

## Development and validation of osteoporosis risk-assessment model for Korean postmenopausal women

Sun Min Oh · Byung-Ho Nam · Yumie Rhee ·  
Seong-Hwan Moon · Deog Young Kim ·  
Dae Ryong Kang · Hyeon Chang Kim

Received: 20 September 2012 / Accepted: 3 January 2013 / Published online: 19 February 2013  
© The Japanese Society for Bone and Mineral Research and Springer Japan 2013

**Abstract** Currently, dual-energy X-ray absorptiometry (DXA) is the gold standard for detecting osteoporosis, but is not recommended for general population screening. Therefore, this study aims to develop an osteoporosis risk-assessment model to identify high-risk individuals among Korean postmenopausal women. Data from 1,209 and 1,046 postmenopausal women who participated in the 2009 and 2010 Korean National Health and Nutrition Examination Survey, respectively, were used for development and validation of an osteoporosis risk-assessment model. Osteoporosis was defined as *T* score less than or equal to  $-2.5$  at either the femoral neck or lumbar spine. Performance of the candidate models and the Osteoporosis Self assessment Tool for Asians (OSTA) were compared with respect to sensitivity, specificity, and area under the receiver operating characteristics curve (AUC). To compare the developed Korean Osteoporosis Risk-Assessment

Model (KORAM) with OSTA, a net reclassification improvement was further calculated. In the development dataset, the prevalence of osteoporosis was 33.9 %. KORAM, consisting of age, weight, and hormone therapy, had a sensitivity of 91.2 %, a specificity of 50.6 %, and an AUC of 0.709 with a specific cut-off score of  $-9$ . Comparable results were shown in the validation dataset: sensitivity 84.8 %, specificity 51.6 %, and AUC 0.682. Additionally, risk categorization with KORAM showed improved reclassification over that of OSTA from 7.4 to 41.7 %. KORAM can be easily used as a pre-screening tool to identify candidates for DXA tests. Further studies investigating cost-effectiveness and replicability in other datasets are required to establish the clinical utility of KORAM.

**Keywords** Osteoporosis · Risk assessment · Postmenopausal women · Korea

S. M. Oh · H. C. Kim (✉)  
Department of Preventive Medicine,  
Yonsei University College of Medicine, 50 Yonsei-ro,  
Seodaemun-gu, Seoul 120-752, Republic of Korea  
e-mail: hckim@yuhs.ac

B.-H. Nam  
Biometric Research Branch, National Cancer Center,  
323 Ilsan-ro, Ilsandong-gu, Goyang,  
Gyeonggi 410-769, Republic of Korea

Y. Rhee  
Department of Internal Medicine, Endocrine Research Institute,  
and Brain Korea 21 Project for Medical Science,  
Yonsei University College of Medicine, 50 Yonsei-ro,  
Seodaemun-gu, Seoul 120-752, Republic of Korea

S.-H. Moon  
Department of Orthopedics Surgery, Yonsei University  
College of Medicine, 50 Yonsei-ro, Seodaemun-gu,  
Seoul 120-752, Republic of Korea

D. Y. Kim  
Department and Research Institute of Rehabilitation Medicine,  
Yonsei University College of Medicine, 50 Yonsei-ro,  
Seodaemun-gu, Seoul 120-752, Republic of Korea

D. R. Kang  
Biostatistics Collaboration Unit, Yonsei University College  
of Medicine, 50 Yonsei-ro, Seodaemun-gu,  
Seoul 120-752, Republic of Korea

H. C. Kim  
Severance Institute for Vascular and Metabolic Research,  
Yonsei University College of Medicine, 50 Yonsei-ro,  
Seodaemun-gu, Seoul 120-752, Republic of Korea

H. C. Kim  
Department of Preventive Medicine,  
Northwestern University Feinberg School of Medicine,  
420 East Superior Street,  
Chicago, IL 60611, USA

## Introduction

Osteoporosis is a disease characterized by low bone mass and structural deterioration of bone tissue, leading to bone fragility and an increased susceptibility to fractures [1]. As osteoporosis and related fractures occur primarily because of aging, they are a growing health care burden in the aging population [2]. However, only 40 % of women with osteoporosis in Europe [2] and 32 % in the USA [3] reported being treated. In Korea, also, osteoporosis is an important public health problem. According to the 2009 Korean Health Statistics, among participating women aged 50 years or older, the prevalence of osteoporosis was 38.7 % [4]. According to the National Health Insurance database, which includes almost all prescription and treatment claims with diagnostic codes in Korea, the number of patients being treated for osteoporosis substantially increased by 26.7 % between 2005 and 2008 [5]. However, osteoporosis is currently under-diagnosed and under-treated. Among women with osteoporosis, 30.3 % were diagnosed with osteoporosis by a physician, and only 14.3 % reported being treated [4]. Accordingly, it is critical that further effort is given to find more effective methods for the prevention and early detection of osteoporosis in both men and women.

At present, dual-energy X-ray absorptiometry (DXA) is the most widely accepted method of diagnosing osteoporosis and of monitoring changes in bone density [6]. However, due to its relatively high cost, DXA is not recommended as a routine screening test for the general population [7]. Most guidelines limit the use of DXA for postmenopausal women aged 65 years or older, with the exception of younger adults with known risk factors [8]. Therefore, several osteoporosis risk-assessment models have been developed for pre-screening using simple variables. However, most tools have been developed and validated for the Western population [9–16]. For Asian women, the Osteoporosis Screening Tool for Asians (OSTA) was developed in community-dwelling participants who visited 21 clinics in eight Asian countries, and was validated in a Japanese population [17]. OSTA is based only on age and weight and had a sensitivity of 91 % and a specificity of 45 % in the original development dataset. OSTA has been replicated in various ethnic populations including Korean postmenopausal women [18], Filipino women [19], and even in Caucasian [20] and African–American women [21]. However, the subjects used in the Korean OSTA studies were limited to patients in a few clinics. Meanwhile, the Korea National Health and Nutrition Examination Survey (KNHANES) included DXA tests from the second half of 2008, which is the first nationwide bone mineral density (BMD) dataset for the Korean population. Therefore, this study aims to develop

and validate an osteoporosis risk-assessment model to identify high-risk Korean women who may benefit from further evaluation of osteoporosis based on a nationwide dataset.

## Materials and methods

### Participants

The KNHANES, conducted by the Korea Centers for Disease Control and Prevention and the Ministry of Health and Welfare, is a nationwide survey to assess the health and nutritional status of a non-institutionalized representative sample of the Korean population. Participants from residential districts throughout Korea were selected using a stratified, multi-stage, clustered probability sampling design; for the 2009 survey (KNHANES IV-3), household units were selected using the 2005 census in Korea [4]. For the 2010 survey (KNHANES V-1), sampling was either based on the registered market value of apartment building complexes or a registered database of the Korean government system that includes all registered citizens [22]. For development of the osteoporosis risk-assessment model, 1,917 postmenopausal women aged 50 years or older who participated in the 2009 KNHANES were included. Of them, 708 participants were excluded from the present analysis due to at least one of the following reasons: absence of BMD measurement ( $n = 149$ ), previously diagnosed osteoporosis or treatment for osteoporosis ( $n = 473$ ), missing blood tests ( $n = 199$ ), and being in a bed-ridden state ( $n = 36$ ). Finally, 1,209 women were eligible for this study. For validation of the model developed, the 2010 KNAHNS dataset was used. Of 1,657 postmenopausal women aged 50 years or older, 611 participants were excluded in the same manner as the development dataset: absence of BMD measurement ( $n = 128$ ), previously diagnosed osteoporosis or treatment for osteoporosis ( $n = 359$ ), missing blood tests ( $n = 196$ ), and being in a bed-ridden state ( $n = 27$ ). Finally, data from 1,046 participants were used for validation of the model developed. All survey participants signed an informed consent form. This study was approved by the Institutional Review Board of Korea Centers for Diseases Control and Prevention (2009-01CON-03-2C, 2010-02CON-21-C) and Yonsei University Health System (4-2011-0222) and was monitored by the Human Research Protection Center of Severance Hospital, Yonsei University Health System.

### Measurements

The survey consisted of a health interview survey, a health behavior survey, a nutrition survey, and a health

examination survey. Household interviews and self-reported questionnaires were used to assess health behavior, past or current history of disease, and family history. Smoking status was categorized as either current smoker or non-smoker (past or never). Drinking status was classified into either current drinker or non-drinker (past or never). Regular exercise was defined as moderate-to-high intensity of physical activity at least three times per week. Trained examiners in specially equipped mobile examination centers performed anthropometrics, blood tests, and BMD measurements. Body mass index (BMI) was calculated as an individual's weight in kilograms divided by their height in meters squared. Blood samples were obtained after a minimum fasting time of 8 h and handled according to standard procedures. All samples were analyzed within 24 h after arriving at the Central Testing Institute in Seoul, Korea. Serum 25-hydroxyvitamin D [25(OH)D] levels were measured using a gamma counter (1470 Wizard, Perkin-Elmer, Finland) with a radioimmunoassay (Diasorin, Still Water, MN, USA) [23]. Low vitamin D was defined as a serum 25(OH)D level of less than 20 ng/mL [23]. Serum parathyroid hormone (PTH) levels were assessed using a chemiluminescence assay (Diasorin) for the measurement of intact PTH. Elevated PTH was operationally defined as a serum PTH level of 80 pg/mL or greater. Elevated alkaline phosphatase (ALP) was operationally defined as a serum ALP level of 300 IU/L or greater. BMD was measured using a QDR Discovery fan beam densitometer (Hologic Inc., Bedford, MA, USA) at total femur, femoral neck, and L1–L4 spine. The results of DXA measurements were analyzed at the Korean Society of Osteoporosis using Hologic Discovery software (version 13.1) [24]. The stability of the DXA measurements was maintained by daily calibration [25]. *T* scores were calculated using sex-specific normal values for young Japanese women; the reference means (standard deviations) of BMD at the femoral neck and lumbar were 0.803 (0.107) and 1.006 (0.115), respectively. Osteoporosis and low BMD were defined as a *T* score less than or equal to  $-2.5$  and  $-2.0$ , respectively, at either the femoral neck or lumbar spine.

### Statistical analyses

Potential risk and protective factors for osteoporosis were selected based on previous studies and statistical investigation of the development dataset. Age [9–17], body weight [9–13, 15–17], BMI [14], hormone replacement therapy (HRT) [10–16], previous (low impact) fracture [10, 12–15], smoking [15], and history of rheumatoid arthritis [15] were evaluated because they were components of previously developed osteoporosis risk-assessment models [9–17]. Additionally, all components of the WHO Fracture Risk Assessment Tool (FRAX<sup>®</sup>) [26] except corticosteroid

use, secondary osteoporosis, and parent fractured hip were evaluated. There were no relevant data available to evaluate these variables in the KNHANES. In addition, regular exercise [27], history of type 2 diabetes [28], depression [29], and biomarkers [30, 31] including 25(OH)D, PTH, and ALP were evaluated. Simple linear regression analyses were performed to detect variables that achieved borderline statistical significance ( $P < 0.15$ ). Among the potential risk factors identified, covariates for the multiple linear regression model were selected by tenfold cross-validation. In detail, the development dataset was randomly divided into 10 subsamples. Nine subsamples were used to select significant covariates using stepwise addition and deletion ( $P < 0.15$ ), and the remaining one subsample was used for validation. This process was repeated 10 times with different subsamples to determine the optimal number of covariates [32]. Based on the results of the multiple linear regression analysis with the selected covariates, variables that did not reach a statistical significance ( $P \geq 0.05$ ) were excluded. Multicollinearity among the investigated variables was assessed by computing a variance inflation factor. Subsequently, a final multiple linear regression model with the selected covariates was computed. The regression coefficient of each covariate was used to calculate its index weight. To standardize the effect of each variable, a ratio using a coefficient for each covariate divided by the reference value, the absolute value of the coefficient for age (per 10 years), was calculated. Each standardized coefficient was then multiplied by an integer that was able to discriminate the effect of each variable and the final value was rounded off as an integer.

To develop a simple and effective model, three candidate models were tested; Model 1 included age and weight, Model 2 added health behavior, and Model 3 added blood test(s). Even though blood tests are invasive, costly and complicated to apply to a general population, Model 3 was tested to confirm whether blood tests add to the prediction of osteoporosis. In addition, these three models were compared to OSTA, which was a model available in the Korean population. Correlations of the scores from the three candidate models and OSTA with actual BMD *T* scores (lower values at either the femoral neck or lumbar spine) were evaluated using Spearman's correlation analyses. The goodness of fit of each model was assessed using the Hosmer–Lemeshow test [33]. The ability of each model to discriminate those with osteoporosis from those without osteoporosis was compared using area under the receiver operating characteristics curves (AUC) with sensitivity on the *y*-axis and  $(1 - \text{specificity})$  on the *x*-axis for all possible cut-off values. Next, a cut-off score was chosen which yielded 90 % sensitivity or greater for detecting those with osteoporosis in each model [10, 11, 17, 20]. The cut-off score was applied for each model and the final model was

selected based on sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, and negative likelihood ratio with their exact binomial confidence intervals. Additionally, false negatives (i.e., the number of missed cases that represents the number of undetected osteoporotic subjects per 1,000 subjects), false positives (i.e., the number of unnecessary DXA tests that represents the number of subjects without osteoporosis referred for DXA testing per 1,000 subjects), and AUC were compared. The model that showed the best performance using the fewest variables was selected and named the Korean Osteoporosis Risk-Assessment Model (KORAM) for postmenopausal women. KORAM was then validated using an independent dataset, KNAHNS V-1. Sensitivity analyses were performed with an outcome of low BMD *T* score of  $-2.0$  or lower at either the femoral neck or lumbar spine [11].

According to the KORAM scores, three risk categories were created operationally: low, intermediate, and high risk of having osteoporosis. We defined low risk as having a less than 10 % probability of osteoporosis, high risk as having more than a 70 % probability, and intermediate risk as in between these values. The net reclassification improvement (NRI) [34] was then calculated to evaluate whether the KORAM risk categories provided a benefit in discriminating participants with osteoporosis from those without osteoporosis over the risk categories of OSTA. NRI was calculated by constructing  $3 \times 3$  tables according to the risk categories of KORAM and OSTA, separately in participants with or without osteoporosis. Any upward movement in categories for participants with osteoporosis implied improved reclassification, and any downward movement indicated poor reclassification. The interpretation was opposite for participants without osteoporosis [34].

$$\text{NRI} = [P(\text{up}|D = 1) - P(\text{down}|D = 1)] \\ - [P(\text{up}|D = 0) - P(\text{down}|D = 0)]$$

where *D* represents osteoporosis, 1 for osteoporosis, and 0 for normal

A *P* value less than 5 % was considered significant. All statistical analyses were performed using the SAS software package (version 9.2.1; SAS Institute, Cary, NC, USA).

## Results

### Characteristics of study participants

Baseline characteristics of Korean postmenopausal women in the development dataset (2009 KNHANES) and the validation dataset (2010 KNHANES) are summarized in Table 1. A total of 1,209 and 1,046 postmenopausal

women in the development dataset and the validation dataset, respectively, were eligible for the present study. Selected potential risk factors of osteoporosis were old age, low body weight, short height, current smoking, diabetes, low serum 25(OH)D, elevated ALP, and elevated PTH. Meanwhile, protective factors were current drinking, regular exercise, and currently taking hormones. The mean age was 63.5 years in the development dataset and 62.3 years in the validation dataset. However, the percentage of participants who were 75 years or older was higher in the development dataset (13.8 %) than in the validation dataset (8.9 %). The prevalence of osteoporosis was 33.9 % in the development dataset and 29.6 % in the validation dataset (Table 1).

### Development and validation of KORAM

According to tenfold cross-validation, seven variables associated with BMD *T* scores were selected: age, weight, HRT, regular exercise, current smoking status, current drinking status, and serum ALP levels. After adjustment for covariates, age, weight, HRT, and serum ALP levels had significant and independent associations with BMD *T* scores. No significant multicollinearity was observed among those variables. Regression coefficient, standard error, and index weight of each variable in the final multiple regression model are shown in Table 2. Based on the selected variables and their index weights, three candidate models were developed as follows:

$$\text{Model 1} = [(\text{age in years}/10) \times (-4) \\ + (\text{weight in kilograms}/10) \times 3]$$

$$\text{Model 2} = [(\text{age in years}/10) \times (-4) \\ + (\text{weight in kilograms}/10) \times 3 \\ + |(\text{if no HRT})| \times (-3)]$$

$$\text{Model 3} = [(\text{age in years}/10) \times (-4) \\ + (\text{weight in kilograms}/10) \times 3 + (\text{if no HRT}) \\ \times (-3) + (\text{if elevated ALP}) \times (-2)]$$

The ranges of scores in Model 1, Model 2, and Model 3 in postmenopausal women were from  $-23$  to  $9$  (median  $-8$ ), from  $-26$  to  $9$  (median  $-10$ ), and from  $-27$  to  $9$  (median  $-11$ ), respectively. In the development dataset, all models for postmenopausal women showed reasonable fitness according to the Hosmer–Lemeshow test (*P* for OSTA = 0.541, *P* for Model 1 = 0.773, *P* for Model 2 = 0.791, and *P* for Model 3 = 0.504). When we analyzed correlations between the predicted scores and actual BMD *T* scores, Spearman's correlation coefficients were 0.635 for OSTA, 0.650 for Model 1, 0.660 for Model 2, and 0.665 for Model 3 (*P* for all  $<0.001$ ). In terms of

**Table 1** Baseline characteristics of participants in the development and the validation dataset

Variables	Development dataset ( <i>N</i> = 1,209)	Validation dataset ( <i>N</i> = 1,046)
Age (years)	63.5 ± 8.9	62.3 ± 8.2
50–54 years	239 (19.8)	216 (20.7)
55–59 years	215 (17.8)	240 (22.9)
60–64 years	239 (19.8)	201 (19.2)
65–69 years	193 (16.0)	164 (15.7)
70–74 years	156 (12.9)	132 (12.6)
≥75 years	167 (13.8)	93 (8.9)
Weight (kg)	57.1 ± 8.8	57.5 ± 8.6
≥60 kg	413 (34.2)	363 (34.7)
55–59 kg	296 (24.5)	258 (24.7)
50–54 kg	258 (21.3)	235 (22.5)
<50 kg	242 (20.0)	190 (18.2)
Height (cm)	153.2 ± 5.9	153.6 ± 5.5
Body mass index (kg/m <sup>2</sup> )	24.3 ± 3.2	24.3 ± 3.3
Current smoking	56 (4.6)	38 (3.6)
Current drinking	548 (45.3)	509 (48.7)
Regular exercise (≥3 times/week)	402 (33.3)	288 (27.5)
Hormone therapy	167 (13.8)	189 (18.1)
History of rheumatoid arthritis	58 (4.8)	43 (4.1)
Diabetes	194 (16.1)	143 (13.7)
Depression	41 (3.4)	36 (3.4)
Serum 25-OH vitamin D (ng/mL)	17.5 [13.2, 22.2]	16.8 [13.2, 21.7]
Serum alkaline phosphatase (IU/L)	251 [211, 300]	249 [210, 302]
Serum parathyroid hormone (pg/mL)	64.8 [50.9, 82.2]	63.7 [52.1, 80.5]
BMD at femoral neck (g/cm <sup>2</sup> )	0.63 ± 0.11	0.64 ± 0.11
<i>T</i> score at femoral neck	−1.64 ± 1.06	−1.53 ± 0.98
<i>T</i> score < −1.0	901 (74.5)	763 (72.9)
<i>T</i> score ≤ −2.0	435 (36.0)	341 (32.6)
<i>T</i> score ≤ −2.5	250 (20.7)	155 (14.8)
BMD at lumbar spine (g/cm <sup>2</sup> )	0.81 ± 0.14	0.82 ± 0.14
<i>T</i> score at lumbar spine	−1.70 ± 1.25	−1.63 ± 1.21
<i>T</i> score < −1.0	889 (73.5)	741 (70.8)
<i>T</i> score ≤ −2.0	514 (42.5)	418 (40.0)
<i>T</i> score ≤ −2.5	319 (26.4)	252 (24.1)
Any of those		
Lower BMD at any of those (g/cm <sup>2</sup> ) <sup>a</sup>	0.63 ± 0.11	0.64 ± 0.10
Lower <i>T</i> score at any of those <sup>a</sup>	−2.04 ± 1.07	−1.96 ± 1.03
<i>T</i> score < −1.0	1,014 (83.9)	872 (83.4)
<i>T</i> score ≤ −2.0	633 (52.4)	514 (49.1)

**Table 1** continued

Variables	Development dataset ( <i>N</i> = 1,209)	Validation dataset ( <i>N</i> = 1,046)
<i>T</i> score ≤ −2.5	410 (33.9)	310 (29.6)

Data are expressed as mean ± standard deviation or number (%) or median [interquartile range]

BMD bone mineral density

<sup>a</sup> Lower value at either the femoral neck or lumbar spine *T* score

discriminative performance, Model 2 (AUC = 0.835) was superior to Model 1 (AUC = 0.828, *P* = 0.017) or OSTA (AUC = 0.816, *P* < 0.001), but not inferior to Model 3 (AUC = 0.836, *P* = 0.638).

We then compared the performance of OSTA and the candidate models with the selected cut-off scores to yield 90 % or greater sensitivity based on sensitivity, specificity, PPV, NPV, false negative, false positive, positive likelihood ratio, negative likelihood ratio, and AUC (Table 3). In the case of OSTA, the predefined cut-off score of −1 showed a relatively low sensitivity (79.5 %). Therefore, a cut-off score of 0 to yield 90 % or greater sensitivity was used as a reference. Compared with the other models, Model 2 with a cut-off score of −9 showed improved specificity, PPV, false negative, false positive, positive likelihood ratio, and negative likelihood ratio. Additionally, the AUC of Model 2 was significantly higher than those of OSTA, with a cut-off score of 0, and Model 1, with a cut-off score of −6, but comparable to that of Model 3, with a cut-off score of −9. Similar findings were shown in the validation dataset and in sensitivity analyses of both the development and validation datasets. Therefore, Model 2 with a cut-off score of −9 was finally selected and named the Korean Osteoporosis Risk-Assessment Model (KORAM) for postmenopausal women.

#### Performance of KORAM by risk category

To define the clinical implications of KORAM, three risk categories were created with cut-off values of −9 and −15: greater than −9 for low risk, between −15 and −9 for intermediate risk, and less than −15 for high risk. In the development dataset, 36.4, 43.1, and 20.5 % were classified into low-, intermediate-, and high-risk categories, respectively. Among women in the low-, intermediate-, and high-risk categories, 8.2, 35.7, and 75.8 %, respectively, had osteoporosis. In the validation dataset, the percentage of women in the low-, intermediate-, and high-risk categories was 40.8, 45.4, and 13.8 %, respectively. The prevalence of osteoporosis was 11.0, 34.1, and 70.1 % in the low-, intermediate-, and high-risk categories, respectively (Table 4).

The net improvement of KORAM compared to OSTA with the suggested cut-off scores (−4 and −1) was



**Table 2** Regression coefficients and index weights in the final multiple regression model

Variables	Regression coefficient	Standard error	P value	Standardized coefficient	Index weight
Intercept	(−) 0.852	0.262	0.001		–
Age (10 years)	(−) 0.516	0.028	<0.001	(−) 1.0	(−) 4
Weight (10 kg)	0.430	0.027	<0.001	0.8	3
No HRT	(−) 0.358	0.069	<0.001	(−) 0.7	(−) 3
Elevated ALP <sup>a</sup>	(−) 0.234	0.054	<0.001	(−) 0.5	(−) 2

HRT hormone replacement therapy

<sup>a</sup> Elevated ALP: serum alkaline phosphatase  $\geq 300$  IU/L

28.5 % (117 out of 410) for participants with osteoporosis, −21.2 % (169 out of 799) for those without osteoporosis, and 7.4 % [95 % confidence interval (CI) 1.1–13.6] overall. Further, the NRIs of KORAM compared to OSTA showed improved classification from 7.4 to 41.7 % across the different cut-off scores of OSTA: −4 and −1, −3 and 0, −4 and 0, and −5 and 0 (Table 5).

#### Comparison of KORAM with the current Korean NHIC guidelines

Currently, Korean National Health Insurance Corporation (NHIC) guidelines for osteoporosis screening using DXA tests are limited to women aged 65 years or older, with the following exceptions: younger women with low body weight (BMI < 18.5), early menopause (<40 years), surgical menopause, and a past history or family history of non-traumatic fracture. KORAM was developed using each individual's age, weight, and HRT status; thus, it may provide an efficient targeting guideline for the use of DXA tests. According to current NHIC guidelines, 64.1 % of women in the development dataset would be candidates for the measurement of BMD by DXA. With NHIC guidelines, 85.1 % of osteoporosis cases could be detected. The number of missed cases (false negatives) was estimated to be 149 per 1,000 subjects, and the number of unnecessary DXA tests (false positives) was estimated to be 534 per 1,000 subjects. With KORAM, 91.2 % of patients with osteoporosis could be detected. Moreover, the number of missed cases and unnecessary DXA tests per 1,000 subjects would be reduced to 88 and 495 cases, respectively (Table 6).

#### Discussion

In this study, the Korean osteoporosis risk-assessment model for postmenopausal women was developed and validated in a Korean population based on a nationally representative BMD and health examination dataset:  $\text{KORAM} = [(\text{age in years}/10) \times (-4) + (\text{weight in kilograms}/10) \times 3 + (\text{if no HRT}) \times (-3)]$ .

To develop KORAM, we investigated clinically or statistically significant factors that were associated with BMD: age, body weight, height, current smoking status, currently drinking status, regular exercise, HRT, diabetes, depression, low serum 25(OH)D, elevated ALP, and elevated PTH. Of these factors, four variables in women including age, weight, HRT, and ALP levels were selected as potential components of the risk-assessment models. To form a simple and effective model, we first defined a baseline model for age and weight, which were core variables in the previously developed models [10–13, 17]. We then assessed the incremental effect of adding HRT status to predict osteoporosis. Currently, HRT is a protective factor for osteoporosis and has been applied in previous models including simple calculated osteoporosis risk estimation (SCORE) [10], osteoporosis risk assessment instrument (ORAI) [11], osteoporosis index of risk (OSIRIS) [12], osteoporosis risk assessment by composite linear estimate (ORACLE) [14], and age, body size, no estrogen (ABONE) [16]. Additionally, serum ALP levels were negatively correlated with BMD in this study. Increased levels of bone-specific serum ALP, a biomarker of bone turnover, have been reported to be associated with an increased risk of osteoporotic fracture in community-dwelling postmenopausal women of Japanese ancestry [35]. Thus, we evaluated whether an invasive laboratory test such as serum ALP levels further improved the prediction of osteoporosis, even though serum ALP is less specific than bone-specific ALP [36]. When comparing the sensitivity, specificity, PPV, NPV, false positive, false negative, positive likelihood ratio, negative likelihood ratio, and AUC of each model, the model with age, weight, and HRT showed better performance than the baseline model. However, the addition of elevated ALP levels did not improve performance. Finally, KORAM was developed based on simple variables including age, weight, and HRT. Therefore, KORAM can be easily used in a primary care setting for pre-screening to decide whether to use DXA testing as well as in the general population for self-screening purposes.

In the present study, we evaluated both linear and logistic regression models to calculate an index weight of each variable to create a scoring system to predict

**Table 3** Discriminatory performance of OSTA and the candidate models to identify postmenopausal women with osteoporosis and low BMD

	Sensitivity % (95 % CI)	Specificity % (95 % CI)	PPV % (95 % CI)	NPV % (95 % CI)	NMC	NUDT	LR (+) (95 % CI)	LR (-) (95 % CI)	ROC	
									AUC (SE)	Statistical difference
<b>Osteoporosis (<math>T_{FN} \leq -2.5</math> or <math>T_{LS} \leq -2.5</math>)</b>										
<b>Development dataset</b>										
OSTA ( $\leq -1$ )	79.5 (75.3, 83.3)	66.8 (63.5, 70.1)	55.2 (51.1, 59.2)	86.4 (83.5, 89.0)	205	332	2.40 (2.15, 2.68)	0.31 (0.25, 0.37)		
OSTA ( $\leq 0$ )	96.8 (94.6, 98.3)	28.3 (25.2, 31.6)	40.9 (37.8, 44.1)	94.6 (90.9, 97.1)	32	718	1.35 (1.29, 1.42)	0.11 (0.06, 0.19)	0.626 (0.009)	(ref)
Model 1 ( $\leq -6$ )	92.7 (89.7, 95.0)	44.1 (40.6, 47.6)	46.0 (42.5, 49.4)	92.2 (89.0, 94.6)	74	560	1.66 (1.55, 1.77)	0.17 (0.12, 0.24)	0.684 (0.011)	<0.001 (Ref)
Model 2 ( $\leq -9$ )	91.2 (88.1, 93.8)	50.6 (47.0, 54.1)	48.6 (45.1, 52.2)	91.8 (88.9, 94.2)	88	495	1.85 (1.71, 1.99)	0.17 (0.13, 0.24)	0.709 (0.011)	<0.001 (Ref)
Model 3 ( $\leq -9$ )	93.4 (90.6, 95.6)	46.4 (42.9, 50.0)	47.2 (43.7, 50.7)	93.2 (90.3, 95.5)	66	536	1.74 (1.63, 1.87)	0.14 (0.10, 0.21)	0.699 (0.011)	<0.001 0.025 0.056
<b>Validation dataset</b>										
OSTA ( $\leq -1$ )	76.1 (71.0, 80.8)	67.1 (63.6, 70.5)	49.4 (44.8, 54.0)	87.0 (83.9, 89.6)	239	329	2.32 (2.05, 2.61)	0.36 (0.29, 0.44)		
OSTA ( $\leq 0$ )	94.2 (91.0, 96.5)	29.2 (26.0, 32.6)	35.9 (32.6, 39.3)	92.3 (88.1, 95.4)	59	708	1.33 (1.26, 1.40)	0.20 (0.13, 0.32)	0.617 (0.011)	(Ref)
Model 1 ( $\leq -6$ )	87.1 (82.9, 90.6)	42.1 (38.5, 45.8)	38.8 (35.2, 42.5)	88.6 (84.8, 91.7)	129	579	1.50 (1.40, 1.62)	0.31 (0.23, 0.41)	0.646 (0.013)	0.002 (Ref)
Model 2 ( $\leq -9$ )	84.8 (80.4, 88.6)	51.6 (48.0, 55.3)	42.5 (38.6, 46.5)	89.0 (85.6, 91.8)	152	484	1.75 (1.61, 1.92)	0.29 (0.22, 0.39)	0.682 (0.014)	<0.001 (Ref)
Model 3 ( $\leq -9$ )	88.7 (84.7, 92.0)	46.1 (42.4, 49.7)	40.9 (37.2, 44.8)	90.6 (87.2, 93.4)	113	540	1.64 (1.52, 1.78)	0.25 (0.18, 0.34)	0.674 (0.013)	<0.001 0.003 0.220
<b>Low BMD (<math>T_{FN} \leq -2.0</math> or <math>T_{LS} \leq -2.0</math>)</b>										
<b>Development dataset</b>										
OSTA ( $\leq -1$ )	71.4 (67.7, 74.9)	75.9 (72.2, 79.3)	76.5 (72.9, 79.8)	70.7 (67.0, 74.3)	286	242	2.96 (0.54, 3.45)	0.38 (0.33, 0.43)		
OSTA ( $\leq 0$ )	93.7 (91.5, 95.5)	34.6 (30.7, 38.6)	61.1 (58.0, 64.2)	83.3 (77.9, 87.8)	64	655	1.43 (1.34, 1.52)	0.18 (0.13, 0.25)	0.641 (0.011)	(Ref)
Model 1 ( $\leq -6$ )	87.7 (84.9, 90.1)	52.8 (48.6, 56.9)	67.1 (63.8, 70.3)	79.6 (75.2, 83.5)	124	473	1.86 (1.70, 2.03)	0.23 (0.19, 0.29)	0.702 (0.012)	<0.001 (Ref)
Model 2 ( $\leq -9$ )	85.2 (82.1, 87.8)	60.1 (55.9, 64.1)	70.1 (66.7, 73.3)	78.6 (74.5, 82.4)	149	340	2.13 (1.92, 2.37)	0.25 (0.20, 0.30)	0.726 (0.012)	<0.001 (Ref)
Model 3 ( $\leq -9$ )	88.0 (85.2, 90.4)	55.9 (51.7, 60.0)	68.7 (65.4, 71.9)	80.9 (76.7, 84.7)	121	441	1.20 (1.81, 2.20)	0.21 (0.17, 0.27)	0.720 (0.012)	<0.001 0.027 0.214
<b>Validation dataset</b>										
OSTA ( $\leq -1$ )	66.0 (61.7, 70.1)	73.9 (69.9, 77.6)	70.9 (66.6, 75.0)	69.2 (65.2, 73.0)	341	262	2.52 (2.16, 2.95)	0.46 (0.40, 0.52)		
OSTA ( $\leq 0$ )	90.9 (88.0, 93.2)	35.0 (30.9, 39.2)	57.4 (54.0, 60.9)	79.8 (74.1, 84.8)	92	651	1.40 (1.31, 1.50)	0.26 (0.19, 0.35)	0.629 (0.012)	(Ref)
Model 1 ( $\leq -6$ )	82.9 (79.3, 86.0)	49.3 (44.9, 53.6)	61.2 (57.5, 64.8)	74.9 (70.0, 79.3)	172	508	1.63 (1.49, 1.79)	0.35 (0.28, 0.43)	0.661 (0.014)	0.001 (Ref)
Model 2 ( $\leq -9$ )	79.2 (75.4, 82.6)	60.2 (55.9, 64.3)	65.8 (61.9, 69.5)	74.9 (70.6, 79.0)	209	399	1.99 (1.77, 2.23)	0.35 (0.29, 0.42)	0.697 (0.014)	<0.001 (Ref)
Model 3 ( $\leq -9$ )	84.1 (80.6, 87.1)	54.9 (50.6, 59.2)	64.3 (60.5, 67.9)	78.1 (73.5, 82.2)	160	452	1.86 (1.68, 2.06)	0.29 (0.23, 0.36)	0.695 (0.014)	<0.001 0.005 0.769

Model 1 included age and weight, model 2 added hormone replacement therapy, and model 3 added hormone replacement therapy and elevated alkaline phosphatase

NMC the number of missed cases which represents the number of undetected osteoporotic subjects (i.e., false negatives) per 1000 subjects; NUDT the number of unnecessary DXA tests which represents the number of subjects referred for DXA testing (i.e., false positives) per 1000 subjects

OSTA osteoporosis self-assessment tool for Asians, BMD bone mineral density, PPV positive predictive value, NPV negative predictive value, LR(+) positive likelihood ratio, LR(-) negative likelihood ratio, AUC area under the curve,  $T_{FN}$  T score at femoral neck,  $T_{LS}$  T score at lumbar spine

**Table 4** Performance of KORAM to predict osteoporosis according to its risk categories

KORAM Korean Osteoporosis Risk-Assessment Model

Risk category	Development dataset		Validation dataset	
	Total N (column %)	Osteoporosis N (row %)	Total N (column %)	Osteoporosis N (row %)
High (< -15)	248 (20.5)	188 (75.8)	144 (13.8)	101 (70.1)
Intermediate (-15 to -9)	521 (43.1)	186 (35.7)	475 (45.4)	162 (34.1)
Low (> -9)	440 (36.4)	36 (8.2)	427 (40.8)	47 (11.0)
Total	1209 (100.0)	410 (33.9)	1046 (100.0)	310 (29.6)

**Table 5** Net reclassification improvement of KORAM compared to OSTA risk categories with the different cut-off values

KORAM Korean Osteoporosis Risk-Assessment Model, OSTA osteoporosis self-assessment tool for Asians, NRI net reclassification improvement, CI confidence interval

OSTA risk category	Total N (column %)	Osteoporosis N (row %)	Participants with osteoporosis, %		Participants without osteoporosis, %		NRI (95 % CI)
			Up	Down	Up	Down	
Cut-off: -4, -1							
High (< -4)	215 (17.8)	168 (78.1)	0.0	0.2	0.0	0.0	7.4 (1.1, 13.6)
Intermediate (-4 to -1)	376 (31.1)	158 (42.0)	17.1	0.2	4.9	1.3	
Low (> -1)	618 (51.1)	84 (13.6)	12.0	0.0	17.5	0.0	
Cut-off: -3, 0							
High (< -3)	320 (26.5)	233 (72.8)	0.0	0.7	0.0	0.1	19.9 (15.1, 24.7)
Intermediate (-3 to 0)	650 (53.8)	164 (25.2)	5.6	5.6	1.8	22.3	
Low (>0)	239 (19.8)	13 (5.4)	0.0	0.0	0.0	0.0	
Cut-off: -4, 0							
High (< -4)	215 (17.8)	168 (78.1)	0.0	0.1	0.0	0.0	28.6 (22.7, 34.5)
Intermediate (-4 to 0)	755 (62.4)	229 (30.3)	17.1	5.6	4.9	22.3	
Low (>0)	239 (19.8)	13 (5.4)	0.0	0.0	0.0	0.0	
Cut-off: -5, 0							
High (< -5)	140 (11.6)	119 (85.0)	0.0	0.0	0.0	0.0	41.7 (34.8, 48.7)
Intermediate (-5 to 0)	830 (68.7)	278 (33.5)	31.7	5.6	6.6	22.3	
Low (>0)	239 (19.8)	13 (5.4)	0.0	0.0	0.0	0.0	

**Table 6** Comparison between the current Korean NHIC guidelines and KORAM

NHIC National Health Insurance Corporation, KORAM Korean Osteoporosis Risk-Assessment Model, DXA dual-energy X-ray absorptiometry, PPV positive predictive value, NPV negative predictive value

	NHIC guideline	KORAM
No. of subjects recommended for DXA testing	775/1,209 (64.1 %)	769/1,209 (63.6 %)
Sensitivity	349/410 (85.1 %)	374/410 (91.2 %)
Specificity	373/799 (46.7 %)	404/799 (50.6 %)
PPV	349/775 (45.0 %)	374/769 (48.6 %)
NPV	373/434 (85.9 %)	404/440 (91.8 %)
No. of missed cases	61/410 (149 per 1,000 subjects)	36/410 (88 per 1,000 subjects)
No. of unnecessary DXA tests	426/799 (534 per 1,000 subjects)	395/799 (495 per 1,000 subjects)

osteoporosis. However, when we compared the performance of the models, the scoring system using the weights from linear regression analyses slightly outperformed those from logistic regression analyses (data not shown).

Additionally, all candidate models were compared with OSTA, which has been validated in many countries [17–21].

Overall, KORAM showed better performance for the detection of osteoporosis than OSTA. The discriminative performance of KORAM was significantly superior to that of OSTA (AUC = 0.835 and 0.816, respectively,  $P < 0.001$ ). In addition, KORAM demonstrated improved net reclassification from 7.4 % to 41.7 % over OSTA across different



cut-off scores. Furthermore, when compared with current Korean NHIC guidelines, KORAM provided a more efficient targeting guideline for the use of DXA tests because it considers each individual's age, weight, and HRT status (Table 6). Osteoporosis itself presents no specific symptoms, but the burden of osteoporotic fracture continues to grow with age. Therefore, the potential benefits and risks of DXA testing should be investigated. However, no clinical trials are available to evaluate the effectiveness of osteoporosis screening or any potential risks that can result from screening [7]. In false-negative cases, a diagnosis of osteoporosis can be overlooked and further treatment can be delayed. Conversely, false-positive cases can lead to unnecessary DXA tests, unnecessary exposure to radiation, and increased health care costs. However, even if participants who underwent DXA testing did not have osteoporosis (false-positive cases), the result of the DXA test will still assess the status of their bone health and help to estimate the proper interval for BMD testing. According to a recent study in white postmenopausal women aged 65 years or older, the estimated BMD testing interval for 10 % of participants to develop osteoporosis was over 15 years in women with normal BMD or mild osteopenia ( $T$  scores  $\geq -1.5$ ) [37]. However, this interval would be shortened to 5 years in women with moderate osteopenia ( $-1.99 \leq T$  scores  $\leq -1.5$ ) and even to 1 year in women with severe osteopenia ( $-2.49 \leq T$  scores  $\leq -2.0$ ) [37].

This study has several limitations. First, the cost-effectiveness of KORAM was not considered in the present study. Although more than 90 % of people with osteoporosis could be detected by KORAM, approximately 50 % would undergo unnecessary DXA tests. Therefore, to use KORAM in clinical practice, its potential benefits and drawbacks should be further evaluated. Second, despite using a nationwide representative dataset, the KNHANES is a cross-sectional study. Therefore, KORAM is limited to estimating prevalent cases of osteoporosis. Considering that the ultimate goal for improving bone health is to prevent osteoporotic fracture and that BMD scores provide only a marginal benefit for predicting osteoporotic fracture [32, 38], the role of KORAM in preventing osteoporotic fracture should be further evaluated. Currently, the Korean version of FRAX<sup>®</sup> to predict the 10-year probability of hip fracture and major osteoporotic fracture (clinical spine, humerus or wrist fracture) is available. Since KORAM and FRAX share two major components, age and weight, studies on correlations of a calculated risk of osteoporosis by KORAM with an estimated risk of osteoporotic fracture by FRAX and prevalent cases of osteoporotic fracture are needed in well-designed cohorts with the Korean population. Third, to calculate sex-specific  $T$  scores at the femoral neck and lumbar spine, site and sex-specific reference means and standard deviations were adopted from the Japanese data because no Korean

reference data are currently available. Thus, KORAM should be further adjusted when Korean reference BMD data are available.

To our knowledge, this is the first development of an osteoporosis risk-assessment model for Korean postmenopausal women using a nationally representative dataset that includes BMD measurements and other relevant risk factors of osteoporosis. This study suggests that KORAM is a useful pre-screening tool for screening osteoporosis by DXA in the Korean population. Since KORAM is easy to calculate with simple variables, it can be used in either a primary care setting or in general use as a self-screening tool. However, prior to using KORAM in these settings, its cost-effectiveness, especially compared to current NHIC guidelines, should be investigated. In addition, replication studies using other Korean BMD datasets are recommended. Finally, further adjustment of KORAM using BMD data as a reference in the Korean population is necessary.

**Acknowledgments** This work is based on the doctoral dissertation of Sun Min Oh entitled “Development and validation of osteoporosis risk assessment model for Koreans (KORAM)” at Yonsei University. There was no source of funding.

**Conflict of interest** All authors have no conflicts of interest.

## References

1. World Health Organization (2003) Prevention and management of osteoporosis. WHO Press, Switzerland
2. Ström O, Borgström F, Kanis JA, Compston J, Cooper C, McCloskey EV, Jönsson B (2011) Osteoporosis: burden, health care provision and opportunities in the EU: a report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Arch Osteoporos* 6:59–155
3. Dawson-Hughes B, Looker AC, Tosteson AN, Johansson H, Kanis JA, Melton LJ 3rd (2012) The potential impact of the National Osteoporosis Foundation guidance on treatment eligibility in the USA: an update in NHANES 2005–2008. *Osteoporos Int* 23:811–820
4. Ministry of Health and Social Welfare and Korea Centers for Disease Control & Prevention (2010) Korea Health Statistics 2009: Korean National Health and Nutrition Examination Survey (KNHANES IV-3). Republic of Korea, Osong
5. Choi HJ, Shin CS, Ha YC, Jang SM, Jang SH, Park CM, Yoon HK, Lee SS (2012) Burden of osteoporosis in adults in Korea: a national health insurance database study. *J Bone Miner Metab* 30:54–58
6. Krieg MA, Barkmann R, Gonnelli S et al (2008) Quantitative ultrasound in the management of osteoporosis: the 2007 ISCD Official Positions. *J Clin Densitom* 11:163–187
7. Nelson HD, Haney EM, Dana T, Bougatsos C, Chou R (2010) Screening for osteoporosis: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 153:99–111
8. Lim LS, Hoeksema LJ, Sherin K (2009) Screening for osteoporosis in the adult U.S. population: ACPM position statement on preventive practice. *Am J Prev Med* 36:366–375

9. Michaëlsson K, Bergström R, Mallmin H, Holmberg L, Wolk A, Ljunghall S (1996) Screening for osteopenia and osteoporosis: selection by body composition. *Osteoporos Int* 6:120–126
10. Lydick E, Cook K, Turpin J, Melton M, Stine R, Byrnes C (1998) Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density. *Am J Manag Care* 4:37–48
11. Cadarette SM, Jaglal SB, Kreiger N, McIsaac WJ, Darlington GA, Tu JV (2000) Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. *CMAJ* 162:1289–1294
12. Sedrine WB, Chevallier T, Zegels B, Kvasz A, Micheletti MC, Gelas B, Reginster JY (2002) Development and assessment of the Osteoporosis Index of Risk (OSIRIS) to facilitate selection of women for bone densitometry. *Gynecol Endocrinol* 16:245–250
13. Salaffi F, Silveri F, Stancati A, Grassi W (2005) Development and validation of the osteoporosis prescreening risk assessment (OPERA) tool to facilitate identification of women likely to have low bone density. *Clin Rheumatol* 24:203–211
14. Richey F, Deceulaer F, Ethgen O, Bruyère O, Reginster JY (2004) Development and validation of the ORACLE score to predict risk of osteoporosis. *Mayo Clin Proc* 79:1402–1408
15. Black DM, Palermo L, Pearson J, Abbott T, Johnell O (1998) SOFSURF: A simple, useful risk factor system can identify the large majority of women with osteoporosis. *Bone* 23:S605 [abstract]
16. Weinstein L, Ullery B (2000) Identification of at-risk women for osteoporosis screening. *Am J Obstet Gynecol* 183:547–549
17. Koh LK, Sedrine WB, Torralba TP, Kung A, Fujiwara S, Chan SP, Huang QR, Rajatanavin R, Tsai KS, Park HM, Reginster JY; Osteoporosis Self-Assessment Tool for Asians (OSTA) Research Group (2001) A simple tool to identify Asian women at increased risk of osteoporosis. *Osteoporos Int* 12:699–705
18. Park HM, Ben Sedrine W, Reginster JY, Ross PD; OSTA (2003) Korean experience with OSTA risk index for osteoporosis: a validation study. *J Clin Densitom* 6:247–250
19. Li-Yu JT, Llamado LJ, Torralba TP (2005) Validation of OSTA among Filipinos. *Osteoporos Int* 16:1789–1793
20. Rud B, Hilden J, Hylidstrup L, Hróbjartsson A (2009) The Osteoporosis Self-Assessment Tool versus alternative tests for selecting postmenopausal women for bone mineral density assessment: a comparative systematic review of accuracy. *Osteoporos Int* 20:599–607
21. Wallace LS, Ballard JE, Holiday D, Turner LW, Keenum AJ, Pearman CM (2004) Evaluation of decision rules for identifying low bone density in postmenopausal African-American women. *J Natl Med Assoc* 96:290–296
22. Ministry of Health and Social Welfare and Korea Centers for Disease Control & Prevention (2012) Korea Health Statistics 2010: Korean National Health and Nutrition Examination Survey (KNHANES V-1). Republic of Korea, Osong
23. Choi HS, Oh HJ, Choi H, Choi WH, Kim JG, Kim KM, Kim KJ, Rhee Y, Lim SK (2011) Vitamin D insufficiency in Korea—a greater threat to younger generation: the Korea National Health and Nutrition Examination Survey (KNHANES) 2008. *J Clin Endocrinol Metab* 96:643–651
24. Hong S, Oh HJ, Choi H, Kim JG, Lim SK, Kim EK, Pyo EY, Oh K, Kim YT, Wilson K, Choi WH (2011) Characteristics of body fat, body fat percentage and other body composition for Koreans from KNHANES IV. *J Korean Med Sci* 26:1599–1605
25. Korea Centers for Disease Control & Prevention (2009) Education and quality control of bone mineral density in the Fourth National Health and Nutrition Examination Survey (KNHANES IV-3). Republic of Korea, Osong
26. Kanis JA, Hans D, Cooper C, Baim S, Bilezikian JP, et al.; Task Force of the FRAX Initiative (2011) Interpretation and use of FRAX in clinical practice. *Osteoporos Int* 22:2395–2411
27. Langsetmo L, Hitchcock CL, Kingwell EJ, Davison KS, Berger C, Forsmo S, Zhou W, Kreiger N, Prior JC; Canadian Multicentre Osteoporosis Study Research Group (2012) Physical activity, body mass index and bone mineral density—associations in a prospective population-based cohort of women and men: the Canadian Multicentre Osteoporosis Study (CaMos). *Bone* 50:401–408
28. Ma L, Oei L, Jiang L, Estrada K, Chen H, Wang Z, Yu Q, Zillikens MC, Gao X, Rivadeneira F (2012) Association between bone mineral density and type 2 diabetes mellitus: a meta-analysis of observational studies. *Eur J Epidemiol* 27:319–332
29. Wu Q, Magnus JH, Liu J, Bencaz AF, Hentz JG (2009) Depression and low bone mineral density: a meta-analysis of epidemiologic studies. *Osteoporos Int* 20:1309–1320
30. Lee J, Vasikaran S (2012) Current recommendations for laboratory testing and use of bone turnover markers in management of osteoporosis. *Ann Lab Med* 32:105–112
31. Joo NS, Dawson-Hughes B, Kim YS, Oh K, Yeum KJ (2012) Impact of calcium and vitamin D insufficiencies on serum parathyroid hormone and bone mineral density: analysis of the fourth and fifth Korea National Health and Nutrition Examination Survey (KNHANES IV-3, 2009 and V-1, 2010). *J Bone Miner Res*. doi:10.1002/jbmr.1790
32. Robbins J, Aragaki AK, Kooperberg C, Watts N, Wactawski-Wende J, Jackson RD, LeBoff MS, Lewis CE, Chen Z, Stefanick ML, Cauley J (2007) Factors associated with 5-year risk of hip fracture in postmenopausal women. *JAMA* 298:2389–2398
33. Hosmer D, Lemeshow S (2000) Applied logistic regression, 2nd edn. Wiley, New York
34. Pencina MJ, D'Agostino RB, Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics Med* 27:157–172 (discussion 207)
35. Ross PD, Kress BC, Parson RE, Wasnich RD, Armour KA, Mizrahi IA (2000) Serum bone alkaline phosphatase and calcaneus bone density predict fractures: a prospective study. *Osteoporos Int* 11:76–82
36. Takahashi M, Kushida K, Hoshino H, Miura M, Ohishi T, Inoue T (1997) Comparison of bone and total alkaline phosphatase activity on bone turnover during menopause and in patients with established osteoporosis. *Clin Endocrinol* 47:177–183
37. Gourlay ML, Preisser JS, Lui L, Cauley JA, Ensrud KE; Study of Osteoporosis Fractures Research Group (2012) BMD screening in older women: initial measurement and testing interval. *J Bone Miner Res* 27:743–746
38. Leslie WD, Morin S, Lix LM, Johansson H, Oden A, McCloskey E, Kanis JA, Manitoba Bone Density Program (2012) Fracture risk assessment without bone density measurement in routine clinical practice. *Osteoporos Int* 23:75–85