

Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks

N. D. Nguyen · S. A. Frost · J. R. Center ·
J. A. Eisman · T. V. Nguyen

Received: 9 November 2007 / Accepted: 2 January 2008 / Published online: 7 March 2008

© International Osteoporosis Foundation and National Osteoporosis Foundation 2008

Abstract

Summary We have developed clinical nomograms for predicting 5-year and 10-year fracture risks for any elderly man or woman. The nomograms used age and information concerning fracture history, fall history, and BMD T-score or body weight.

Introduction Although many fracture risk factors have been identified, the translation of these risk factors into a prognostic model that can be used in primary care setting has not been well realized. The present study sought to develop a nomogram that incorporates non-invasive risk factors to predict 5-year and 10-year absolute fracture risks for an individual man and woman.

Methods The Dubbo Osteoporosis Epidemiology Study was designed as a community-based prospective study, with 1358 women and 858 men aged 60+ years as at 1989. Baseline measurements included femoral neck bone mineral density (FNBMD), prior fracture, a history of falls and body weight. Between 1989 and 2004, 426 women and 149 men had sustained a low-trauma fracture (not including morphometric vertebral fractures). Two prognostic models based on the Cox's proportional hazards analysis were considered: model I included age, BMD, prior fracture and falls; and model II included age, weight, prior fracture and fall.

Results Analysis of the area under the receiver operating characteristic curve (AUC) suggested that model I (AUC=0.75 for both sexes) performed better than model II (AUC=0.72 for women and 0.74 for men). Using the models' estimates, we constructed various nomograms for individualizing the risk of fracture for men and women. If the 5-year risk of 10% or greater is considered "high risk", then virtually all 80-year-old men with BMD T-scores <-1.0 or 80-year-old women with T-scores <-2.0 were predicted to be in the high risk group. A 60-year-old woman's risk was considered high risk only if her BMD T-scores \leq -2.5 and with a prior fracture; however, no 60-year-old men would be in the high risk regardless of their BMD and risk profile.

Conclusion These data suggest that the assessment of fracture risk for an individual cannot be based on BMD alone, since there are clearly various combinations of factors that could substantially elevate an individual's risk of fracture. The nomograms presented here can be useful for individualizing the short- and intermediate-term risk of fracture and identifying high-risk individuals for intervention to reduce the burden of fracture in the general population.

Source of financial support: This work has been supported by the National Health and Medical Research Council of Australia, and untied educational grants from GE-Lunar, Merck Australia, Eli Lilly International and sanofi-aventis Australia.

N. D. Nguyen · S. A. Frost · J. R. Center · J. A. Eisman ·
T. V. Nguyen (✉)

Bone and Mineral Research Program, Garvan Institute of Medical Research, St Vincent's Hospital,
384 Victoria Street, Darlinghurst,
Sydney, NSW 2010, Australia
e-mail: t.nguyen@garvan.org.au

N. D. Nguyen
e-mail: n.nguyen@garvan.org.au

S. A. Frost
e-mail: s.frost@uws.edu.au

J. R. Center
e-mail: j.center@garvan.org.au

J. A. Eisman
e-mail: j.eisman@garvan.org.au

J. A. Eisman · T. V. Nguyen
Faculty of Medicine, The University of New South Wales,
Sydney, NSW, Australia

Keywords Bone mineral density · Epidemiology · Fall · Fracture risk · Nomogram · Osteoporosis · Prior fracture

Introduction

During the past two decades, epidemiologic studies have identified an array of fracture risk factors [1]. These studies have also underscored the fact that osteoporotic fracture is a multifactorial disorder, making it difficult to discriminate those who will sustain a fracture from those who will not. This multifactorial nature suggests that the risk of fracture for an individual cannot be accurately estimated without considering effects of multiple risk factors. The challenge is how to develop workable prognostic models that incorporate relevant risk factors and identify those at high-risk of fracture [2].

A number of prognostic models have recently been developed [3–5] using bone mineral density measurement and common fracture risk factors such as age, weight or body mass index, smoking, chronic disease, use of central nervous system medication, and history of early menopause. However, all of these models are more applicable to groups of individuals rather than to an individual, because they were built based on the concept of risk stratification in which continuous variables were categorized into subgroups. For example, a good prognostic model should distinguish a 65-year-old woman with BMD T-score of -2.5 from a 67-year-old woman with a BMD T-score of -2.4, because the two women should theoretically have different risk levels even though the difference is modest. In fact, at the patient level, an individual is likely to be unique in risk profile and the risk of fracture for the individual should take that uniqueness into account [6].

Furthermore, all current prognostic models of fracture require complex computation, typically performed by computer software or a website, which may not be accessible or convenient in the primary care setting. A prognostic model that takes the form of a simple paper-based algorithm without data entry to a computer may be particularly useful [7].

For a risk factor to be useful in primary care, it has to ideally possess the following criteria: accuracy, reliability, responsiveness, and to be non-invasive. Among the fracture risk factors that have been identified during the past 30 years or so, advancing age, low body weight, low bone mineral density (BMD), a history of fractures, and falls fulfill these criteria [8–12]. The translation of these risk factors into clinical tools for primary care use has been applied for predicting of hip fracture risk [6], but not for any osteoporotic fracture risk. The present study sought to develop a nomogram that incorporates these non-invasive risk factors to predict five-year and ten-year

absolute risks of any osteoporotic fracture for an individual man and woman.

Materials and methods

Setting and subjects

This study is part of the on-going longitudinal Dubbo Osteoporosis Epidemiology Study (DOES), for which details of protocol and study design have been previously described [8, 13–15]. Briefly, in 1989, all men and women aged 60 or above (as of 1989) living in Dubbo, a city of approximately 32,000 people 400 km north west of Sydney (Australia), were invited to participate in an epidemiological study. At that time, the population comprised 1,581 men and 2,095 women aged ≥ 60 years, of whom 98.6% were Caucasian and 1.4% indigenous Aboriginal. These individuals were all invited to participate in DOES. This study was approved by the St Vincent's Campus Research Ethics Committee and informed written consent was obtained from each participant.

Dubbo had been selected for the study as the age and gender distribution of the population closely resembled the Australian population [14], and it is relatively isolated in terms of medical care, so that virtually complete ascertainment of all fractures in the target population is possible. During the follow-up period, approximately 5% of women were on anti-osteoporosis treatment, with the minority (4.5%) being prescribed calcium and vitamin D.

Ascertainment of fracture

Low trauma and non-pathological fractures were considered the primary outcome of this study. Fractures occurring during the study period were identified for residents of the Dubbo local government area through radiologists' reports from the two centers providing X-ray services as previously described [8, 13]. Fractures were only included if the report of fracture was definite and, on interview, had occurred with low trauma (e.g., fall from standing height or less). Fractures clearly due to major trauma (such as motor vehicle accidents), those due to underlying diseases (such as cancer or bone-related diseases) or those of digit, skull or cervical spine were excluded from the analysis. The present study did not include morphometric vertebral fractures.

Fractures were classified as any fracture (any first osteoporotic fracture), including those of hip, vertebrae (symptomatic), wrist, meta-carpal, humerus, scapula, clavicle, distal femur, proximal tibia, patella, pelvis and sternum. Not all individuals who sustained a fracture had had or agreed to have bone density measurements. The total number of individuals with

fracture reported in this study accounted for 92% of all fractured subjects from the entire DOES population.

Measurements of risk factors

Individuals were interviewed by a nurse coordinator who administered a structured questionnaire to obtain data, including age, life-style factors such as duration of smoking intake and alcohol consumption, physical activity, any history of falls in the preceding 12 months, any history of fractures after the age of 50, family history of osteoporosis, including maternal history and paternal history, condition of rheumatoid arthritis and a history of using corticosteroids. Current use of corticosteroids was defined as using oral prednisolone of ≥ 7.5 mg per day or equivalent dose of another corticosteroid for ≥ 6 months [16] either at the time of fracture or at the time of analysis to those without a fracture. Anthropometric variables (height, weight) were measured, and a dietary assessment was performed based on a frequency questionnaire for calcium intake, as described elsewhere [17].

At baseline, bone mineral density (BMD, g/cm^2) was measured at the lumbar spine or femoral neck (FN) by dual energy x-ray absorptiometry initially using a DPX densitometer (GE-LUNAR Corp, Madison, WI). The radiation dose with this method is < 0.1 μGy . The coefficient of reliability of BMD in our institution in normal subjects is 0.96 and 0.98 at the proximal femur and lumbar spine, respectively [18]. In the present analysis, baseline femoral neck BMD, not lumbar spine BMD, was used, since the femoral neck is minimally affected by degenerative changes that may artificially elevate BMD. For comparison with other studies, each subject was also classified as “osteoporotic”, with a BMD being 2.5 SD or more below the young normal level, “osteopenic”, with a BMD between 2.5 to 1.0 SD below the young normal level, or as “normal”, being 1.0 SD below or above. The “young normal” BMD was obtained from a referent database for Australian women [19]. The “young normal” BMD was obtained from a sample of Australian men and women aged between 20 to 32 years of age. These values are identical to those of LUNAR Caucasian database [20].

Development of predictive models

Given many potential risk factors, the number of “candidate models” for predicting fracture risk can be large. The Bayesian model average (BMA) [21] was applied to search for most parsimonious models with consistent and maximum discriminatory power. In terms of model consistency and accuracy, it has been shown that the BMA approach performed better than traditional algorithms, such as stepwise regression [22, 23], because it can account for

model uncertainty in both predictions and parameter estimates [21, 24].

The prognostic performance of parsimonious models was assessed by the area under the receiver operating characteristic (ROC) curve [25–28], which reflects the model’s ability to discriminate between those who will sustain a fracture from those who will not.

The Cox’s proportional hazards regression model [29, 30] was used to estimate the magnitude of association between fracture risk and the risk factors. The association between each risk factor and fracture risk was initially expressed by the hazard ratio (HR) and its 95% confidence interval (CI) for each standard deviation (SD) or unit change with ordinal risk factors.

From the estimates of the models’ parameters, a series of nomograms was constructed for predicting 5-year and 10-year fracture risks for men and women separately. The nomograms were internally validated by the bootstrap method, which assesses how accurately the model will predict fracture in a new similar sample of subjects. In this method, 1,000 sub-samples, each with 150 subjects, of the entire sample were repeatedly re-sampled (with replacement) and analyzed, from which biases can be estimated [31]. The predictive accuracy of a model was assessed by the concordance index [31], which is the probability that, given two randomly drawn individuals, the individual who will sustain a fracture first had a lower probability of non-fracture. The interpretation of the concordance index is similar to that of the area under the receiver operating characteristic curve [25–28]. In a further validation, the predicted probability of fracture was compared with the actual probability (i.e., nomogram calibration) on the entire sample, again using 1,000 bootstrap resample to reduce overfit bias, which would overstate the accuracy of the nomogram. All analyses were performed using the R language on the Windows platform [32] with the Design and Hmisc packages [31, 33].

Results

Characteristics of participants and univariate analysis

Data were analyzed from 1358 women and 858 men who had been followed up between 1989 and 2004. The median duration of follow-up was 13 years (inter-quartile range, IQR: 8–14) for women and 12 years (7–14) for men, yielding 14,443 and 8,695 person-years of observation in women and men, respectively. The average age (\pm SD) at baseline was 71 ± 8 years for women and 70 ± 6 for men.

During the follow-up period, 426 women and 149 men sustained at least one fracture; making the overall incidence of fracture 35 per 1,000 person-years in women and 18 per

Table 1 Characteristics of any fracture and non-fracture individuals

Risk factor	Any fracture	Non-fracture	Hazard ratio for risk factor		P-value
			Unit	HR (95% CI)	
Women (n)	426	932			
Age (y)	73±8	70±7	+5 y	1.43(1.34, 1.53)	<0.0001
Weight (kg)	63±12	66±13	-10 kg	1.25(1.15, 1.37)	<0.0001
Height (cm)	159±7	160±6	-5 cm	1.19(1.10, 1.29)	<0.0001
BMI (kg/m ²)	25±4	26±5	-5 kg/m ²	1.25(1.11, 1.40)	0.0002
FNBMd (g/cm ²)	0.73±0.12	0.80±0.13	-0.12 g/cm ²	1.68(1.55, 1.82)	<0.0001
LSBMd (g/cm ²)	0.96±0.18	1.05±0.19	-0.2 g/cm ²	1.68(1.49, 1.89)	<0.0001
Daily calcium intake (g/d) ^a	569 (399, 803)	623 (332, 563)	-350 mg/d	0.94(0.85, 1.04)	0.2178
Home physical activity (METs) ^a	77 (59, 104)	79 (62, 100)	-30METs	0.97(0.88, 1.07)	0.5309
Current/ex-smoking (yes)	126 (29.6)	269 (28.8)	Yes	1.05(0.85, 1.29)	0.6782
Multiple fracture (n,%) ^b			Plus one	2.06(1.87, 2.26)	<0.0001
1	104 (24.4)	64 (6.9)	Fracture		
2	29 (9.2)	4 (0.4)			
≥3	34 (5.6)	0			
Fall in the last 12 mo (n,%)			Plus one fall	1.23(1.10, 1.38)	0.0005
1	97 (22.8)	132 (14.2)			
2	32 (7.5)	49 (5.3)			
≥3	10 (2.4)	33 (3.5)			
Current use of steroid (n,%)	17 (3.9)	19 (2.1)	Yes	1.34(0.35, 2.20)	0.2202
Rheumatoid arthritis (n,%)	27 (6.3)	36 (3.9)	Yes	1.49(1.00, 2.19)	0.047
Maternal history ^c	80 (18.8)	143 (15.3)	Yes	1.07(0.83, 1.36)	0.6221
Paternal history ^c	16 (3.8)	34 (3.7)	Yes	0.97(0.59, 1.59)	0.8895
Men (n)	149	709			
Age (y)	73±7	70±6	+5 y	1.67(1.48, 1.88)	<0.0001
Weight (kg)	75±13	79±13	-10 kg	1.31(1.14, 1.51)	0.0002
Height (cm)	171±7	174±7	-5 cm	1.28(1.13, 1.44)	<0.0001
BMI (kg/m ²)	25±4	26±4	-5 kg/m ²	1.30(1.03, 1.65)	0.0282
FNBMd (g/cm ²)	0.83±0.16	0.93±0.15	-0.12 g/cm ²	1.62(1.43, 1.83)	<0.0001
LSBMd (g/cm ²)	1.16±0.21	1.26±0.21	-0.2 g/cm ²	1.54(1.30, 1.82)	<0.0001
Daily calcium intake (g/d) ^a	541 (351, 811)	549 (410, 802)	-350 mg/d	1.08(0.90, 1.31)	0.4129
Home physical activity (METs) ^a	79 (44, 110)	67 (41, 102)	-30METs	0.90(0.82, 0.99)	0.0422
Current/ex-smoking (yes)	94 (63.1)	436 (61.5)	Yes	1.14(0.82, 1.59)	0.4350
Multiple fracture (n,%) ^b			Plus one	2.92(2.43, 3.52)	<0.0001
1	33 (22.15)	27 (3.8)	Fracture		
≥2	9 (6.0)	4 (0.6)			
≥3	5 (3.4)	0			
Fall in the last 12 mo (n,%)			Plus one fall	1.38(1.13, 1.69)	0.0016
1	27 (18.1)	69 (9.7)			
2	7 (4.7)	28 (4.0)			
≥3	5 (3.4)	15 (2.1)			
Current use of steroid (n,%)	5 (3.4)	21 (3.0)	Yes	0.92(0.38, 2.23)	0.8493
Rheumatoid arthritis (n,%)	4 (2.7)	14 (1.6)	Yes	1.23(0.47, 3.45)	0.6296
Maternal history ^c	21 (14.1)	91 (12.9)	Yes	1.03(0.65, 1.63)	0.9104
Paternal history ^c	1 (0.7)	20 (2.8)	Yes	0.26(0.04, 1.84)	0.1769

Values are mean ± SD, unless otherwise specified

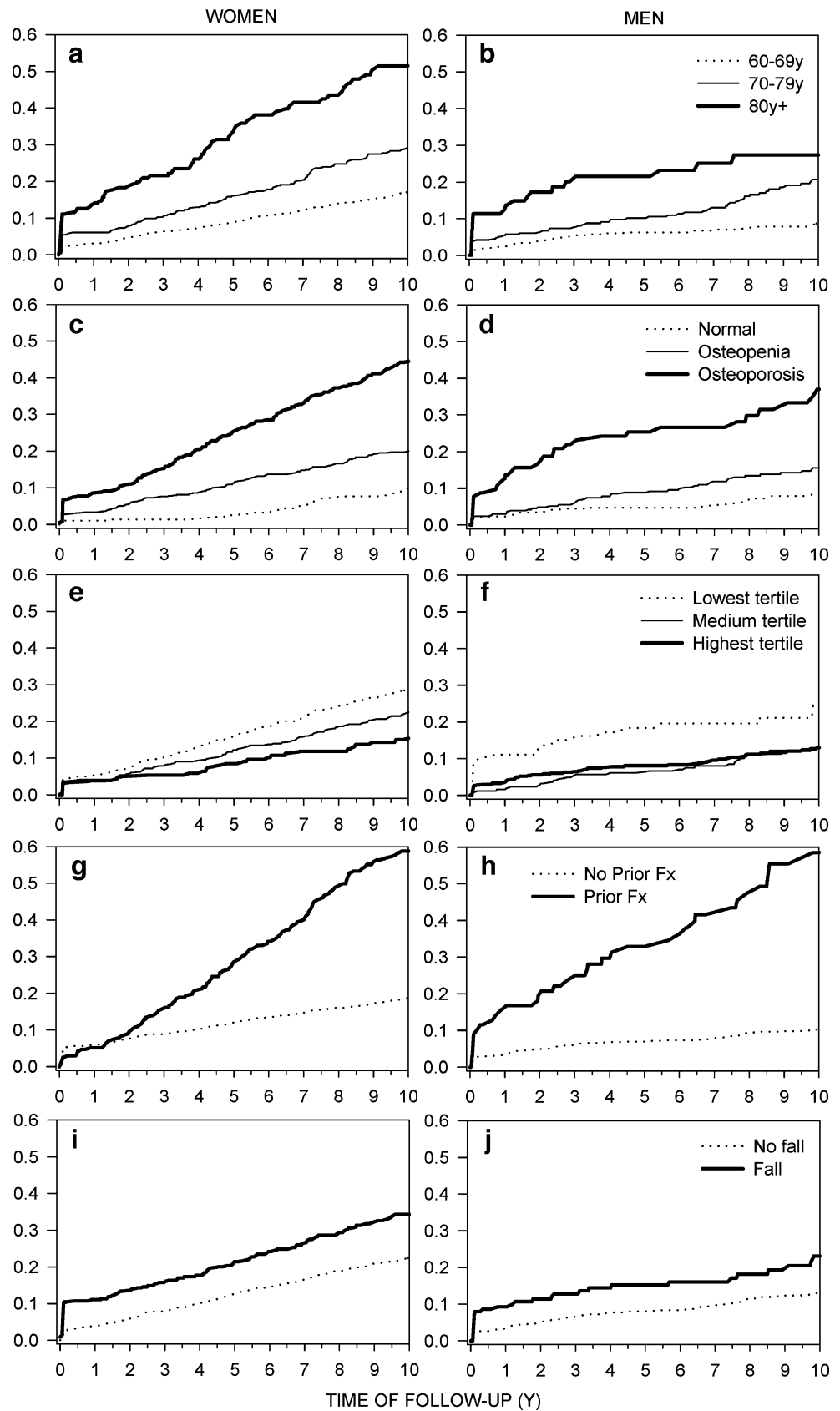
^amedian (Q1, Q3)

^bNumbers of any fracture from the age of 50 years

^cMaternal and paternal history of osteoporosis

BMI, body mass index; FNBMd, femoral neck bone mineral density; LSBMD, lumbar spine bone mineral density; METs, metabolic equivalents

Fig. 1 Cumulative probability of fracture in women (left panel) and men (right panel) classified by age group (**a, b**), BMD category (**c, d**), body weight tertile (**e, f**), history of fracture (**g, h**), and falls during the past 12 months (**i, j**)



1,000 person-years in men. In both sexes, the most common sites of fracture were symptomatic vertebral (28% in women and 34% in men), hip (17% for both sexes), forearm (2% in women and 4% in men), rib (5% in women and 23% in men). In univariate analysis, advancing age, lower weight, shorter current height, lower BMI, lower BMD, prior fracture, and fall during the past 12 months were each associated with increased fracture risk in men and women (Table 1 and Fig. 1).

In addition, there were no significant differences between those with and without a fracture in terms of current use of corticosteroid, a condition of rheumatoid arthritis, a parental history of osteoporosis and lifestyle factors, such as daily calcium intake, physical activity and smoking status (Table 1).

Development models for predicting fracture risk

All potential risk factors were included in the analysis by using Bayesian model average to select optimal variables for predicting fracture risk. It was confirmed that the five risk factors: age, femoral neck BMD T-scores, weight, prior fracture, and falls during the past 12 months were most significantly associated with fracture risk. Two most parsimonious models were considered: Model I includes age, BMD T-scores, prior fracture and fall, and Model II replaces the BMD variable in Model I with body weight.

In multivariable analysis with sex being considered a covariate, the hazard ratio (95% CI) of fracture relating to BMD was 1.36 (1.26–1.47) for each SD lower, age 1.25 (1.17–1.30) for each SD older, prior fracture 1.84 (1.68–2.03), previous falls 1.22 (1.11–1.36). For model II, the hazard ratio of fracture was: body weight 1.10 (1.02–1.19), age 1.34 (1.26–1.43), prior fracture 2.05 (1.87–2.25), previous falls 1.22 (1.10–1.36).

The prognostic performance as assessed by the area under the receiver operating characteristic curve (AUC) is shown in Table 2. Model I (which included age, femoral neck BMD, prior fracture, and fall) performed better than model II (which replaced BMD in model I with body weight). In women, the AUC value of model I was 0.76,

which was significantly better ($p=0.0007$) than model II's (AUC=0.72). In men, there was no statistically significant difference in discrimination ($p=0.146$) between model I (AUC=0.75) and model II (AUC=0.74). In men, the additional of age (model III) significantly improved the discriminatory power from the model with BMD alone (model IV).

When the continuous BMD variable in model I was replaced by categorized BMD (normal, osteopenia, and osteoporosis according to the WHO classification system), the AUC value for model I reduced to 0.717 (SE 0.016) for women and 0.704 (SE 0.025) for men, which was significantly lower ($p=0.001$ for women and $p=0.018$ for men) than the model with continuous BMD measurements. Thus, models based on continuous measurements performed consistently better than the models based on categorical measurements.

Using the parameter estimates of model I and model II, we constructed two nomograms for women (Figs. 2 and 3) and two nomograms for men (Figs. 4 and 5). Each of these nomograms predicts the risk of fracture within 5 years and 10 years. In women, internal validation by the bootstrap method suggested that the bias-corrected estimate of predictive discrimination of 0.51 for model I and 0.43 for model II; the corresponding estimates in men were 0.50 for model I and 0.47 for model II. The maximum calibration error in predicting probability of fracture was about 0.4% for both models in women and 0.6–1.9% in men.

Some typical point estimates of 5-year and 10-year risks are presented for model I (Table 3) and for model II (Table 4). For illustration, 5-year risk of more than 10% and 10-year risk of greater than 20% are highlighted. As expected from a multivariable prognostic model, there are multiple combinations of risk factors that can raise an individual's risk. For example, if the 5-year absolute risk of 10% or greater is considered "high risk" then, virtually all women and men aged 80+ with BMD T-scores being less than -1.0 and regardless of their prior fracture and fall are at high risk. Women aged 70 years old with BMD T-scores being less than -2.0 can also be considered high-risk regardless of their prior fracture and fall history. Virtually

Table 2 Area under the receiver operating characteristic curves of different models

	Women		Men	
	AUC (SE)	P-value	AUC (SE)	P-value
Model I: age+baseline BMD+prior fracture+fall	0.757 (0.015)	0.0007	0.754 (0.024)	0.1436
Model II: age+baseline weight+prior fracture+fall	0.721 (0.016)	0.0026	0.739 (0.024)	0.0240
Model III: age+baseline BMD	0.669 (0.016)	0.3755	0.698 (0.024)	0.0191
Model IV: baseline BMD only	0.668 (0.016)	-	0.661 (0.026)	-

AUC, area under the curve; SE, standard error
P-value was between each model with model IV

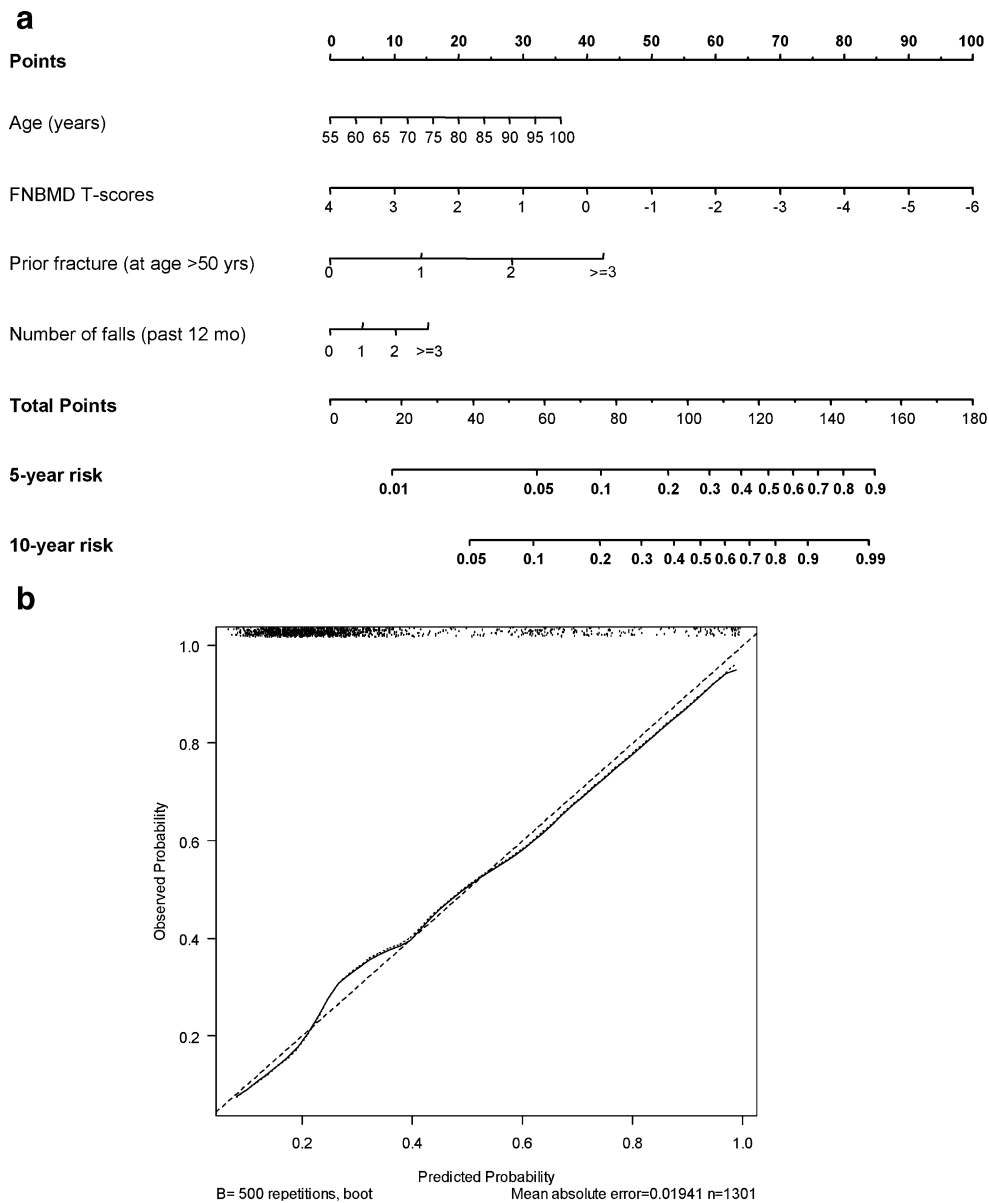


Fig. 2 a Nomogram for predicting the 5-year and 10-year probability of any fracture for a woman, based on Model I. Instruction for usage: Mark the age of an individual on the “Age” axis and draw a vertical line to the “Point” axis to determine how many points toward the probability of fracture the individual receives for his/her age value. Repeat the process for each additional risk factor. Sum the points of the risk factors. Locate the final sum on the “Total points” axis. Draw a vertical line down to the 5-year or 10-year risk line to find the individual’s probability of sustaining a fracture within next 5 or 10 years. Example: Mrs. A, 70 years old, has a BMD T-score of -2.5, had a prior fracture and a fall

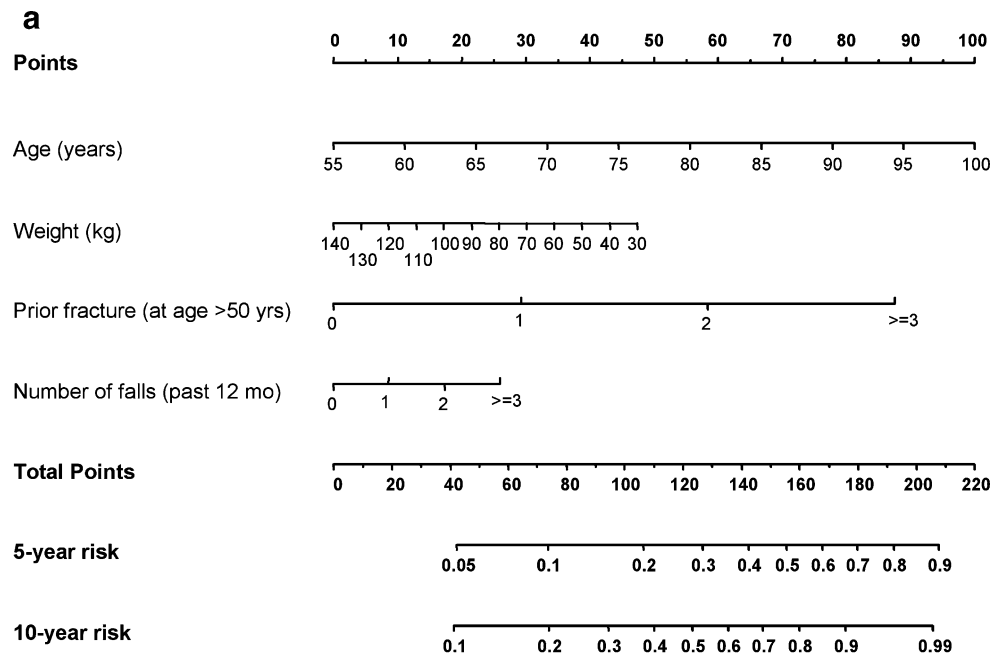
in the past 12 months; her points for age is approximately 12, her BMD points is 65; prior fracture point is 13 and fall point is 4. Her total points are, therefore, 12+65+13+4=94, and her probability of having a hip fracture is around 0.21 in the next 5 years and 0.40 in the next 10 years. In other words, in 100 women like her, one would expect 21 and 40 of them will have a fracture in the next 5 years and next 10 years, respectively. **b:** Calibration of nomogram for any fracture for women (Model I). The diagonal dotted line indicates reference line on which ideal nomogram would lie (perfect prediction). Solid line indicates current nomogram performance

all men aged 70 years and with a prior fracture can also be considered high risk, regardless of their BMD measurement. Among women aged 60 years, their 5-year risk of fracture is greater than 10% only when their T-scores are ≤ -3.0 . None of 60-year-old men have 5-year risk of fracture greater than 10%.

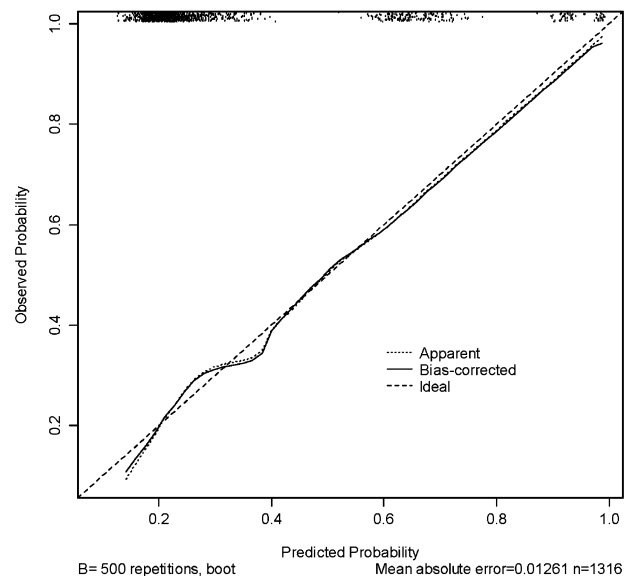
Discussion

Despite the fact that several fracture risk factors have been identified by epidemiologic research, the synthesis of these risk factors into a prognostic model for clinical use has not been realized. Part of the problems is that many risk factors

Fig. 3 a Nomogram for predicting the 5-year and 10-year probability of any fracture for a woman, based on Model II. Instruction for usage is similar to Fig. 2a. **b**: Calibration of nomogram for any fracture for women, model II. The diagonal dotted line indicates reference line on which ideal nomogram would lie (perfect prediction). Solid line indicates current nomogram performance



b

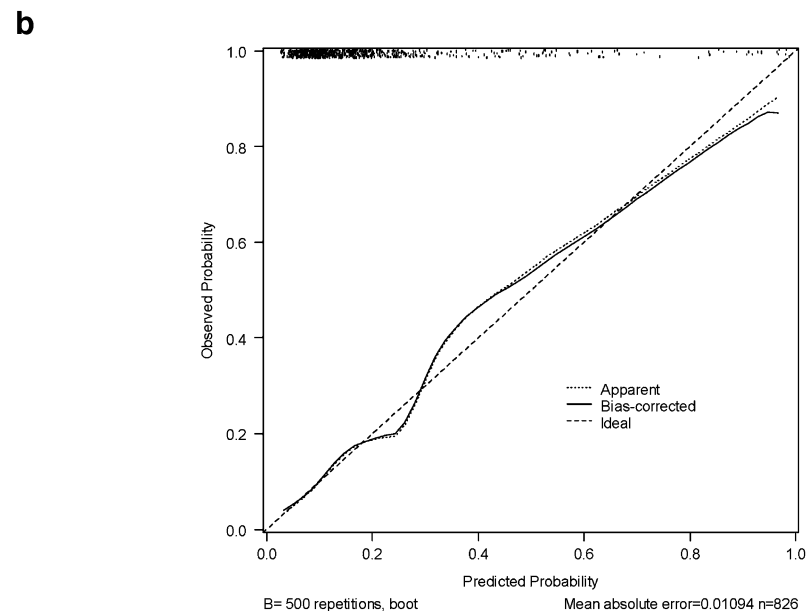
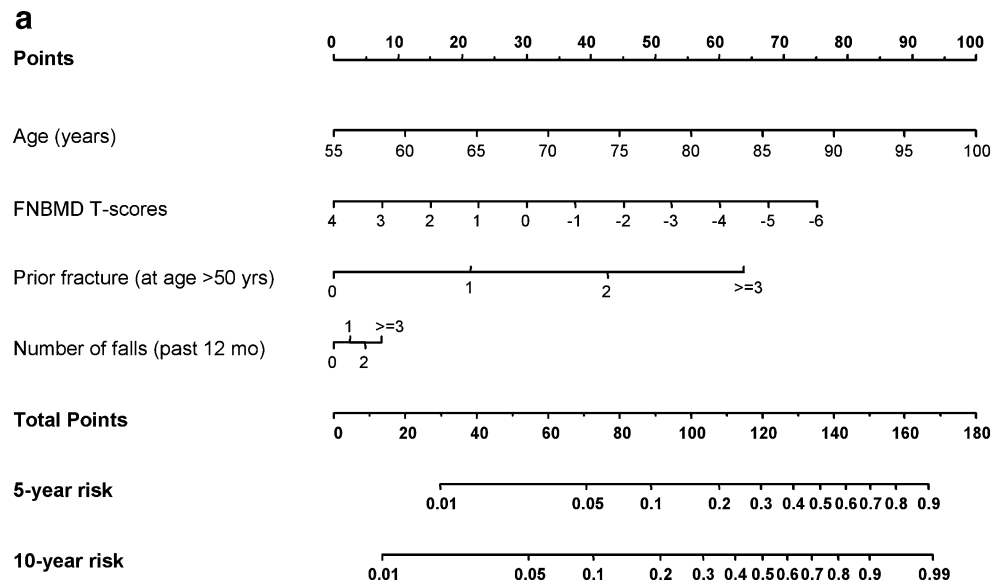


require sophisticated measurements that are beyond resources of a typical primary care setting. Thus, a practically useful prognostic model should make use of relevant clinical data that are easily obtained from an individual. The present study, built on previous research of risk factors, analyzed two models of prognosis, which incorporate the established risk factors of age, prior fracture, history of falls, and BMD or body weight. While weight is highly correlated with BMD, it is not surprising to observe that the model with the clinical risk factors and BMD performed better than the model with the clinical risk factors and body weight, because BMD has been shown to be more sensitive and specific in terms of fracture prediction [34, 35].

However, the difference in predictive accuracy between the two models is modest, and given the relatively low incidence of fracture, the difference is of limited practical importance.

The ultimate aim of developing a prognostic model is to provide clinicians and each individual with their risk estimate to guide clinical decisions. At present, individuals with low bone mineral density (i.e., T-scores being less than -2.5) or with a history of prior low trauma fracture are recommended for therapeutic intervention [36, 37]. This recommendation is logical and appropriate, since these individuals—as shown in this study and previous studies [12, 38]—have higher risk of fracture, and treatment can

Fig. 4 a Nomogram for predicting the 5-year and 10-year probability of any fracture for an elderly man, based on Model I. Instruction for usage is similar to Fig. 2a. **b**: Calibration of nomogram for any fracture for men (Model I). The diagonal dotted line indicates reference line on which ideal nomogram would lie (perfect prediction). Solid line indicates current nomogram performance

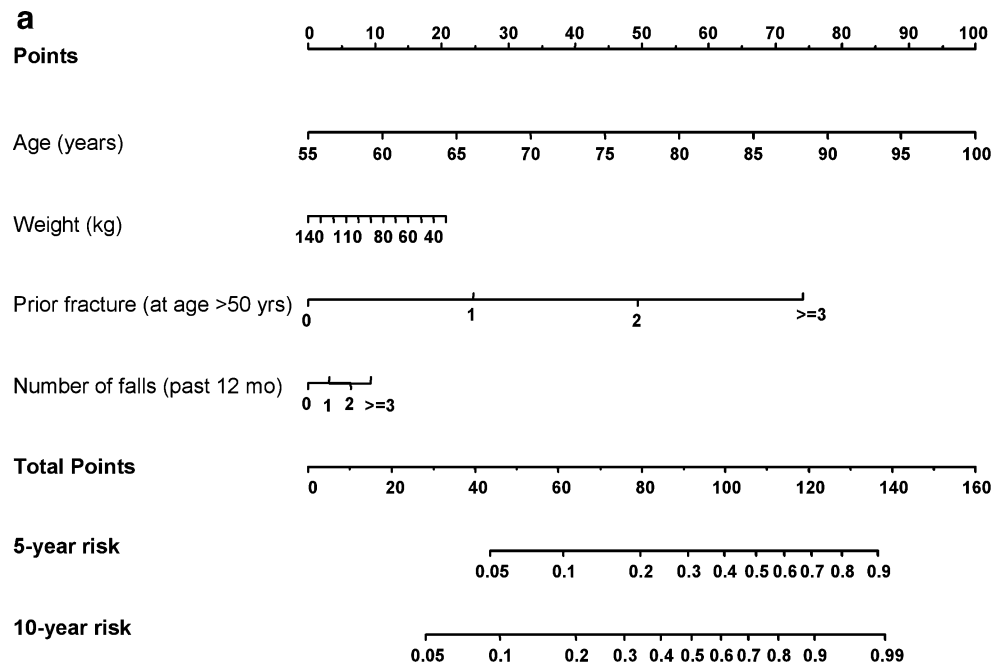


reduce their risk of fracture [39–41]. However, because fracture is a multifactorial event, there is more than one way that an individual can attain the risk conferred by either low BMD or a prior fracture. Indeed, virtually all women aged 70 years with BMD T-scores less than -1.5 and all 80-year-old men with BMD T-scores less than -1.0 can be considered “high risk”. On the other hand, no 60-year-old men or women without a prior fracture and a fall are considered high risk, even when their BMD T-scores were below -2.5. This demonstrates the informativeness of a multivariable prognostic model, and the limitation of a risk stratification-based approach for risk assessment for an individual.

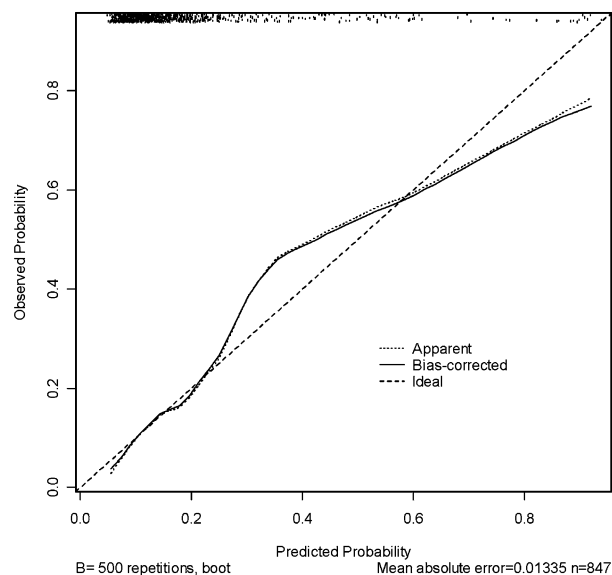
Other clinical factors, such as corticosteroid use, family history of fracture and a condition of underlying rheuma-

toid arthritis, have been shown to be predictors for fracture risk [42–44]. However, in the present study, there were no significant associations between those risk factors and fracture risk in both women and men. Therefore, these factors were not included in the prognostic model, as they did not significantly contribute to the predictiveness of fracture risk. It is likely, nevertheless, that a family history of fracture may in general improve the risk prediction. Vertebral deformity (morphometric fracture) has been shown to be a risk factor for fracture [45]. In the present study, vertebral deformity was not assessed at baseline for all participants; therefore, it was not included in the prognostic models. Nevertheless, there is room for further improvement in the prognosis of fracture by incorporating these clinical risk factors into the prognostic model.

Fig. 5 a Nomogram for predicting the 5-year and 10-year probability of any fracture for a man, based on Model II. Instruction for usage is similar to Fig. 2a. **b**: Calibration of nomogram for any fracture for men, model II. The diagonal dotted line indicates reference line on which ideal nomogram would lie (perfect prediction). Solid line indicates current nomogram performance



b



Each individual is important and unique. Individualization of risk—or the prediction of risk for an individual given a risk profile—is a fundamental aspect of the present models. The present models considered all continuous risk factors (e.g., BMD, body weight and age) in their original units of measurement. This consideration is different from previous models [3, 4], which categorized continuous risk factors into distinct groups based on some thresholds. While the categorization is appealing for its simplicity, it implicitly assumes a discontinuous relationship, which is unlikely to be true for well-known risk factors, such as BMD and body weight. Such a categorization is also known to reduce statistical power [46, 47]. Furthermore, the

risk estimates based on categorization of continuous risk factors can only be applied to a group of individuals, not to an individual. Prognosis is about imparting information of fracture risk to an individual and each individual is a unique case, because there exists no “average individual” in the population. The more risk factors are considered, the greater likelihood of uniqueness of an individual’s profile being defined. Therefore, by modeling risk factors in their continuous scale the present models can be uniquely tailored to an individual.

The idea of using a nomogram to develop a prognosis model for an individual is not new with more than 1,700 nomograms being advocated [48]. Several nomograms

Table 3 Five-year and 10-year risks of any fracture for a woman and for a man, based on age, BMD T-scores, a history of prior fracture and falls (Model I)

Age	T-scores	5-year risk (%)				10-year risk (%)				
		No prior fracture		Prior fracture		No prior fracture		Prior fracture		
		No fall	Fall	No fall	Fall	No fall	Fall	No fall	Fall	
Women										
60	0.0	2.9	3.5	5.0	6.1	6.1	7.4	10.5	12.7	
	-1.0	4.3	5.2	7.4	9.0	8.9	10.9	15.3	18.4	
	-1.5	5.2	6.3	9.0	10.9	10.8	13.1	18.3	22.0	
	-2.0	6.3	7.7	10.9	13.2	13.1	15.8	21.9	26.2	
	-2.5	7.7	9.3	13.2	15.9	15.8	19.0	26.1	31.0	
	-3.0	9.3	11.3	15.9	19.1	18.9	22.7	30.9	36.5	
	70	0.0	3.9	4.8	6.9	8.4	8.3	10.1	14.2	17.1
		-1.0	5.8	7.1	10.1	12.2	12.1	14.7	20.4	24.4
		-1.5	7.1	8.6	12.2	14.7	14.6	17.6	24.4	29.0
		-2.0	8.6	10.5	14.7	17.7	17.6	21.1	28.9	34.2
80	-2.5	10.4	12.6	17.7	21.2	21.0	25.1	34.1	40.1	
	-3.0	12.6	15.2	21.2	25.3	25.1	29.8	40.0	46.5	
	0.0	5.4	6.6	9.3	11.3	11.2	13.6	19.0	22.7	
	-1.0	8.0	9.7	13.6	16.5	16.3	19.6	27.0	32.0	
	-1.5	9.7	11.7	16.4	19.7	19.6	23.4	31.9	37.6	
	-2.0	11.7	14.1	19.7	23.6	23.4	27.9	37.5	43.8	
	-2.5	14.1	17.0	23.5	28.0	27.8	32.9	43.7	50.6	
	-3.0	16.9	20.3	28.0	33.1	32.8	38.6	50.5	57.8	
	Men									
	60	0.0	1.6	1.7	3.6	4.0	3.0	3.3	6.8	7.5
-1.0		2.1	2.3	4.8	5.3	4.0	4.4	9.0	9.9	
-1.5		2.4	2.7	5.6	6.1	4.6	5.1	10.4	11.4	
-2.0		2.8	3.1	6.5	7.1	5.3	5.9	12.0	13.1	
-2.5		3.3	3.6	7.5	8.2	6.2	6.8	13.8	15.1	
-3.0		3.8	4.2	8.6	9.5	7.1	7.8	15.8	17.3	
70		0.0	3.7	4.1	8.5	9.3	7.0	7.7	15.6	17.1
		-1.0	5.0	5.5	11.3	12.4	9.4	10.3	20.5	22.3
		-1.5	5.8	6.4	13.0	14.2	10.8	11.8	23.4	25.4
		-2.0	6.7	7.4	14.9	16.3	12.4	13.6	26.6	28.9
80	-2.5	7.7	8.5	17.1	18.7	14.3	15.6	30.1	32.7	
	-3.0	8.9	9.8	19.6	21.3	16.4	17.9	34.1	36.8	
	0.0	8.8	9.7	19.3	21.1	16.2	17.7	33.7	36.5	
	-1.0	11.7	12.8	25.2	27.4	21.2	23.1	42.6	45.7	
	-1.5	13.4	14.7	28.6	31.0	24.1	26.3	47.5	50.8	
	-2.0	15.4	16.9	32.3	35.0	27.4	29.8	52.6	56.1	
	-2.5	17.7	19.3	36.5	39.4	31.1	33.7	58.0	61.6	
	-3.0	20.2	22.1	40.9	44.0	35.1	37.9	63.5	67.1	

developed and used in the field of oncology exhibited better performance than risk-grouping categorization [49, 50].

Statistical prognostic models have been shown to outperform clinical judgment [51], because they can objectively incorporate many risk data. Because of their objectivity, multivariable prognostic models can reduce the variability in risk estimates.

No prognostic models are able to discriminate all individuals who will fracture from those who will not, and the current models are not an exception. The accuracy of a predictive model is normally quantified by its capacity to separate individuals who will sustain a fracture along a continuum from those who will not. In terms of the capacity of discrimination, both models proposed in this paper performed reasonably well with the area under the

Table 4 Five-year and 10-year risks of any fracture for a woman and for a man, based on age, body weight, a history of prior fracture and falls (Model II)

Age	Weight (kg)	5-year risk (%)				10-year risk (%)				
		No prior fracture		Prior fracture		No prior fracture		Prior fracture		
		No fall	Fall	No fall	Fall	No fall	Fall	No fall	Fall	
Women										
60	100	3.7	4.4	7.0	8.5	7.5	9.1	14.2	17.0	
	90	4.0	4.9	7.7	9.3	8.3	10.0	15.5	18.6	
	80	4.4	5.4	8.5	10.3	9.1	11.0	17.0	20.3	
	70	4.9	5.9	9.3	11.3	10.0	12.0	18.6	22.2	
	60	5.4	6.5	10.3	12.4	11.0	13.2	20.3	24.2	
	50	5.9	7.2	11.3	13.6	12.0	14.5	22.2	26.4	
	70	100	6.0	7.3	11.4	13.8	12.2	14.7	22.5	26.7
		90	6.6	8.0	12.5	15.1	13.4	16.1	24.5	29.0
		80	7.3	8.8	13.7	16.5	14.7	17.6	26.7	31.5
		70	8.0	9.7	15.1	18.1	16.1	19.2	29.0	34.2
80	60	8.8	10.6	16.5	19.7	17.6	21.0	31.5	37.0	
	50	9.7	11.7	18.0	21.6	19.2	22.9	34.2	39.9	
	100	9.8	11.8	18.3	21.9	19.5	23.3	34.6	40.4	
	90	10.8	13.0	20.0	23.8	21.3	25.3	37.4	43.5	
	80	11.8	14.2	21.8	26.0	23.3	27.6	40.4	46.8	
	70	13.0	15.6	23.8	28.2	25.3	30.0	43.5	50.2	
	60	14.2	17.1	26.0	30.7	27.6	32.5	46.8	53.7	
	50	15.6	18.7	28.2	33.3	30.0	35.2	50.2	57.2	
	Men									
	60	100	1.8	2.1	5.0	5.6	3.4	3.9	9.1	10.2
90		2.0	2.3	5.3	6.1	3.7	4.2	9.7	11.0	
80		2.1	2.4	5.8	6.5	4.0	4.5	10.5	11.8	
70		2.3	2.6	6.2	7.0	4.3	4.8	11.3	12.7	
60		2.5	2.8	6.7	7.6	4.6	5.2	12.1	13.6	
50		2.7	3.0	7.2	8.1	5.0	5.6	13.0	14.6	
70		100	4.5	5.1	11.8	13.3	8.2	9.3	21.0	23.5
		90	4.8	5.5	12.7	14.3	8.9	10.0	22.4	25.1
		80	5.2	5.9	13.7	15.4	9.5	10.8	24.0	26.8
		70	5.6	6.4	14.7	16.5	10.2	11.6	25.6	28.6
80	60	6.1	6.9	15.7	17.7	11.0	12.4	27.4	30.5	
	50	6.5	7.4	16.9	18.9	11.8	13.3	29.2	32.4	
	100	10.8	12.1	26.8	29.9	19.2	21.5	44.2	48.5	
	90	11.6	13.0	28.6	31.8	20.5	23.0	46.7	51.1	
	80	12.4	14.0	30.5	33.9	22.0	24.6	49.3	53.8	
	70	13.4	15.0	32.5	36.0	23.5	26.2	52.0	56.6	
	60	14.3	16.1	34.6	38.2	25.1	28.0	54.7	59.3	
	50	15.4	17.3	36.7	40.6	26.8	29.8	57.5	62.2	

ROC curve being 0.75, which is considered clinically acceptable [52]. Although the AUC is a useful measure of discrimination, it does not provide much insight into calibration, which refers to the concordance between predicted and actual probabilities [53]. In the present study, using the bootstrap technique, it was shown that the overfitting-corrected calibration was very good, particularly in women with the mean absolute calibration errors being 1–2% for all models. In summary, although the models are not perfect, their predictive accuracy and capacity of discrimination would be considered clinically useful in the real-world situation.

The predicted risk of fracture is a continuous probabilistic variable ranging from 0 to 1. This raises the issue of selecting an optimal cut-off predicted probability to classify an individual into fracture or non-fracture. This is not an easy task, because the cut-off value—if it exists at all—depends on the complex risk-benefit consideration, and perhaps more importantly, an individual's perception of risk, which is beyond the scope of the present study. The level of predicted risk at which that individual is prepared to take action is dependent on the risk perception of the individual, which is not easily quantified. Nevertheless, the predicted probability of fracture from the present prognostic models can be viewed as a measure of severity of osteoporosis for an individual.

It is logical that individuals with high risk of fracture, regardless of their BMD levels, should be considered for treatment because there was evidence suggesting that treating these individuals could yield clinical benefit [54]. However, at what level (or levels) of risk should be regarded as “high risk”, so that an intervention can be considered cost-effective? The issue of cost-effectiveness has been shown to be a function of national per-capita income, age, and future costs excluded [55]. For example, it was estimated that for a 50-year-old Australian woman, treatment is considered cost-effective if the woman's 10-year risk of hip fracture is at least 1.93% [55]. However, for a 90-year-old woman, the treatment is cost-effective if her 10-year risk is 10.8% or higher. The present nomograms can help identify such individuals for intervention.

All prognosis models are preliminary in the sense that they are yet to be tested in a totally independent population. The present models, as well as previous models of risk assessment developed from observational studies, may not be readily extrapolated or generalized to other populations, because the population under study might differ from the general population. Therefore, both external and internal validation should be a priority of research in the application of risk assessment models.

An important weakness of prognostic models of fracture risk is that they are based on a single measurement of risk factors, with the underlying but not stated assumption that

the risk factors do not change with time. Obviously, this assumption is not true in many risk factors such as BMD and body weight that are known to decline or change with time. Moreover, the rates of decline in BMD varied substantially among individuals. Therefore, one important aspect of future model development should take the time-varying nature of risk factors into account to achieve a better estimate of risk for an individual.

Notwithstanding these common limitations, the present models have the advantage of flexibility and ease of use without tedious computation, which can be impractical in primary care settings or inappropriate categorization. The present models visually translate clinical risk factors into a paper-based nomogram so that it can be readily used in clinical practice.

Acknowledgments We gratefully acknowledge the assistance of Janet Watters, Donna Reeves, Shaye Field and Jodie Ratleg for the interview, data collection and measurement bone mineral density. We also appreciate the invaluable help of the staff of Dubbo Base Hospital. We thank David Hayes and the IT group of the Garvan Institute of Medical Research for the management of the database.

Conflict of interest statement Dr. John Eisman serves as a consultant and receives corporate appointment from Amgen, deCode, Eli Lilly and Company, GE-Lunar, Merck Sharp & Dohme Ltd., Novartis, Organon, Roche-GSK, sanofi-aventis and Servier. All other authors have no conflict of interest.

References

- Cummings SR, Melton LJ (2002) Epidemiology and outcomes of osteoporotic fractures. *Lancet* 359:1761–1767
- Raisz LG (2005) Pathogenesis of osteoporosis: concepts, conflicts, and prospects. *J Clin Invest* 115:3318–3325
- Black DM, Steinbuch M, Palermo L, Dargent-Molina P, Lindsay R, Hoseyni MS, Johnell O (2001) An assessment tool for predicting fracture risk in postmenopausal women. *Osteoporos Int* 12:519–528
- van Staa TP, Geusens P, Kanis JA, Leufkens HG, Gehlbach S, Cooper C (2006) A simple clinical score for estimating the long-term risk of fracture in post-menopausal women. *QJM* 99:673–682
- Ettinger B, Hillier TA, Pressman A, Che M, Hanley DA (2005) Simple computer model for calculating and reporting 5-year osteoporotic fracture risk in postmenopausal women. *J Womens Health (Larchmt)* 14:159–171
- Nguyen ND, Frost SA, Center JR, Eisman JA, Nguyen TV (2007) Development of a nomogram for individualizing hip fracture risk in men and women. *Osteoporos Int* 18:1109–1117
- Wyatt JC, Altman DG (1995) Commentary: prognostic models: clinically useful or quickly forgotten? *BMJ* 311:1539–1541
- Nguyen T, Sambrook P, Kelly P, Jones G, Lord S, Freund J, Eisman J (1993) Prediction of osteoporotic fractures by postural instability and bone density. *BMJ* 307:1111–1115
- Marshall D, Johnell O, Wedel H (1996) Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* 312:1254–1259

10. Johnell O, Kanis JA, Oden A, Johansson H, De Laet C, Delmas P, Eisman JA, Fujiwara S, Kroger H, Mellstrom D, Meunier PJ, Melton LJ 3rd, O, Neill T, Pols H, Reeve J, Silman A, Tenenhouse A (2005) Predictive value of BMD for hip and other fractures. *J Bone Miner Res* 20:1185–1194
11. Kanis JA, Johnell O, De Laet C, Johansson H, Oden A, Delmas P, Eisman J, Fujiwara S, Garnero P, Kroger H, McCloskey EV, Mellstrom D, Melton LJ, Pols H, Reeve J, Silman A, Tenenhouse A (2004) A meta-analysis of previous fracture and subsequent fracture risk. *Bone* 35:375–382
12. Nguyen ND, Pongchaiyakul C, Center JR, Eisman JA, Nguyen TV (2005) Identification of high-risk individuals for hip fracture: a 14-year prospective study. *J Bone Miner Res* 20:1921–1928
13. Jones G, Nguyen T, Sambrook PN, Kelly PJ, Gilbert C, Eisman JA (1994) Symptomatic fracture incidence in elderly men and women: the Dubbo Osteoporosis Epidemiology Study (DOES). *Osteoporos Int* 4:277–282
14. Simons LA, McCallum J, Simons J, Powell I, Ruys J, Heller R, Lerba C (1990) The Dubbo study: an Australian prospective community study of the health of elderly. *Aust N Z J Med* 20:783–789
15. Nguyen TV, Center JR, Sambrook PN, Eisman JA (2001) Risk factors for proximal humerus, forearm, and wrist fractures in elderly men and women: the Dubbo Osteoporosis Epidemiology Study. *Am J Epidemiol* 153:587–595
16. National Osteoporosis Society (1998) Guidance on the prevention and management of corticosteroid induced osteoporosis. National Osteoporosis Society, Camerton, Bath, UK
17. Angus RM, Sambrook PN, Pocock NA, Eisman JA (1989) A simple method for assessing calcium intake in Caucasian women. *J Am Diet Assoc* 89:209–214
18. Nguyen TV, Sambrook PN, Eisman JA (1997) Sources of variability in bone mineral density measurements: implications for study design and analysis of bone loss. *J Bone Miner Res* 12:124–135
19. Henry MJ, Pasco JA, Pocock NA, Nicholson GC, Kotowicz MA (2004) Reference ranges for bone densitometers adopted Australia-wide: Geelong osteoporosis study. *Australas Radiol* 48:473–475
20. Pocock NA, Sambrook PN, Nguyen T, Kelly P, Freund J, Eisman JA (1992) Assessment of spinal and femoral bone density by dual X-ray absorptiometry: comparison of lunar and hologic instruments. *J Bone Miner Res* 7:1081–1084
21. Hoeting JA (1999) Bayesian model averaging: a tutorial. *Stat Sci* 14:382–147
22. Raftery AE, Madigan D, Hoeting JA (1997) Bayesian model averaging dor linear regression models. *J Am Stat A* 92:179–191
23. Volinsky CT, Madigan D, Raftery AE, Kronmal RA (1997) Bayesian model averaging in proportional hazard models: assessing the risk of a stroke. *Appl Statist* 48:433–448
24. Raftery AE (1995) Bayesian model selection in social research. In: Marsden PV (ed) *Sociological methodology*. Blackwell, Cambridge, MA
25. Swets JA (1986) Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance. *Psychol Bull* 99:181–198
26. Swets JA (1986) Indices of discrimination or diagnostic accuracy: their ROCs and implied models. *Psychol Bull* 99:100–117
27. Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
28. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143:29–36
29. Cox DR (1972) Regression models and life tables. *J R Stat Soc (B)* 34:187–220
30. Cox DR, Snell EJ (1989) *Analysis of binary data*. Chapman and Hall, London, U.K
31. Harrell FE Jr, Lee KL, Mark DB (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 15:361–387
32. R Development Core Team (2007) R: a language and environment for statistical computing. In: *R Foundation for Statistical Computing*, Vienna, Austria
33. Harrell FE Jr, Margolis PA, Gove S, Mason KE, Mulholland EK, Lehmann D, Muhe L, Gatchalian S, Eichenwald HF (1998) Development of a clinical prediction model for an ordinal outcome: the World Health Organization Multicentre Study of Clinical Signs and Etiological agents of Pneumonia, Sepsis and Meningitis in Young Infants. WHO/ARI young infant multicentre study group. *Stat Med* 17:909–944
34. Ensrud KE, Lipschutz RC, Cauley JA, Seeley D, Nevitt MC, Scott J, Orwoll ES, Genant HK, Cummings SR (1997) Body size and hip fracture risk in older women: a prospective study. Study of osteoporotic fractures research group. *Am J Med* 103:274–280
35. Henry MJ, Pasco JA, Sanders KM, Nicholson GC, Kotowicz MA (2006) Fracture Risk (FRISK) Score: Geelong osteoporosis study. *Radiology* 241:190–196
36. Cummings SR (2006) A 55-year-old woman with osteopenia. *JAMA* 296:2601–2610
37. Sambrook PN, Eisman JA (2000) Osteoporosis prevention and treatment. *Med J Aust* 172:226–229
38. Dargent-Molina P, Douchin MN, Cormier C, Meunier PJ, Breart G (2002) Use of clinical risk factors in elderly women with low bone mineral density to identify women at higher risk of hip fracture: The EPIDOS prospective study. *Osteoporos Int* 13:593–599
39. Cranney A, Guyatt G, Griffith L, Wells G, Tugwell P, Rosen C (2002) Meta-analyses of therapies for postmenopausal osteoporosis. IX: summary of meta-analyses of therapies for postmenopausal osteoporosis. *Endocr Rev* 23:570–578
40. Nguyen ND, Eisman JA, Nguyen TV (2006) Anti-hip fracture efficacy of bisphosphonates: a bayesian analysis of clinical trials. *J Bone Miner Res* 21:340–349
41. Vestergaard P, Jorgensen NR, Mosekilde L, Schwarz P (2007) Effects of parathyroid hormone alone or in combination with antiresorptive therapy on bone mineral density and fracture risk—a meta-analysis. *Osteoporos Int* 18:45–57
42. Kanis JA, Borgstrom F, De Laet C, Johansson H, Johnell O, Jonsson B, Oden A, Zethraeus N, Pflieger B, Khaltayev N (2005) Assessment of fracture risk. *Osteoporos Int* 16:581–589
43. Kanis JA, Johansson H, Oden A, Johnell O, De Laet C, Eisman JA, McCloskey EV, Mellstrom D, Melton LJ 3rd, Pols HA, Reeve J, Silman AJ, Tenenhouse A (2004) A family history of fracture and fracture risk: a meta-analysis. *Bone* 35:1029–1037
44. Kanis JA, Johansson H, Oden A, Johnell O, de Laet C, Melton IL, Tenenhouse A, Reeve J, Silman AJ, Pols HA, Eisman JA, McCloskey EV, Mellstrom D (2004) A meta-analysis of prior corticosteroid use and fracture risk. *J Bone Miner Res* 19:893–899
45. Pongchaiyakul C, Nguyen ND, Jones G, Center JR, Eisman JA, Nguyen TV (2005) Asymptomatic vertebral deformity as a major risk factor for subsequent fractures and mortality: a long-term prospective study. *J Bone Miner Res* 20:1349–1355
46. Harrell FE (2001) *Regression modeling strategies with applications to linear models, logistic regression, and survival analysis*. Springer, New York, NY
47. Royston P, Altman DG, Sauerbrei W (2006) Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 25:127–141
48. Bianco FJ Jr (2006) Nomograms and medicine. *Eur Urol* 50:884–886
49. Kattan MW (2003) Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: preoperative application in prostate cancer. *Curr Opin Urol* 13:111–116

50. Wong SL, Kattan MW, McMasters KM, Coit DG (2005) A nomogram that predicts the presence of sentinel node metastasis in melanoma with better discrimination than the American Joint Committee on Cancer staging system. *Ann Surg Oncol* 12:282–288
51. Ross PL, Gerigk C, Gonen M, Yossepowitch O, Cagiannos I, Sogani PC, Scardino PT, Kattan MW (2002) Comparisons of nomograms and urologists' predictions in prostate cancer. *Semin Urol Oncol* 20:82–88
52. Hosmer DW, Lemeshow S (1999) Assessing the fit of the model. In *Applied logistic regression*. Wiley J, New York, pp 135–175
53. Kattan M (2002) Statistical prediction models, artificial neural networks, and the sophism "I am a patient, not a statistic". *J Clin Oncol* 20:885–887
54. Ensrud KE, Black DM, Palermo L, Bauer DC, Barrett-Connor E, Quandt SA, Thompson DE, Karf DB (1997) Treatment with alendronate prevents fractures in women at highest risk: results from the Fracture Intervention Trial. *Arch Intern Med* 157:2617–2624
55. Borgstrom F, Johnell O, Kanis JA, Jonsson B, Rehnberg C (2006) At what hip fracture risk is it cost-effective to treat? International intervention thresholds for the treatment of osteoporosis. *Osteoporos Int* 17:1459–1471