2 (V2) or visit 3 (V3)

**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. 3
&Range of C-stat chi^2 (range of p-values)	35.2 – 42.2 (all <0.001)	0.89 – 11.5 (<0.001 – 0.35)	0.02 – 2.5 (0.02 – 0.90)
NRI – 5% ^ (p-value)	0.000 (0.16)	0.002 (0.07)	0.003 (0.003)
NRI – 10%** (p-value)	0.050 (<0.001)	0.026 (<0.001)	0.008 (0.10)
NRI – 15%^^ (p-value)	0.053 (<0.001)	0.025 (0.01)	0.005 (0.44)

&Comparisons across five pairs of bootstrapped models

^Net Reclassification Index Score (Pepe Method), with a cut point prevalence of 5%

**Net Reclassification Index Score (Pepe Method), with a cut point prevalence of 10%

^^Net Reclassification Index Score (Pepe Method), with a cut point prevalence of 15%

Table 4

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

	Screening Pre-Test Probability Cutpoint[**]	Percent Screened[^]	Percent of Women with PVFx Detected	Prevalence PVFx Among Those Screened
Model 1[#]	5%	100%	100%	21.8%
	10%	98.3%	99.4%	22.0%
	15%	75.6%	85.6%	25.0%
Model 2[&]	5%	100%	100%	23.0%
	10%	91.9%	96.7%	21.7%
	15%	67.9%	83.1%	26.7%
Model 3[##]	5%	99.8%	99.9%	21.6%
	10%	88.4%	95.6%	23.3%
	15%	64.7%	82.1%	27.3%
Model 4[&&]	5%	99.7%	100%	21.6%
	10%	87.9%	95.6%	23.5%
	15%	64.7%	82.4%	27.5%

[%] Analyses limited to women with no self-reported prior spine fracture and with femoral neck BMD T-score ≤ −1.0

[*] 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, back pain over the past two years, and grip strength

[&&] Model 4: Age, Femoral Neck BMD, historical height loss, prior non-vertebral fracture, body mass index, back pain over the past two years, grip strength, smoking, estrogen use, and glucocorticoid use