J Chin Med Assoc

ORIGINAL ARTICLE

Comparison between linear regression and four different machine learning methods in selecting risk factors for osteoporosis in a Chinese female aged cohort

Shiow-Jyu Tzou[a,b], Chung-Hsin Peng[c], Li-Ying Huang[d], Fang-Yu Chen[d], Chun-Heng Kuo[d], Chung-Ze Wu[e], Ta-Wei Chu[f,g,*]

[a]Teaching and Researching Center, Kaohsiung Armed Forces General Hospital, Kaohsiung, Taiwan, ROC; [b]Institute of Medical Science and Technology, National Sun Yat-sen University, Kaohsiung, Taiwan, ROC; [c]Department of Urology, Cardinal Tien Hospital, School of Medicine, Fu Jen Catholic University, New Taipei City, Taiwan, ROC; [d]Division of Endocrinology and Metabolism, Department of Internal Medicine, Fu Jen Catholic University Hospital, New Taipei, Taiwan; [e]Department of Internal Medicine, Shuang Ho Hospital, New Taipei City, Division of Endocrinology and Metabolism, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, ROC; [f]Department of Obstetrics and Gynecology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, ROC; [g]MJ Health Research Foundation, Taipei, Taiwan, ROC

Abstract

Background: Population aging is emerging as an increasingly acute challenge for countries around the world. One particular manifestation of this phenomenon is the impact of osteoporosis on individuals and national health systems. Previous studies of risk factors for osteoporosis were conducted using traditional statistical methods, but more recent efforts have turned to machine learning approaches. Most such efforts, however, treat the target variable (bone mineral density [BMD] or fracture rate) as a categorical one, which provides no quantitative information. The present study uses five different machine learning methods to analyze the risk factors for T-score of BMD, seeking to (1) compare the prediction accuracy between different machine learning methods and traditional multiple linear regression (MLR) and (2) rank the importance of 25 different risk factors.

Methods: The study sample includes 24 412 women older than 55 years with 25 related variables, applying traditional MLR and five different machine learning methods: classification and regression tree, Naïve Bayes, random forest, stochastic gradient boosting, and eXtreme gradient boosting. The metrics used for model performance comparisons are the symmetric mean absolute percentage error, relative absolute error, root relative squared error, and root mean squared error.

Results: Machine learning approaches outperformed MLR for all four prediction errors. The average importance ranking of each factor generated by the machine learning methods indicates that age is the most important factor determining T-score, followed by estimated glomerular filtration rate (eGFR), body mass index (BMI), uric acid (UA), and education level.

Conclusion: In a group of women older than 55 years, we demonstrated that machine learning methods provide superior performance in estimating T-Score, with age being the most important impact factor, followed by eGFR, BMI, UA, and education level.

Keywords: Chinese; Machine learning; Osteoporosis

1. INTRODUCTION

Reduced birth rates and increased life expectancy have driven population aging in the developed world, but these trends in recent years are increasingly apparent in the developing world as well, with the general exception of Africa.[1] By 2030, the World Health Organization estimates the global population older than 60 years will be 1.4 billion, a 40% increase from 2019.[2] As of 2018, 14.3% of Taiwan's population was older than 65 years.[3] Aging is related to many comorbidities such as cancer, metabolic disease, and cardiovascular disease, making it an imperative concern for governments and healthcare providers.[4] One key comorbidity is osteoporosis, a degenerative condition that particularly affects women. The World Health Organization defines osteoporosis as bone mineral density (BMD) more than 2.5 SDs below that of the mean of young adults (T-score ≤ –2.5) based on the dual-energy x-ray absorptiometry measurements.[3] Kanis[5] reported that the elderly experiences a 10-fold increase in 10-year fracture rate compared with younger individuals. Besides from the pain and suffering osteoporosis directly imposes on sufferers, it also creates a huge financial burden for governments and national health systems. Kemmak et al[6] noted that the treatment of osteoporosis-related fractures costs

* *Address correspondence. Dr. Ta-Wei Chu, Department of Obstetrics and Gynecology, Tri-Service General Hospital, National Defense Medical Center, 325, Section, 2, Chenggong Road, Taipei 114, Taiwan, ROC. E-mail address: taweichu@gmail.com (T.-W. Chu).*

Conflicts of interest: The authors declare that they have no conflicts of interest related to the subject matter or materials discussed in this article.

Journal of Chinese Medical Association. (2023) 86: 1028-1036.

Received April 17, 2023; accepted July 3, 2023.

doi: 10.1097/JCMA.0000000000000999

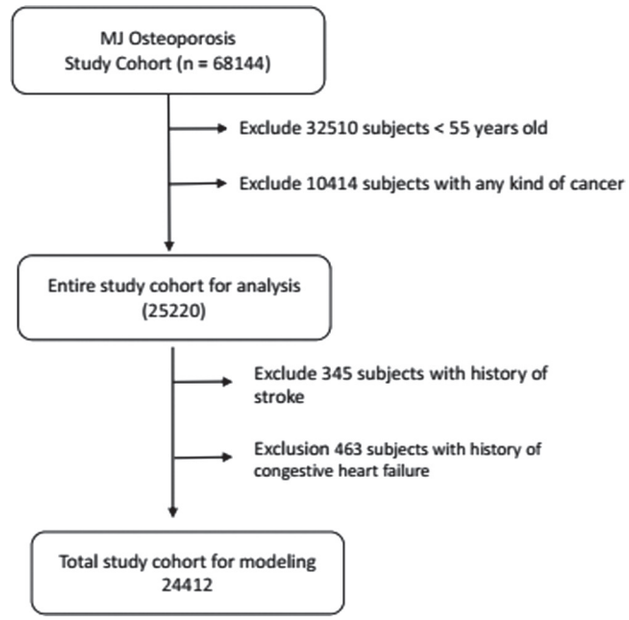

Fig. 1 Flowchart of sample selection from the MJ osteoporosis study cohort. MJ = MJ Health Database.

Western countries (Canada, Europe, and United States) an average of USD 5000 to 6500 billion annually, without accounting for costs related to subsequent disability.

Several studies have investigated risk factors for osteoporosis, with a particularly high concentration conducted in Korea,[7–10] and most of which treat osteoporosis as a binary variable[7,8,11] using logistic regressions to calculate the receiver operation curve. In such calculations, a larger area indicates a higher degree of model accuracy. However, logistic regressions are considerably less informative than multiple linear regressions (MLRs). In assessing the relationship between risk factors and BMD, greater accuracy can be obtained using the T-score of the BMD as an independent variable (y) to provide quantitative observations. Recently, machine learning methods have emerged as a new data analysis method that is competitive with MLR.[12,13] Machine learning can capture nonlinear relationships in data and complex interactions among multiple predictors, and therefore, can potentially outperform conventional MLR in disease prediction.[14]

The present study enrolled 24 412 women in Taiwan older than >55 years, collecting information on 25 BMD-related risk factors, applying traditional MLR and five machine learning methods to investigate the relationships between risk factors and T-score. Our purposes were as follows:

1. To compare the prediction accuracy between machine learning and traditional MLR.
2. To rank the relative importance of the 25 risk factors.

2. METHODS

2.1. Participant and study design

The data for this study were obtained from the Taiwan MJ cohort, an ongoing prospective cohort of health examinations performed by the MJ Health Screening Centers in Taiwan.[15] These health examinations include more than 100 crucial biological indicators, such as anthropometric measurements, blood tests, and imaging tests. In addition, each participant completed a self-administered questionnaire to gather information on personal and family medical history, current health status, lifestyle, physical exercise, sleep habits, and dietary habits.[16] The MJ Health Database only comprises individuals who have given informed consent. All or part of the data used in this research were authorized by and received from the MJ Health Research Foundation (Authorization Code MJHRF2020018A). Any interpretations or conclusions described in this article are those of the authors alone and do not represent the views of the MJ Health Research Foundation.[17] The study protocol was approved by the Institutional Review Board of the Kaohsiung Armed Forces General Hospital (IRB No. KAFGHIRB 109-041). In total, 68 144 healthy participants were enrolled. After excluding subjects with different causes, 24 412 subjects remained for analysis, as shown in Fig. 1.

MJ senior medical staff documented each subject's medical history, including details of their current medications, and conducted a comprehensive physical examination. The waist circumference was measured at the natural waist level in a horizontal position. To calculate the body mass index (BMI), the participant's weight (in kg) was divided by the square of their height (in meters). The systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured while the subject was seated using standard mercury sphygmomanometers on their right arm.

The procedures used to collect demographic and biochemical data have been previously documented.[18] Participants had fasted for 10 hours before blood draw for biochemical analyses. Within 1 hour of collection, plasma was separated from the blood and kept at 30°C until analysis for fasting plasma glucose (FPG) and lipid profiles. FPG was measured using the glucose oxidase method (YSI 203 glucose analyzer; Yellow Springs Instruments, Yellow Springs, OH). Total cholesterol and triglyceride (TG) levels were measured using the dry multilayer analytical slide method with a Fuji Dri-Chem 3000 analyzer (Fuji Photo Film, Tokyo, Japan). The serum concentrations of high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C) were determined through enzymatic cholesterol assays, following dextran sulfate precipitation. Urine albumin to creatinine ratio (ACR) was determined using turbidimetry and a Beckman Coulter AU 5800 biochemical analyzer.

Table 1 defines the 25 baseline clinical variables used as predictor variables, with T-score used as the dependent (target) variable. The independent variables include sex, age, BMI, duration of diabetes, smoking, alcohol use, FPG, glycated hemoglobin, TG, HDL-C, LDL-C, alanine aminotransferase, creatinine, SBP, and DBP (Table 2). Pearson correlation is used to evaluate the simple correlation between T-score and all other variables. Wilcoxon sign rank test is used to compare the performance of MLR and the other five machine learning methods.

2.2. Proposed scheme

This study proposes a predictive scheme for T-score using five machine learning methods, including classification and regression tree (CART), Naïve Bayes (NB), random forest (RF), stochastic gradient boosting (SGB), and eXtreme gradient boosting (XGBoost). These methods were selected as they have been used in different healthcare applications and do not require any prior assumptions about data distribution.[19–28] To evaluate the efficacy of our proposed scheme, we used MLR as a benchmark for comparison. We also identify the importance of various risk factors for predicting T-score.

The first method, CART, is a tree-structure method,[29] comprising root nodes, branches, and leaf nodes that grow recursively based on the tree structures from the root nodes and split at each node using the Gini index to produce branches and leaf nodes. The pruning node in the overgrown tree generates different decision rules to create an optimal tree size using the cost-complexity criterion, resulting in a complete tree structure.[30,31]

NB, another machine learning model used in this study, is widely used for classification tasks. This algorithm can sort objects according to specific characteristics and variables based on the Bayes theorem, estimating the values of dependent variable (y).[32]

The third method in this study is RF, an ensemble learning decision tree algorithm that combines bootstrap resampling and bagging.[33] RF works by randomly generating many different and unpruned CART decision trees, where the decrease in Gini impurity is used as the splitting criterion. All the trees generated are combined into a forest and then averaged or voted to generate output probabilities and a final model that provides a robust prediction.[34]

The fourth method is SGB, a tree-based gradient boosting learning algorithm that uses a combination of bagging and boosting techniques to address the overfitting problem of traditional decision trees.[35,36] SGB generates many stochastic weak learner trees through multiple iterations. Each tree focuses on correcting or explaining the errors of the tree generated in the previous iteration, using the residual of the previous iteration tree as input for the newly generated tree. This iterative process is repeated until the convergence condition or the maximum number of iterations is reached. Finally, the cumulative results of many trees are used to determine the final robust model.

XGBoost is an optimized extension of SGB that utilises gradient boosting technology.[37] The algorithm trains many weak models sequentially and ensembles them to achieve better prediction performance. XGBoost uses Taylor binomial expansion to approximate the objective function and generate arbitrary differentiable loss functions to accelerate model convergence.[38] It also applies a regularized boosting technique to penalize model complexity and prevent overfitting, which helps to improve model accuracy.[39]

Fig. 2 depicts the proposed scheme for prediction and variable identification, which incorporates four different machine learning methods. Initially, patient data were collected and used to prepare the dataset, which was then randomly split into a training dataset and a testing dataset on an 80/20 ratio. Hyperparameters for each machine learning method were tuned using a 10-fold cross-validation technique. The training dataset was further divided into a training dataset for model building and a validation dataset for model validation, using grid search to explore all possible hyperparameter combinations. The best model for each machine learning method was selected based on the lowest root mean square error for the validation dataset, and the variable importance ranking information was obtained for CART, NB, RF, SGB, and XGBoost.

During the testing phase, the performance of the best machine learning models was evaluated using the testing dataset. Because the target variable in this study is a numerical variable, the model performance was compared using different metrics, including symmetric mean absolute percentage error, relative absolute error, root relative squared error, and root mean squared error. The values for these metrics are listed in Table 3. The machine learning methods and MLR were compared using the Wilcoxon signed rank test because only 10 values were derived from each method so they are nonparametric variables.

To ensure a more reliable and stable comparison, the training and testing processes were repeated 10 times. The performance metrics of these five machine learning models were then averaged to compare with the performance of the benchmark MLR model. The same training and testing datasets were used for both the machine learning methods and the MLR model. A model with an average metric lower than that of the MLR model was considered a more convincing model.

Because all the machine learning methods used can rank the importance of each predictor variable, we defined the priority demonstrated in each model ranked 1 as the most critical risk factor and 25 as the last selected risk factor. The machine learning methods used in this study may produce different rankings of variable importance due to their unique modeling characteristics. To increase the stability and reliability of our findings, we integrated the variable importance rankings of the pricier machine learning models. In the final stage of our proposed

Table 1

Participant demographics

Variables	Mean ± SD	n
Age, y	62.5 ± 6.4	24,411
Body mass index, kg/m^2	24.3 ± 3.6	24,405
Leukocyte, 103/μL	5.7 ± 1.6	24,402
Hemoglobin, g/dL	13.4 ± 1.0	24,401
Fasting plasma glucose, mg/dL	109.9 ± 29.2	24,402
Serum glutamate oxaloacetic transaminase, IU/L	26.7 ± 24.6	24,392
Serum glutamate pyruvate transaminase, IU/L	27.6 ± 27.5	24,393
Estimated glomerular filtration rate, mL/min/1.73 m^2	73.3 ± 13.5	11,664
Uric acid, mg/dL	5.3 ± 1.3	24,392
Triglycerides, mg/dL	125.6 ± 81.2	24,401
High-density lipoprotein cholesterol, mg/dL	60.4 ± 15.0	24,161
Low-density lipoprotein cholesterol, mg/dL	129.0 ± 33.5	24,150
Plasma calcium concentration, mg/dL	9.4 ± 0.4	22,095
Plasma phosphate concentration, mg/dL	3.8 ± 0.4	22,092
Thyroid stimulating hormone, μIU/mL	2.0 ± 3.9	23,055
C-reactive protein, mg/dL	0.3 ± 0.6	22,817
Sport hour/week, h	3.3 ± 4.1	21,174
Systolic blood pressure, mmHg	130.7 ± 20.3	24,409
Diastolic blood pressure, mmHg	74.6 ± 11.6	24,409
T-score	−1.5 ± 1.6	24,411
	n (%)	N
Gender		24,411
Female	24,411 (100)	
Marriage status		21,889
No	6,916 (31.6)	
Yes	14,973 (68.4)	
Education		22,847
Illiterate	3,411 (14.9)	
Elementary school	9,610 (42.1)	
Junior high school (vocational)	2,862 (12.5)	
High school	3,846 (16.8)	
Junior college	1,442 (6.3)	
University	1,420 (6.2)	
Graduate school or above	256 (1.1)	
Family income		22,495
None	2,943 (13.1)	
No	5,394 (24.0)	
Below $12 493	4,843 (21.5)	
$12 805-$24 986	4,448 (19.8)	
$25 298-$37 478	2,842 (12.6)	
$37 790-$49 971	940 (4.2)	
$50 283-$62 464	481 (2.1)	
More than $62 776	604 (2.7)	
Sleeping time/day, h		22,110
0-4	910 (4.1)	
4-6	7,543 (34.1)	
6-8	12,426 (56.2)	
More than 8 h	1,231 (5.6)	
Smoking status		22,518
No	21,431 (95.17)	
Yes	1,087 (4.8)	
Drinking		20,874
No	19,909 (95.4)	
Yes	965 (4.6)	

Table 2

Simple correlations between BMD and other factors

Variables	BMI	UA	Calcium	Income	Sport	HDL-C	DBP	GPT	FPG	HB	TG
BMD	0.177***	0.095***	0.094***	0.071***	0.060***	0.059***	0.058***	0.036***	0.031***	0.027***	0.026***
Variables	TSH	Phosphate	Education	Leukocyte	SBP	Age	LDL-C	eGFR	Sleep	GOT	CRP
BMD	0.025***	0.023***	0.167***	−0.019***	−0.062***	−0.348***	0.006	0.005	−0.004	−0.005	−0.012

*$p < 0.05$,
** $p < 0.01$,
*** $p < 0.005$.

BMD = bone mineral density; BMI = body mass index; CRP = C-reactive protein; DBP = diastolic blood pressure; eGFR = estimated glomerular filtration rate; FPG = fasting plasma glucose; GOT = serum glutamate transaminase; GPT = serum glutamate pyruvate transaminase; HB = hemoglobin; HDL-C = high-density lipoprotein cholesterol; LDL-C = low-density lipoprotein cholesterol; SBP = systolic blood pressure; TG = triglyceride; TSH = thyroid stimulating hormone; UA = uric acid.

scheme, we summarize and discuss our significant findings based on the pricier machine learning models and identify the most important variables.

The study was conducted using R software, version 4.0.5, and RStudio, version 1.1.453, with the required packages installed.[40,41] The RF, SGB, CART, and XGBoost methods were, respectively, implemented using the "randomForest" R package, version 4.6-14[42]; "gbm" R package, version 2.1.8[43]; "rpart" R package, version 4.1-15[44]; and "XGBoost" R package, version 1.5.0.2.[45] The "caret" R package, version 6.0-90, was used to determine the best hyperparameters for the developed CART, RF, SGB, and XGBoost methods.[46] MLR was implemented using the "stats" R package, version 4.0.5, with the default settings.

3. RESULTS

A total of 24 412 participants were enrolled in the study, with demographic data summarized in Table 1 (mean ± SD). The results of Pearson correlation are presented in Table 2, showing that BMI, uric acid (UA), plasma calcium level, income, HDL-C, GPT, FPG, hemoglobin, TG, TSH, plasma phosphate level, and sport were positively correlated to T-score, whereas negative correlations were found for leukocyte and age.

Table 4 compares traditional MLR and the four machine learning methods in terms of T-score prediction performance. Using Wilcoxon signed rank test, all four machine learning methods significantly outperformed MLR in terms of prediction error and were all convincing machine learning models.

Tables 5 and 6 present the average importance ranking of each factor generated by the CART, SGB, NB, RF, and XGBoost methods. The different machine learning methods generated different relative importance rankings for each factor. The shade of gray indicates the importance of risk factors, with darker shades indicating a more important risk factor. For instance, in the RF method, the most important factors were baseline BMI, age, and UA. To fully integrate the importance rankings of each factor in all the five machine learning methods, the average importance ranking of each risk factor was obtained by averaging the ranking values of each variable in each method (the right-hand column). Fig. 3 showed that age was the most important factor to determine T-score, followed by estimated glomerular filtration rate (eGFR), BMI, UA, education level, and family income in Chinese women older than 55 years.

4. DISCUSSION

Although the threat posed by osteoporosis to postmenopausal women is widely recognized, the present study is the first to apply five machine learning methods to assess the relative importance of risk factors on the BMD T-score by treating the target variable as continuous, whereas previous studies had treated BMD as a categorical variable. Binary regressions only provide the sensitivity, specificity, and area under receiver operation curve, and thus lack quantitative information otherwise available through continuous methods. As demonstrated in the Results section, in order of descending importance, age, eGFR, BMI, UA, education level, and family income are the key impact factors for Chinese women older than 55 years.

The present study found age to be the most important impact factor for T-score, which corresponds with findings from Taiwan's National Health and Nutrition Examination Survey (NHANES) that 16.2% adults older than 65 years had osteoporosis, with incidence considerably higher for women (24.8% vs 5.6%). This correlation can be further extended to show a linear decrease in BMD as age increases.[47] The underlying pathophysiology for this change has been clearly explained by the role of biomarker $p16^{Ink4}$, which increases in bone-derived cells such as osteoblasts and osteocytes.[48] Accumulation of these senescent cells in bone was also confirmed from bone biopsy.[49] Most importantly and interestingly, the causative role of these cells was further supported by using $p16^{Ink4a}$ apoptosis through targeted activation of caspase mice (*p16*-INK-ATTAC) to suppress the expression of $p16^{Ink4a}$ in the senescent cells of mice, with treatment resulting in significant improvement to bone microarchitecture and strength.[50]

The second most important impact factor was eGFR, again consistent with previous findings. For example, Cai et al[51] found that the levels of bone metabolic markers and eGFR were closely correlated in stage 3 chronic renal disease, but this study was limited to only 368 subjects with stages 3 to 5 chronic renal disease. In the same time, another study performed by Choi et al[52] also had the same conclusion. They classified the eGFR into quartile. In both genders, it was observed that as eGFR increased, BMD decreased. For men, BMD values were 1.181, 1.166, 1.152, and 1.149 g/cm^2 (p = 0.001), while for women, BMD values were 0.997, 0.980, 0.979, and 0.982 g/cm^2 (p = 0.005), respectively. This study is more persuasive in that it enrolled 8992 subjects, but used analysis of variance of BMD in the four groups, and thus the methodology used is not as robust as that used in the present study.[52]

Even though the positive relationship between BMI and BMD is well accepted, it has only been documented in a handful studies in PubMed, one of which (Shayganfar et al[53]) supports our findings, separating 1054 participants into three groups (osteoporosis, low bone mass, and normal) and finding that the r values of the correlations between BMI and BMD were 1.24, 1.32, and 1.38 (p = 0.07, 0.07, and 0.19, respectively). It should be noted that these participants had a BMI between 25 and 30 kg/m^2, which is similar to the values in the present study.[53] Another study that is directly related to the present study was performed in Kosovo. Their findings also supported our finding that BMI is a significant and independent factor to increase BMD in both

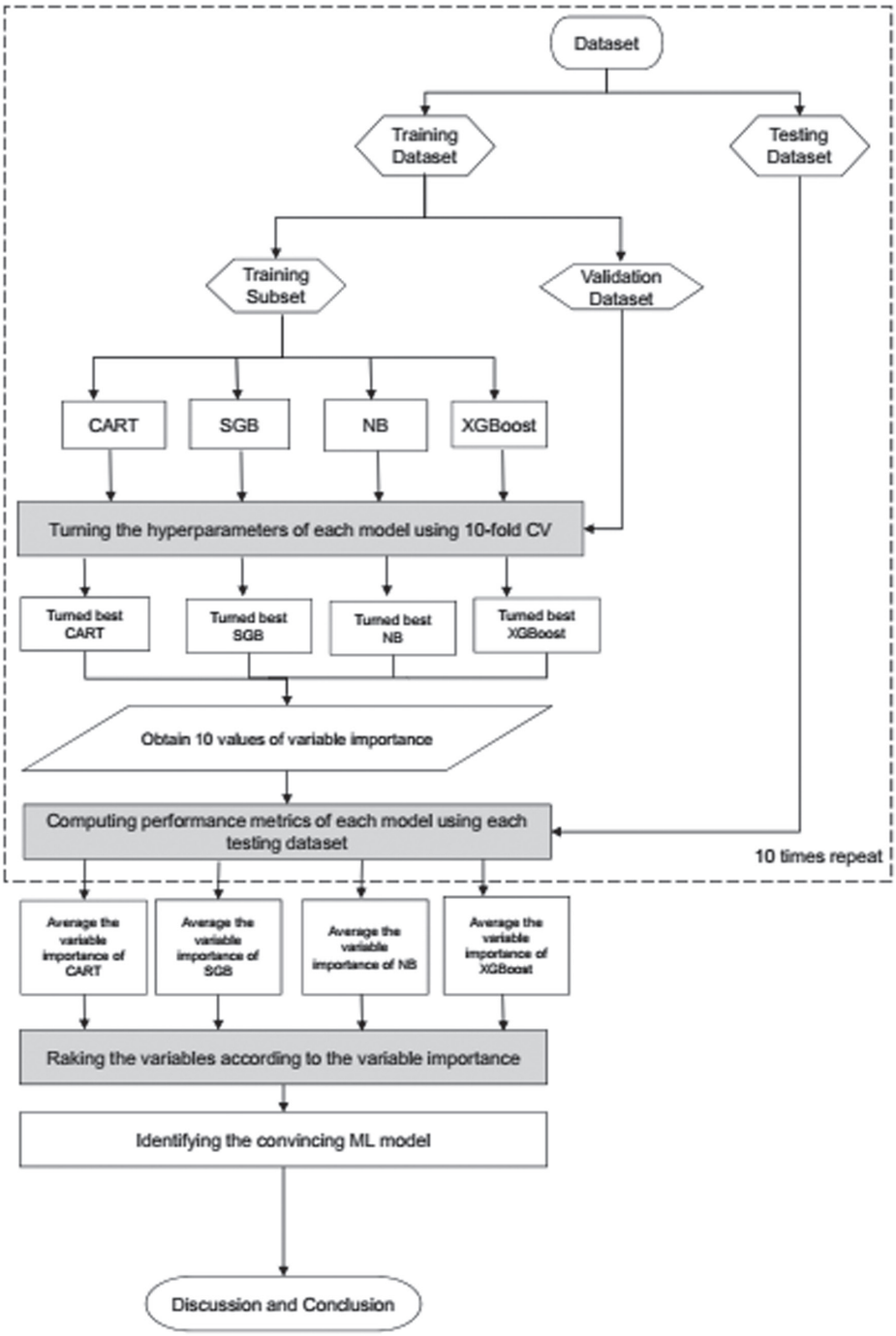

Fig. 2 Proposed Mach-L prediction scheme. CART = classification and regression tree; CV = cross-validation; ML = machine learning; NB = Naïve Bayes; SGB = stochastic gradient boosting; XGBoost = eXtreme gradient boosting.

men and women. However, their study only included 132 subjects.[54] Another study with a larger cohort of 1644 women found that BMD increases with BMI. However, again, they categorized their study participants according to BMI, and thus, less information could be obtained compared to studies using MLR.[55] At the same time, it should be noted that premenopausal subjects were also enrolled in other studies. In the present study, only women older than 55 years were enrolled. This could explain our finding while bone modeling is reactive to mechanical load, and high BMI has been shown to correlate to increased bone mineralization by improving the forces on the bones. The higher the body weight, the more the pressure is on the bone.[56]

Although UA has been shown to be related to BMD, this issue remains controversial. In the present study, UA is the fourth important factor affecting BMD. Other reports have similar findings. One study of 1080 Iranians found that BMD decreased from the lowest

Table 3

Equation of performance metrics

Metrics	Description	Calculation
SMAPE	Symmetric mean absolute percentage error	$\text{SMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\lvert y_i-\hat{y}_i\rvert}{(\lvert y_i\rvert+\lvert\hat{y}_i\rvert)/2}\times 100$
RAE	Relative absolute error	$\text{RAE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i)^2}}$
RRSE	Root relative squared error	$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}}$
RMSE	Root mean squared error	$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$

n = number of instances; yi = actual value; $\hat{y}_i$ = predicted value.

Table 4

The average performance of CART, SGB, XGBoost compared to MLR by Wilcoxon signed rank test

	SMAPE	RAE	RRSE	RMSE
Linear	1.077 ± 0.009	1.131 ± 0.007	1.138 ± 0.005	1.785 ± 0.018
CART*	1.025 ± 0.009	1.099 ± 0.004	1.11 ± 0.004	1.74 ± 0.014
NB*	1.077 ± 0.009	1.131 ± 0.007	1.138 ± 0.005	1.785 ± 0.018
RF*	1.078 ± 0.007	1.127 ± 0.007	1.135 ± 0.007	1.779 ± 0.018
SGB*	1.065 ± 0.008	1.12 ± 0.006	1.128 ± 0.006	1.769 ± 0.018
XGboost*	1.068 ± 0.009	1.121 ± 0.006	1.129 ± 0.005	1.771 ± 0.017

*$p < 0.05$ compared to linear.

CART = classification and regression tree; MAPE = mean absolute percentage error; MLR = multiple linear regression; NB = Naïve Bayes; RAE = relative absolute error; RF = random forest; RMSE = root mean squared error; RRSE = root relative squared error; SGB = stochastic gradient boosting; SMAPE = symmetric mean absolute percentage error; XGBoost = eXtreme gradient boosting.

Table 5

The results of Wilcoxon signed rank test between four machine learning methods and MLR

MLR	CART	RF	SGB	XGBoost
SMAPE	2.521 (0.01)**	–0.771 (0.44)	2.521 (0.01)**	2.521 (0.01)**
RAE	2.521 (0.01)**	2.38 (0.01)**	2.521 (0.01)**	2.521 (0.01)**
RRSE	2.521 (0.01)**	2.1 (0.03)**	2.521 (0.01)**	2.521 (0.01)**
RMSE	2.521 (0.01)**	1.96 (0.04)**	2.521 (0.01)**	2.521 (0.01)**

The results of the negative binomial (NB) model were not displayed, as the performances were presented as numeric values within parentheses, with the corresponding *p*-values.

**$p < 0.05$.

CART = classification and regression tree; MLR = multiple linear regression; RAE = relative absolute error; RF = random forest; RMSE = root mean squared error; RRSE = root relative squared error; SGB = stochastic gradient boosting; SMAPE = symmetric mean absolute percentage error; XGBoost = eXtreme gradient boosting.

to the highest quintile of UA. As previously noted, a categorical analysis provides less information than an MLR.[57] On the contrary, a Japanese study of 615 women aged from 45 to 75 years reported that lumbar spine BMD was positively related to UA after adjusting for other confounding factors (β = 0.078, p = 0.049). This inconsistency might be explained by the fact that UA is considered an antioxidant with metal-chelating properties and could scavenge superoxide. Oxidative stress is known to attenuate osteoblastogenesis and bone formation,[58,59] supporting the hypothesis that UA is beneficial to bone health. On the other hand, this could counteract antioxidant defenses, increasing the probability of osteoporosis onset.[57] However, other studies have found no direct relationship between UA and BMD. Lin et al[60] reported that UA is not associated with BMD at different skeletal sites in elderly men, and the only positive association was found in normal weight groups, implying that BMI has an impact on this relationship. Another study of 328 postmenopausal women found no significant relationship between UA and BMD.[61] From the aforementioned studies, we can conclude that, this relationship, whereas real, is weak, and further study with larger samples and longer observation periods is needed.

Due to the large study cohort size, this study was also the first to identify a correlation between educational attainment and BMD. This is possibly due to less-educated people being less likely to engage in robust self-care practices, such as regular health checkups, prevention of comorbidities, and fall prevention. This suggestion is supported by two studies using both in-person and remote learning methods to teach osteoporosis-related information, finding that the training effectively improved participant performance on the Osteoporosis Knowledge Test.[62,63] However, neither of these two studies followed up on subsequent BMD improvement. Another study of 8151 NHANES participants found that education level could have either positive or negative impact on BMD. Subjects with knowledge of osteoporosis were better able to modify their behavior to improve BMD,[64] whereas other well-informed osteoporosis patients might choose to forego treatment because of the side effects of bisphosphonates.[65] The present study provides additional evidence to support the impact of education on BMD.

The last factor noted in our study is family income, a factor largely overlooked by previous work, but supported by

Table 6

Importance ranking of each risk factor using the five convincing methods

Variables	CART	RF	SGB	NB	XGBoost	Average	Rank value
Age	1.1	2	1	3.9	1	2.3	
Estimated glomerular filtration rate	25	7.4	5.9	1	13.8	4.8	
Body mass index	3.9	1	2	14	5	5.7	
Uric acid	25	3	3.6	16.7	18.5	7.8	
Education	3.9	18.2	3.4	8.9	3	10.2	
Family income	5	20	6.5	10	4	12.2	
Fasting plasma glucose	25	8	7.5	22	8	12.5	1.0-3.5
Thyroid stimulating hormone	25	4.1	21.4	13	25	12.8	3.5-5.9
Alanine aminotransferase	25	13.5	7.9	20	10	13.8	6.0-8.4
Plasma calcium level	25	16.7	18.7	6.2	25	13.9	8.5-12.5
Sport	25	18.8	20.1	3.1	25	14	
High-density lipoprotein cholesterol	25	11.9	19.7	15	25	15.5	
Marriage	4.7	23.8	18	5	2	15.6	
Systolic blood pressure	25	9.9	19.1	18.4	25	15.8	
Triglyceride	25	5.3	21.1	22.2	20.2	16.2	
Low-density lipoprotein	25	7.8	25	16.3	25	16.4	
Drinking	25	16.2	25	8.1	25	16.4	
Plasma calcium level	25	25	22.1	2	25	16.4	
C-reactive protein	25	21	16.5	12	25	16.5	
Leukocyte	8.2	6.6	19	24.4	6	16.7	
Sleeping time	25	22	24	6.8	25	17.6	
Diastolic blood pressure	25	13.4	22	18.6	25	18	
Hemoglobin	15.8	11.1	22.6	21.8	7	18.5	
Smoking	25	23.2	25	11	25	19.7	
Aspartate aminotransferase	25	15	20	24.7	9	19.9	

CART = classification and regression tree; NB = Naïve Bayes; RF = random forest; SGB = stochastic gradient boosting; XGBoost = eXtreme gradient boosting.

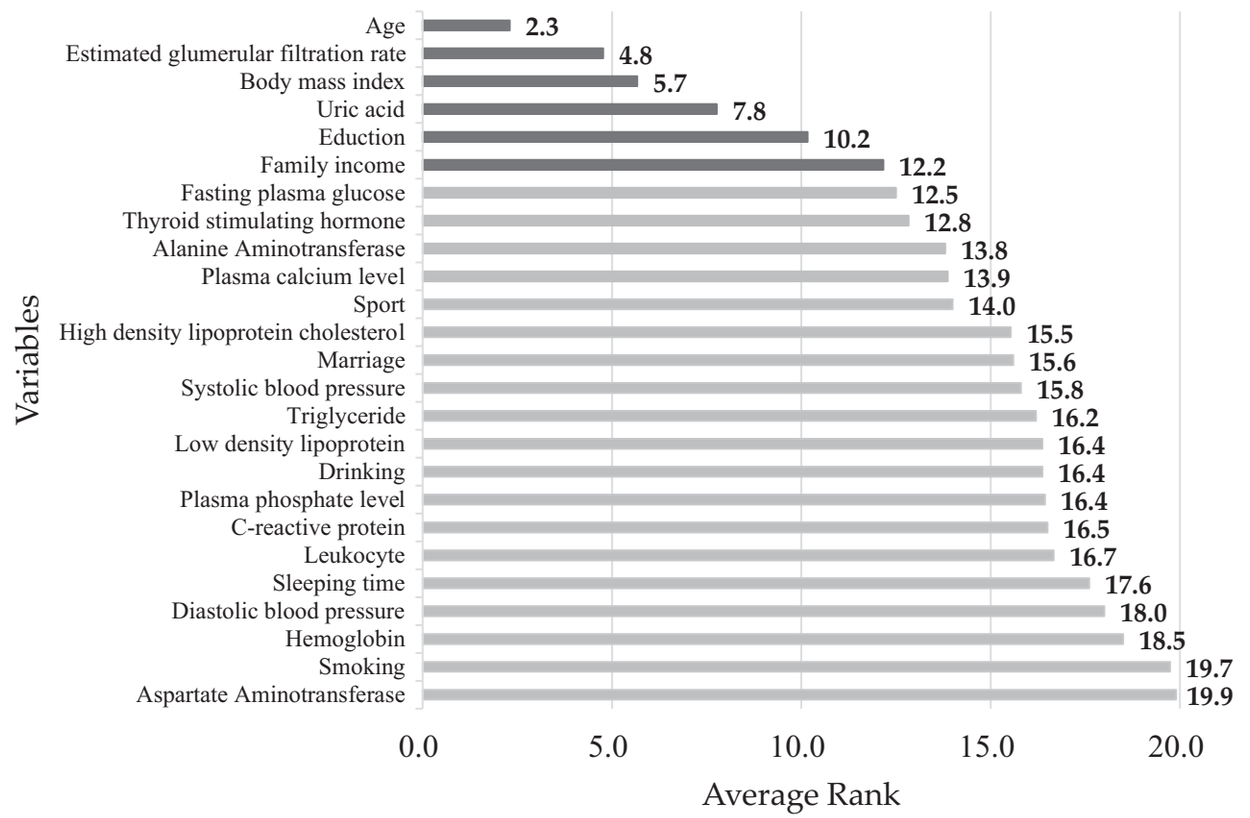

Fig. 3 Risk factors for osteoporosis in an increasing order of averaged ranking values.

several studies. For example, the NHANES found osteoporosis in 60.5% subjects with poverty income ratio of <1.3 compared to 28.1% of those with a poverty income ratio of ≥3.5, and these findings have significant credibility given the large cohort size.[66] After adjusting for confounding factors in 2403 NHANES participants, Lyles et al[67] found that low income was significantly associated with higher osteoporosis risk (odds ratio: 1.9; 95% CI, 1.07-3.37). This could be explained by limited access to general healthcare, osteoporosis treatment, healthy food, and physical activity, but this must be confirmed by further studies.

The present study still has certain limitations. First, this is a cross-sectional study and is therefore less persuasive than a longitudinal one. Second, we only measured BMD of the lumbar spine, which should only be extrapolated to hip or other parts of the body with care. Finally, this study was concentrated among a single ethnic group (Chinese), and care should be taken when generalizing the findings to other ethnic groups.

In conclusion, we find that all five machine learning methods outperformed traditional MLR, with age, eGFR, BMI, UA, education level, and family income being the most important influencers among Chinese women above older than 55 years.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the Kaohsiung Armed Forces General Hospital (KAFGH_A_109009).

REFERENCES

1. *Aging Society—The Global Trend, Its Consequences, and the Role of Technology*. Kenan Foundation Asia 2018. Available at https://www.kenan-asia.org/blog/health/aging-society-the-global-trend/. Accessed October 22, 2022.
2. *Aging*. World Health Organization 2023. Available at https://www.who.int/health-topics/ageing#tab=tab_1. Accessed January 11, 2023.
3. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J* 2017;38:500–7.
4. Yancik R, Ershler W, Satariano W, Hazzard W, Cohen HJ, Ferrucci L. Report of the National Institute on Aging task force on comorbidity. *J Gerontol A Biol Sci Med Sci* 2007;62:275–80.
5. Kanis JA. Diagnosis of osteoporosis and assessment of fracture risk. *Lancet* 2002;359:1929–36.
6. Kemmak AR, Rezapour A, Jahangiri R, Nikjoo S, Farabi H, Soleimanpour S, et al. Economic burden of osteoporosis in the world: a systematic review. *Med J Islam Repub Iran* 2020;34:154.
7. Lim HK, Ha HI, Park SY, Han J. Prediction of femoral osteoporosis using machine-learning analysis with radiomics features and abdomen-pelvic CT: a retrospective single center preliminary study. *PLoS One* 2021;16:e0247330.
8. Kwon Y, Lee J, Park JH, Kim YM, Kim SH, Won YJ, et al. Osteoporosis pre-screening using ensemble machine learning in postmenopausal Korean women. *Healthcare (Basel)* 2022;10:1107.
9. Yoo TK, Kim SK, Kim DW, Choi JY, Lee WH, Oh E, et al. Osteoporosis risk prediction for bone mineral density assessment of postmenopausal women using machine learning. *Yonsei Med J* 2013;54:1321–30.
10. Shim JG, Kim DW, Ryu KH, Cho EA, Ahn JH, Kim JI, et al. Application of machine learning approaches for osteoporosis risk prediction in postmenopausal women. *Arch Osteoporos* 2020;15:1–9.
11. Ou Yang WY, Lai CC, Tsou MT, Hwang LC. Development of machine learning models for prediction of osteoporosis from clinical health examination data. *Int J Environ Res Public Health* 2021;18:7635.
12. Marateb HR, Mansourian M, Faghihimani E, Amini M, Farina D. A hybrid intelligent system for diagnosing microalbuminuria in type 2 diabetes patients without having to measure urinary albumin. *Comput Biol Med* 2014;45:34–42.
13. Nusinovici S, Tham YC, Yan MYC, Ting DSW, Li J, Sabanayagam C, et al. Logistic regression was as good as machine learning for predicting major chronic diseases. *J Clin Epidemiol* 2020;122:56–69.
14. Miller DD, Brown EW. Artificial intelligence in medical practice: the question to the answer? *Am J Med* 2018;131:129–33.
15. Wu X, Tsai SP, Tsao CK, Chiu ML, Tsai MK, Lu PJ, et al. Cohort profile: the Taiwan MJ cohort: half a million Chinese with repeated health surveillance data. *Int J Epidemiol* 2017;46:1744–44g.
16. MJ Health Research Foundation. *The Introduction of MJ Health Data base, MJ Health Research Foundation Technical Report, MJHRF TR* 2016. Available at https://www.mjhrf.org/file/file/report/MJHRF-TR-01%E7%BE%8E%E5%85%86%E5%81%A5%E5%BA%B7%E8%B3%87%E6%96%99%E5%BA%AB%E7%B0%A1%E4%BB%8B.pdf. Accessed October 20, 2020.
17. MJ Health Research Foundation. *MJ Health Survey Database, MJ BioData [Data file], MJ BioBank [Biological Specimen]* 2014. Available at https://www.mjhrf.org/. Accessed October 20, 2020.
18. Lu CH, Pei D, Wu CZ, Kua HC, Liang YJ, Chen YL, et al. Predictors of abnormality in thallium myocardial perfusion scans for type 2 diabetes. *Heart Vessel* 2021;36:180–8.
19. Tseng CJ, Lu CJ, Chang CC, Chen GD, Cheewakriangkrai C. Integration of data mining classification techniques and ensemble learning to identify risk factors and diagnose ovarian cancer recurrence. *Artif Intell Med* 2017;78:47–54.
20. Chang CC, Chen SH. Developing a novel machine learning-based classification scheme for predicting SPCs in breast cancer survivors. *Front Genet* 2019;10:848.
21. Shih CC, Lu CJ, Chen GD, Chang CC. Risk prediction for early chronic kidney disease: results from an adult health examination program of 19,270 individuals. *Int J Environ Res Public Health* 2020;17:4973.
22. Lee TS, Chen IF, Chang TJ, Lu CJ. Forecasting weekly influenza outpatient visits using a two-dimensional hierarchical decision tree scheme. *Int J Environ Res Public Health* 2020;17:4743.
23. Chang CC, Yeh JH, Chen YM, Jhou MJ, Lu CJ. Clinical predictors of prolonged hospital stay in patients with myasthenia gravis: a study using machine learning algorithms. *J Clin Med* 2021;10:4393.
24. Chang CC, Huang TH, Shueng PW, Chen SH, Chen CC, Lu CJ, et al. Developing a stacked ensemble-based classification scheme to predict second primary cancers in head and neck cancer survivors. *Int J Environ Res Public Health* 2021;18:12499.
25. Chiu YL, Jhou MJ, Lee TS, Lu CJ, Chen MS. Health data-driven machine learning algorithms applied to risk indicators assessment for chronic kidney disease. *Risk Manag Healthc Policy* 2021;14:4401–12.
26. Wu TE, Chen HA, Jhou MJ, Chen YN, Chang TJ, Lu CJ, et al. Evaluating the effect of topical atropine use for myopia control on intraocular pressure by using machine learning. *J Clin Med* 2021;10:111.
27. Wu CW, Shen HL, Lu CJ, Chen SH, Chen HY. Comparison of different machine learning classifiers for glaucoma diagnosis based on Spectralis OCT. *Diagnostics (Basel)* 2021;11:1718.
28. Chang CC, Yeh JH, Chiu HC, Chen YM, Jhou MJ, Liu TC, et al. Utilization of decision tree algorithms for supporting the prediction of intensive care unit admission of myasthenia gravis: a machine learning-based approach. *J Pers Med* 2022;12:32.
29. Gordon AD, Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. *Biometrics* 1984;40:874.
30. Patel N, Upadhyay S. Study of various decision tree pruning methods with their empirical comparison in WEKA. *Int J Comput Appl* 2012;60:20–5.
31. Tierney NJ, Harden FA, Harden MJ, Mengersen KL. Using decision trees to understand structure in missing data. *BMJ Open* 2015;5:e007450.
32. Huang YC, Cheng YC, Jhou MJ, Chen M, Lu CJ. Important risk factors in patients with nonvalvular atrial fibrillation taking dabigatran using integrated machine learning scheme—a post hoc analysis. *J Pers Med* 2022;12:756–71.
33. Breiman L. Random forests. *Mach Learn* 2001;45:5–32.
34. Calle ML, Urrea V. Stability of random forest importance measures. *Brief Bioinform* 2011;12:86–9.
35. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001;29:1189–232.
36. Friedman JH. Stochastic gradient boosting. *Comput Stat Data Anal* 2002;38:367–78.
37. Chen T, Guestrin C. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 2016;785–94.

38. Torlay L, Perrone-Bertolotti M, Thomas E, Baciu M. Machine learning—XGBoost analysis of language networks to classify patients with epilepsy. *Brain Inform* 2017;4:159–69.
39. Iwagami M, Qizilbash N, Gregson J, Douglas I, Johnson M, Pearce N, et al. Blood cholesterol and risk of dementia in more than 1.8 million people over two decades: a retrospective cohort study. *Lancet Healthy Longev* 2021;2:e498–e506.40.
40. Tool R. *R Project* 2015. Available at http://www.r-project.org/. Accessed November 18, 2022.
41. RStudio. *Posit* n.d. Available at https://posit.co/products/open-source/rstudio/. Accessed November 18, 2022.
42. Liaw A, Wiener M. Classification and regression by randomForest. *R News* 2002;2:18–22.
43. Greenwell B, Boehmke B, Cunningham J; Developers GBM. *gbm: Generalized Boosted Regression Models, R Package Version 2.1.8* 2020. Available at https://CRAN.R-project.org/package=gbm. Accessed January 1, 2022.
44. Therneau T, Atkinson B. *Rpart: Recursive Partitioning and Regression Trees. R Package Version 4.1.15* 2022. Available at https://CRAN.R-project.org/package=rpart. Accessed January 1, 2022.
45. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. *Xgboost: Extreme Gradient Boosting. R Package Version 1.5.0.2*. Available at https://CRAN.R-project.org/package=xgboost. Accessed January 1, 2022.
46. Kuhn M. *Caret: Classification and Regression Training. R Package Version 6.0-90* 2022. Available at https://CRAN.R-project.org/package=caret. Accessed January 1, 2022.
47. Johnston CB, Dagar M. Osteoporosis in older adults. *Med Clin North Am* 2020;104:873–84.
48. Kaur J, Farr JN. Cellular senescence in age-related disorders. *Transl Res* 2020;226:96–104.
49. Farr JN, Fraser DG, Wang H, Jaehn K, Ogrodnik MB, Weivoda MM, et al. Identification of senescent cells in the bone microenvironment. *J Bone Miner Res* 2016;31:1920–9.
50. Föger-Samwald U, Kerschan-Schindl K, Butylina M, Pietschmann P. Age related osteoporosis: targeting cellular senescence. *Int J Mol Sci* 2022;23:2701.
51. Cai ZY, Hu XF, Chen LQ, Li ZR, Du XG. A cross-sectional study of osteoporosis and cardiovascular calcification in patients with chronic kidney disease at different CKD stages. *Sichuan Da Xue Xue Bao Yi Xue Ban* 2021;52:334–9.
52. Choi SW, Kim HY, Ahn HR, Lee YH, Kweon SS, Choi JS, et al. Association of bone mineral density with albuminuria and estimated glomerular filtration rate: the Dong-gu Study. *Kidney Blood Press Res* 2013;37:132–41.
53. Shayganfar A, Farrokhi M, Shayganfar S, Ebrahimian S. Associations between bone mineral density, trabecular bone score, and body mass index in postmenopausal females. *Osteoporos Sarcopenia* 2020;6:111–4.
54. Rexhepi S, Bahtiri E, Rexhepi M, Sahatciu-Meka V, Rexhepi B. Association of body weight and body mass index with bone mineral density in women and men from Kosovo. *Mater Sociomed* 2015;27:259–62.
55. Zhu K, Hunter M, James A, Lim EM, Cooke BR, Walsh JP. Discordance between fat mass index and body mass index is associated with reduced bone mineral density in women but not in men: the Busselton Healthy Ageing Study. *Osteoporos Int* 2017;28:259–68.
56. GE Healthcare. *Lunar enCORE-based x-ray bone densitometer user manual*. Madison, WI; 2019. Available at https://customer-doc.cloud.gehealthcare.com/copyDoc/LU43616v13.4EN/1. Accessed October 14, 2022.
57. Babaei M, Shamsi R, Heidari B, Bijani A, Hosseini SR. Serum uric acid status and its association with bone mineral density in the elderly people aged 60 years and more. *Int J Endocrinol Metab* 2019;17:e80780.
58. Almeida M, Han L, Martin-Millan M, O'Brien CA, Manolagas SC. Oxidative stress antagonizes Wnt signaling in osteoblast precursors by diverting β-catenin from T cell factor-to forkhead box O-mediated transcription. *J Biol Chem* 2007;282:27298–305.
59. Sahni S, Hannan MT, Gagnon D, Blumberg J, Cupples LA, Kiel DP, et al. High vitamin C intake is associated with lower 4-year bone loss in elderly men. *J Nutr* 2008;138:1931–8.
60. Lin ZC, Wu JF, Chang CY, Lai KM, Yang HY. Association between serum uric acid level and bone mineral density at multiple skeletal sites in middle-aged and elderly men: a cross-sectional study of a healthy population in Taiwan. *Arch Osteoporos* 2022;17:142.
61. Kang S, Kwon D, Lee J, Chung YJ, Kim MR, Namkung J, et al. Association between serum uric acid levels and bone mineral density in postmenopausal women: a cross-sectional and longitudinal study. *Healthcare (Basel)* 2021;9:1681.
62. Kalkım A, Dağhan S. Theory-based osteoporosis prevention education and counseling program for women: a randomized controlled trial. *Asian Nurs Res (Korean Soc Nurs Sci)* 2017;11:119–27.
63. Chotiyarnwong P, Boonnasa W, Chotiyarnwong C, Unnanuntana A. Video-based learning versus traditional lecture-based learning for osteoporosis education: a randomized controlled trial. *Aging Clin Exp Res* 2021;33:125–31.
64. Barcenilla-Wong AL, Chen JS, March LM. Concern and risk perception: effects on osteoprotective behaviour. *J Osteoporos* 2014;2014:142546.
65. Abbasi J. Amid osteoporosis treatment crisis, experts suggest addressing patients' bisphosphonate concerns. *JAMA* 2018;319:2464–6.
66. Wu Q, Xu Y, Lin G. Trends and disparities in self-reported and measured osteoporosis among US adults, 2007–2014. *J Clin Med* 2019;8:2052.
67. Lyles CR, Schafer AL, Seligman HK. Income, food insecurity, and osteoporosis among older adults in the 2007–2008 National Health and Nutrition Examination Survey (NHANES). *J Health Care Poor Underserved* 2014;25:1530–41.