scientific reports

Check for updates

OPEN

Preventive machine learning models incorporating health checkup data and hair mineral analysis for low bone mass identification

Su Jeong Kang[1], Joung Ouk (Ryan) Kim[2], Moon Jong Kim[1], Yang-Im Hur[1], Ji-Hee Haam[1], Kunhee Han[3] & Young-Sang Kim[1]✉

Machine learning (ML) models have been increasingly employed to predict osteoporosis. However, the incorporation of hair minerals into ML models remains unexplored. This study aimed to develop ML models for predicting low bone mass (LBM) using health checkup data and hair mineral analysis. A total of 1206 postmenopausal women and 820 men aged 50 years or older at a health promotion center were included in this study. LBM was defined as a T-score below − 1 at the lumbar, femur neck, or total hip area. The proportion of individuals with LBM was 59.4% (n = 1205). The features used in the models comprised 50 health checkup items and 22 hair minerals. The ML algorithms employed were Extreme Gradient Boosting (XGB), Random Forest (RF), Gradient Boosting (GB), and Adaptive Boosting (AdaBoost). The subjects were divided into training and test datasets with an 80:20 ratio. The area under the receiver operating characteristic curve (AUROC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and an F1 score were evaluated to measure the performances of the models. Through 50 repetitions, the mean (standard deviation) AUROC for LBM was 0.744 (± 0.021) for XGB, the highest among the models, followed by 0.737 (± 0.023) for AdaBoost, and 0.733 (± 0.023) for GB, and 0.732 (± 0.021) for RF. The XGB model had an accuracy of 68.7%, sensitivity of 80.7%, specificity of 51.1%, PPV of 70.9%, NPV of 64.3%, and an F1 score of 0.754. However, these performance metrics did not demonstrate notable differences among the models. The XGB model identified sulfur, sodium, mercury, copper, magnesium, arsenic, and phosphate as crucial hair mineral features. The study findings emphasize the significance of employing ML algorithms for predicting LBM. Integrating health checkup data and hair mineral analysis into these models may provide valuable insights into identifying individuals at risk of LBM.

Keywords Machine learning, Extreme gradient boosting, Hair mineral analysis, Bone mineral density

Osteoporosis is characterized by a systemic impairment of bone mass and microarchitecture, increasing the risk of fractures[1,2]. Bone mineral density (BMD) measurement by dual-energy X-ray absorptiometry (DXA) is the preferred method for diagnosing osteoporosis[3]. Decreased BMD is strongly associated with an increased risk of fractures[4,5]. Prior to performing BMD measurements via DXA, various tools reliant on clinical risk factors are accessible for prescreening individuals with low BMD[6]. The Osteoporosis Self-assessment Tool for Asians, (OSTA) score, predicted solely upon age and body weight, is the simplest tool[7]. The osteoporosis risk assessment instrument[8], simple calculated osteoporosis risk estimation[9], and osteoporosis index of risk[10] are more complex decision tools that incorporate additional risk factors, including rheumatoid arthritis, estrogen therapy, and previous low-energy fractures, in addition to age and weight. Similar efforts are needed to construct models that use clinical risk factors to predict low BMD.

[1]Department of Family Medicine, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam 13496, Republic of Korea. [2]Department of AI and Big Data, Swiss School of Management, 6500 Bellinzona, Switzerland. [3]Department of Family Medicine, Seoul Medical Center, Seoul 02053, Republic of Korea. ✉email: zeroup@cha.ac.kr

Recently, there has been significant attention directed towards the utilization of using machine learning (ML) algorithms to enhance the precision of risk assessment for osteoporosis and osteopenia. ML is one of the artificial intelligence technologies that develop algorithms that recognize patterns in data to make predictions and decisions[11]. It can be applied to clinical datasets to develop robust risk models, often outperforming traditional tools[12–14]. Consequently, investigations are underway to identify new additional risk factors for osteoporosis and osteopenia with the aid of ML algorithms.

Various minerals play critical roles in maintaining bone health. Minerals can be quantified by analysis of biological samples such as blood, urine, hair, and other tissues[15]. Hair mineral analysis is a test that measures the mineral and metal levels in a person's hair. Hair samples can be collected easily in non-invasive manner, and hair mineral analysis provides information on mineral levels over a long period. This test is based on the principle that the concentrations of these substances in the hair are related to their concentrations in the rest of the body and can be used to assess a person's overall mineral status and detect imbalances or deficiencies[16]. Various factors, such as genetic, endocrine functional, exercise-related, and nutritional factors, can influence bone metabolism. Some studies have suggested that hair mineral analysis is useful for assessing the risk of osteoporosis or osteopenia[17–19].

To date, no one has trained ML algorithms using hair minerals, and it is uncertain whether they are feasible and effective for predicting the bone disease. This study aimed to develop predictive models using ML algorithms for identifying low bone mass (LBM), referring to both osteoporosis and osteopenia, among postmenopausal women and men aged 50 years or older using data from health checkups including hair mineral analysis to contribute to the early detection and treatment of osteoporosis and osteopenia.

Materials and methods

Study design and participants

This was a single-center, retrospective cross-sectional study including community-dwelling postmenopausal women and men aged 50 years and older who participated in health checkups between 2008 and 2022 at a health promotion center, Bundang CHA Medical Center, CHA University in South Korea. The exclusion criteria included participants with missing values in their blood samples and those who had not completed the questionnaire. A total of 2061 participants were enrolled in this study. After applying the exclusion criteria, 2026 subjects were included in the analysis. Hair mineral analysis, an option within our medical center's health screening program, which enables the general public to assess their mineral levels and exposure to heavy metals. Figure 1 shows the flowchart of participant selection. This study was approved by the Institutional Review Board (IRB protocol no. 2023-08-038), and informed consent was waived because of the retrospective nature of the study.

Biochemical and anthropometric measurements

A baseline physical examination was conducted, which included measurements of vital signs and assessments of body height, weight, and waist circumference. Blood samples were taken after an overnight 8 h of fasting. Blood tests included complete blood count, biochemical tests, metabolic components, endocrine hormones, tumor markers, and bone turnover markers, which are frequently performed during health checkups. Detailed features are presented in Table 1. BMD was measured using DXA (Hologic QDR-4500, Bedford, MA, USA) at the lumbar spine (L1-4), unilateral femoral neck, and total femur. The T-score represented the standard deviation of BMD from healthy young adults of the same sex and ethnicity. The T-scores at each site were obtained, and the lowest T-score was used to interpret the results. Osteoporosis was defined as T-score ≤ -2.5, osteopenia as $-2.5 <$ T-score < -1, and normal as T-score ≥ -1[20]. The osteopenia and osteoporosis groups were classified as the LBM group. The normal and LBM groups were used to train the binary classification model.

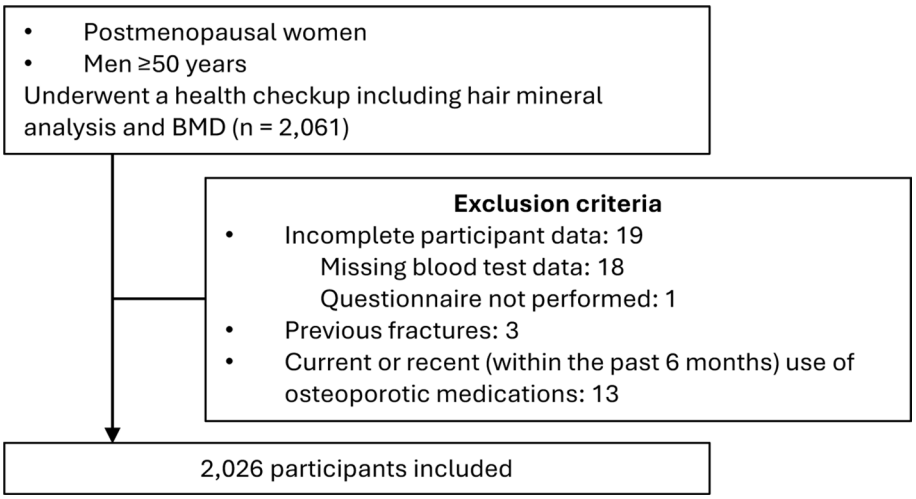

Figure 1. Study participants and the machine learning models. *BMD* bone mineral density.

Category	Features
Demographics	Age and gender
Anthropometry	Height, weight, waist circumference, BMI, SBP, and DBP
Lifestyle habits*	Physical activity, smoking, and alcohol consumption
Medical history*	History of DM, dyslipidemia, and hypertension
Blood tests	Complete blood count (WBC, seg, lym, hemoglobin, MCV, and platelets), biochemistry test (Potassium, sodium, Ca, P, Cl, uric acid, BUN, creatinine, AST, ALT, GGT, ALP, T.bil, protein, and albumin), metabolic components (Fasting plasma glucose, triglyceride, total cholesterol, HDL-C, LDL-C, vitamin B12, and folate), endocrine hormones (TSH, IGF1, and DHEA-S), tumor markers (AFP, CEA, and CA19-9), and bone turnover markers (Osteocalcin and CTX)
Hair mineral analysis	Calcium, magnesium, sodium, potassium, copper, zinc, phosphate, iron, manganese, chromium, selenium, boron, cobalt, molybdenum, sulfur, uranium, arsenic, beryllium, mercury, cadmium, lead, and aluminum

Table 1. Input features used for machine learning (ML) algorithms. *Binary classification. *BMI* body mass index, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *DM* diabetes mellitus, *WBC* white blood cell, *seg* differential count of segment neutrophil, *lym* differential count of lymphocyte, *MCV* mean corpuscular volume, *Ca* calcium, *P* phosphorus, *Cl* chloride, *BUN* blood urea nitrogen, *AST* aspartate aminotransferase, *ALT* alanine aminotransferase, *GGT* gamma-glutamyl transferase, *ALP* alkaline phosphatase, *T.bil* total bilirubin, *HDL-C* high-density lipoprotein cholesterol, *LDL-C* low-density lipoprotein cholesterol, *TSH* thyroid stimulating hormone, *IGF-1* insulin-like growth factor 1, *DHEA-S* dehydroepiandrosterone sulfate, *AFP* alpha-fetoprotein, *CEA* carcinoembryonic antigen, *CA 19–9* carbohydrate antigen, *CTX* C-terminal telopeptide of type 1 collagen.

Demographic and lifestyle parameters

All participants were requested to complete the questionnaire on the day of their health checkup. The questionnaire encompassed various aspects of the participants' sociodemographic characteristics, past medical history, smoking status, alcohol consumption status, and physical activity levels. Regarding alcohol consumption, individuals consuming more than 14 standard drinks were categorized as the excessive alcohol consumption group[21]. Regardless of the duration, smoking status was classified as current smoker, ex-smoker, or non-smoker. Regular physical activity was defined as moderate-intensity physical activity of at least 150–300 min/week[22]. Participants were classified as having hypertension, diabetes mellitus, and dyslipidemia if they were under medication.

Hair mineral analysis

Hair samples were obtained using stainless steel sampling scissors at four different points of the occipital scalp. All participants were asked not to use a chemical process on their hair for at least 8 weeks before sample collection. The hair sample was placed directly into a clean specimen envelope and sent to USA Trace Elements Inc. (TEI, Dallas, TX, USA). Hair mineral analysis revealed the concentrations of nutritional, additional, and toxic elements. A total of 22 elements were used for the analysis (Supplementary Table S1). All mineral levels were reported in milligrams percent. One milligram percent (mg%) equals ten parts per million (ppm).

Machine learning model development and input features

In this study, the four ensemble ML algorithms were used to analyze the data: Random Forest (RF), Extreme Gradient Boosting (XGB), Gradient Boosting (GB), and Adaptive Boosting (AdaBoost)[23–25]. All models were performed using Scikit-learn in Python 3.11 (Python Software Foundation, Wilmington, DE, USA). The ML models were trained with a prediction of LBM, which included the osteopenia and osteoporosis groups. The prediction target of each model was a binary variable, where "1" represented the LBM groups, and "0" represented the normal group.

The health checkup results mentioned above were divided into five categories: demographics, anthropometric measurements, lifestyle, medical history, and blood tests. The checkup results and hair mineral analysis were utilized as features in machine learning algorithms. The Features were randomly divided into training and testing datasets with an 80:20 ratio: 1620 in the training dataset and 406 in the test dataset. The feature weights were not applied to the models.

The average area under the receiver operating characteristic curve (AUROC) was processed through 50 repetitions of fivefold cross-validation to train and validate the four ML algorithms. The receiver operating characteristic (ROC) curve is a plot of true positive rate (sensitivity) on the *y* axis against false positive rate (1-specificity) on the *x* axis. The AUROC is a summary measure that essentially averages diagnostic accuracy across the spectrum of test values. AUROC equals 0.5 when the ROC curve corresponds to random chance and 1.0 for perfect accuracy[26]. From the confusion matrix, accuracy, sensitivity, specificity, precision (positive predictive value), negative predictive value, and an F1 score were calculated and summarized as the performance metrics. A grid search was performed to optimize the hyperparameters. (Supplementary Table S2). In the data preprocessing steps, missing values were excluded (Fig. 1), categorical variables were encoded using one-hot encoding to binarization, and numeric variables were standardized with StandardScaler in Python to maintain their original ranges without distortion[27].

The Shapley additive explanation (SHAP) values were used to compare the effects of the features to detect the important parameters. The SHAP values, rooted in cooperative game theory for fair profit allocation, are adapted to explain individual feature contributions in ML models[28,29]. They can be visualized using summary

plots, which display the average magnitude and direction of each feature's impact on predictions and provide a useful overview of the model.

Statistical analysis

Continuous variables were presented as means ± standard deviation and medians with interquartile range, and categorical variables were presented as numbers (percentages). Comparisons of variables between the two groups were performed using the chi-square test or Fisher's exact test for categorical variables and the independent *t*-test for continuous variables. Hair mineral concentrations were compared between the two groups using the Mann–Whitney U test. All statistical analyses were performed using the SPSS statistical package, version 27.0 (IBM corporation, Armonk, NY, USA), and p-values < 0.05 were considered statistically significant.

Informed consent

Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Institutional review board statement

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of CHA Bundang Medical Center (IRB protocol no. 2023-08-038).

Results

General characteristics

A total of 2,026 participants were included in this study. Among them, 820 (40.5%) were men. The average age of the participants was 59.15 ± 6.8 years. Based on the DXA results, 1205 (59.4%) subjects had LBM and were included in the LBM group. Participants in the decreased bone density group were of advanced age, lower body weight, shorter body height, smaller waist circumference, and lower body mass index (BMI). Other general characteristics are shown in Table 2. Only the characteristics that showed statistically significant differences were presented in the Table 2. The full comparisons were described in the Supplementary Table S1.

Model performance

ROC curves of the models were drawn and compared in every learning process; Fig. 2 shows a sample ROC curve. Out of four ML models (RF, XGB, GB, and AdaBoost), the performance of XGB had the highest average AUROC of 0.744. RF, GB, and AdaBoost had average AUROC of 0.732, 0.733, and 0.737, respectively. The maximum AUROC in 50 repetitions were 0.776, 0.802, 0.798, and 0.804 for RF, XGB, GB, and AdaBoost This represented one result out of 50 iterations. Supplementary Figure illustrated the comparison of AUROC of the four models.

The predictive metrics results of each model were summarized in Table 3. Among the four machine learning models, the F1 scores of XGB, GB, and AdaBoost, which is utilizing boosting methods, were higher compared to RF employing bagging technique. XGB and AdaBoost showed highest scrores for accuracy. When considering that the F1 score, which combines precision and sensitivity, provides a more balanced assessment of a model's performance[27], XGB achieved excellent results not only in higher average AUROC but also in other performance metrics.

Important features

Table 4 lists the top 20 important features of the XGB model evaluated by SHAP values. Body weight was the most important predictor of LBM, followed by age, alkaline phosphatase (ALP), bone turnover marker osteocalcin, body height, and dehydroepiandrosterone sulfate (DHEA-S). Among the hair minerals, sulfur was in the top 20 list of important features.

Important hair minerals

Supplementary Table S3 shows the top 10 hair minerals, their corresponding SHAP values, and the ranking of XGB. The XGB model identified sulfur, sodium, mercury, copper, magnesium, arsenic, and phosphorus, zinc, copper, and calcium as important hair minerals for predicting LBM. The SHAP values for the top 7 significant features and the top 10 important hair minerals including sulfur are shown in Fig. 3.

Discussion

In this study, ML models that predict LBM in community-dwelling adults were investigated using health checkup data. All ML models showed an average AUROC exceeding 0.73. While the XGB model showed the highest average AUROC, other performance metrics did not demonstrate notable differences among the four ensemble ML models. Body weight was the most important feature in all models. Age, height, osteocalcin, and ALP were common important features. The XGB model utilized hair minerals with higher feature importance.

Previous studies have predicted osteoporosis using clinical data with ML algorithms Most Korean studies analyzed the Korea National Health and Nutrition Examination Survey[14,30–32]. The studies were focused on postmenopausal women, showing performance ranging from AUROC 0.6–0.92. A Taiwanese study also developed an ML model to predict osteoporosis using clinical database of approximately 6000 community-dwelling participants[33] and showed similar performance to the Korean studies. A Japanese study developed an ML model that predicts osteoporosis using only age, BMI, and blood tests[34]. Like our study, several bone turnover markers were included as input features. Remarkably, performance measured using AUROC widely varied according to ML models ranged between 0.67 and 0.96. Our study adopted LBM as the prediction target using binary

Characteristics ($n = 2,026$)	Normal ($n = 821$)	LBM ($n = 1,205$)	p-value
Demographics			
Age (years)	58.1 (± 6.7)	59.9 (± 6.8)	< 0.001
Men (%)	462 (56.3%)	358 (29.7%)	< 0.001
Anthropometry			
Body height (cm)	163.5 (± 7.5)	158.7 (± 7.3)	< 0.001
Body weight (kg)	66.7 (± 9.9)	59.7 (± 8.9)	< 0.001
Waist circumference (cm)	85.9 (± 8.9)	82.3 (± 8.7)	< 0.001
BMI (kg/m^2)	24.9 (± 3.0)	23.7 (± 2.9)	< 0.001
Lifestyle habits			
Alcohol consumption (%)	366 (44.6%)	355 (29.5%)	< 0.001
Smoking (%)	219 (26.7%)	195 (16.2%)	< 0.001
Blood tests			
Complete blood count			
Hb (g/dL)	14.6 (± 1.4)	13.9 (± 1.3)	< 0.001
Platelets (10^3/uL)	230.8 (± 54.2)	240.7 (± 55.6)	< 0.001
Biochemistry			
Uric acid (mg/dL)	5.2 (± 1.3)	4.7 (± 1.2)	< 0.001
Creatinine (mg/dL)	0.98 (± 0.3)	0.90 (± 2.0)	< 0.001
ALT (IU/L)	30.3 (± 18.2)	27.7 (± 19.2)	0.002
ALP (IU/L)	165.6 (± 63.0)	185.7 (± 71.7)	< 0.001
Total bilirubin (mg/dL)	0.64 (± 0.36)	0.59 (± 0.36)	< 0.001
Total protein (g/dL)	7.2 (± 0.46)	7.3 (± 0.43)	0.017
Metabolic components			
Fasting glucose (mg/dL)	103.92 (± 23.1)	101.8 (± 21.1)	0.033
Triglyceride (mg/dL)	123.9 (± 76.6)	115.41 (± 68.0)	0.010
Total cholesterol (mg/dL)	201.8 (± 40.0)	208.1 (± 39.1)	< 0.001
HDL-C (mg/dL)	54.1 (± 14.2)	56.7 (± 14.6)	< 0.001
LDL-C (mg/dL)	123.4 (± 35.3)	127.7 (± 34.6)	0.006
Endocrine markers			
IGF-1 (ng/ml)	158.9 (± 50.0)	148.0 (± 47.9)	< 0.001
DHEA-S (μg/dL)	117.1 (± 67.6)	91.0 (± 62.3)	< 0.001
Bone turnover markers			
Osteocalcin (ng/mL)	16.49 (± 6.4)	19.9 (± 8.9)	< 0.001
CTX (ng/mL)	0.47 (± 0.24)	0.57 (± 0.32)	< 0.001
Hair mineral analysis			
Calcium (mg%)	86 (55, 144.5)	101 (57, 174.5)	< 0.001
Magnesium (mg%)	6.2 (3.7, 10.75)	7.1 (4.0, 12.05)	< 0.001
Iron (mg%)	0.7 (0.5, 0.9)	0.7 (0.6, 1.0)	< 0.001
Manganese (mg%)	0.018 (0.011, 0.032)	0.020 (0.012, 0.038)	0.001
Selenium (mg%)	0.06 (0.04, 0.07)	0.05 (0.04, 0.07)	< 0.001
Cobalt (mg%)	0.001 (0.001, 0.002)	0.001 (0.001, 0.002)	0.002
Sulfur (mg%)	4143.0 (3936.5, 4380.0)	4221.0 (4002.0, 4469.0)	< 0.001
Arsenic (mg%)	0.008 (0.005, 0.012)	0.007 (0.005, 0.011)	0.003
Mercury (mg%)	0.110 (0.080, 0.190)	0.090 (0.060, 0.130)	< 0.001
Cadmium (mg%)	0.001 (0.001, 0.002)	0.001 (0.001, 0.002)	0.018

Table 2. General characteristics of the study population. *LBM* low bone mass, *BMI* body mass index, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *DM* diabetes mellitus, *WBC* white blood cell, *seg* differential count of segment neutrophil, *lym* differential count of lymphocyte, *Hb* hemoglobin, *MCV* mean corpuscular volume, *Ca* calcium, *P* phosphorus, *Cl* chloride, *BUN* blood urea nitrogen, *AST* aspartate aminotransferase, *ALT* alanine aminotransferase, *GGT* gamma-glutamyl transferase, *ALP* alkaline phosphatase, *T.bil* total bilirubin, *HDL-C* high-density lipoprotein cholesterol, *LDL-C* low-density lipoprotein cholesterol, *TSH* thyroid stimulating hormone, *IGF-1* Insulin-like growth factor 1, *DHEA-S* dehydroepiandrosterone sulfate, *AFP* alpha-fetoprotein, *CEA* carcinoembryonic antigen, *CA 19–9* carbohydrate antigen, *CTX* C-terminal telopeptide of type 1 collagen.

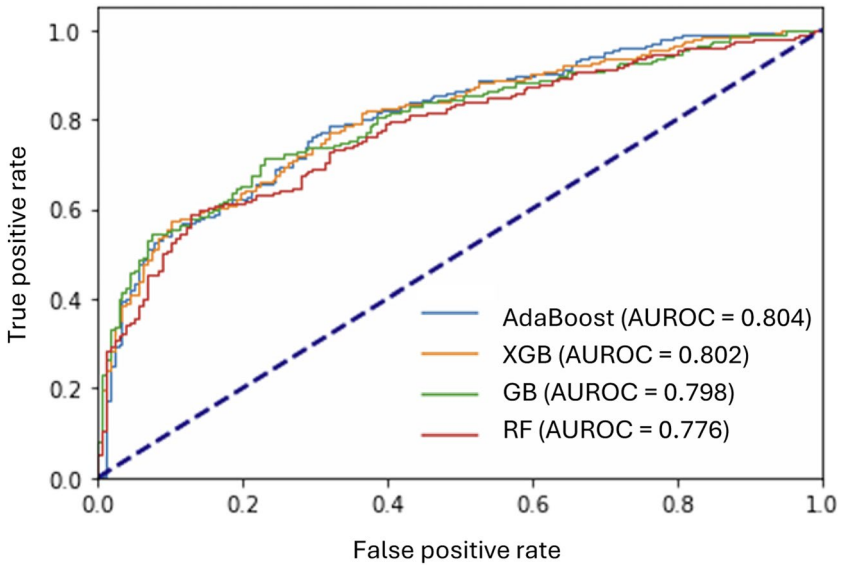

Figure 2. A sample of comparison among the receiver operating characteristic curves of the machine learning models. *AUROC* area under the receiver operating characteristic curve, *RF* Random Forest, *XGB* Extreme Gradient Boosting, *GB* Gradient Boosting, *AdaBoost* Adaptive Boosting.

	RF	XGB	GB	AdaBoost
Accuracy	0.682 (0.677–0.688)	0.687 (0.682–0.691)	0.678 (0.672–0.683)	0.687 (0.681–0.693)
Precision	0.713 (0.704–0.722)	0.709 (0.701–0.717)	0.697 (0.688–0.705)	0.705 (0.696–0.713)
NPV	0.630 (0.613–0.647)	0.643 (0.632–0.654)	0.648 (0.634–0.662)	0.649 (0.637–0.661)
Sensitivity	0.782 (0.773–0.792)	0.807 (0.799–0.815)	0.818 (0.804–0.832)	0.817 (0.810–0.825)
Specificity	0.525 (0.511–0.539)	0.511 (0.500–0.523)	0.469 (0.450–0.489)	0.495 (0.483–0.507)
F1 score	0.745 (0.734–0.750)	0.754 (0.750–0.758)	0.751 (0.745–0.756)	0.757 (0.752–0.761)

Table 3. Summary of the performance metrics of each model. The values are presented as means and 95% confidence interval. *RF* Random Forest, *XGB* Extreme Gradient Boosting, *GB* Gradient Boosting, *AdaBoost* Adaptive Boosting, *NPV* negative predictive value.

classification, instead of osteoporosis. Clinically, osteoporosis and osteopenia are classified based on T-scores measured by DXA. Since fracture risk is elevated also in osteopenia compared to normal bone mass, we focused on LBM. In addition, the relatively small sample size of osteoporosis may result in imbalance in the prediction target classification; however, classification based on LBM and normal could achieve a better balance.

This study has concerns about the applicability to clinical practice due to the large number of input features. For example, OSTA has traditionally calculated the risk of osteoporosis in postmenopausal women based solely on weight and age, which were also the most two important features identified in our study[7]. However with advancements in technology enabling rapid computation of complex calculations, it may be feasible to discover significantly greater number of patients with low bone mass by employing a slightly larger set of variables. In this regard, our study utilized widely collected clinical data in the models. The selection of the most significant features or classes can facilitate the construction of simpler models with comparable performance. Further investigation into feature selection methodologies may contribute to improving the predictive models.

Our study is the first attempt to create a machine learning model predicting low bone mass using hair mineral analysis data. While serum mineral concentrations maintain hemostatic balance, hair mineral concentrations can reflect an individual's overall mineral status and identify imbalances or deficiencies[16]. Several studies have investigated the relationship between bone density and hair mineral content. Kotkowiak reported that hair calcium and magnesium levels were low in patients with osteoporosis[35]. In Korean studies, BMD was associated with hair magnesium or selenium levels.[17,18] Japanese researchers have found that self-reported osteoporosis is significantly associated with advanced age and low phosphorus and calcium concentrations in hair[36]. Given the diversity of results, focusing on the interaction of elements rather than specific hair mineral elements is more advisable. In this regard, ML analysis, which identifies patterns and correlations within large-scale datasets, is believed to be even more helpful. The concept of feature importance in machine learning involves identifying variables that have the most significant impact on model performance and predictions. Higher scores indicate a larger effect on the model. In our study, we were able to identify variables that significantly influenced the prediction of low bone density in patients including several hair minerals. Specifically, hair sulfur have similar impact on the SHAP values to several traditional risk factors such as waist circumference and body height. Although

Ranking	Feature	Feature importance (SHAP values)
1	Body weight	0.01647579
2	Age	0.01496218
3	ALP	0.01445836
4	Osteocalcin	0.01440475
5	Body height	0.01438183
6	DHEA-S	0.01421377
7	Hair sulfur	0.01414776
8	AFP	0.01412552
9	Hb	0.01405549
10	Waist circumference	0.01404894
11	CEA	0.01398897
12	Vitamin B12	0.01398510
13	IGF-1	0.01395223
14	BMI	0.01394938
15	BUN	0.01394516
16	Cl	0.01392897
17	Uric acid	0.01392838
18	Folate	0.01392610
19	Fasting glucose	0.01391230
20	TSH	0.01390031

Table 4. Top 20 important features of the Extreme Gradient Boosting (XGB) model. *ALP* alkaline phosphatase, *DHEA-S* dehydroepiandrosterone sulfate, *AFP* alpha-fetoprotein, *Hb* hemoglobin, *CEA* carcinoembryonic antigen, *IGF-1* insulin-like growth factor 1, *BMI* body mass index, *BUN* blood urea nitrogen, *Cl* chloride, *TSH* thyroid stimulating hormone.

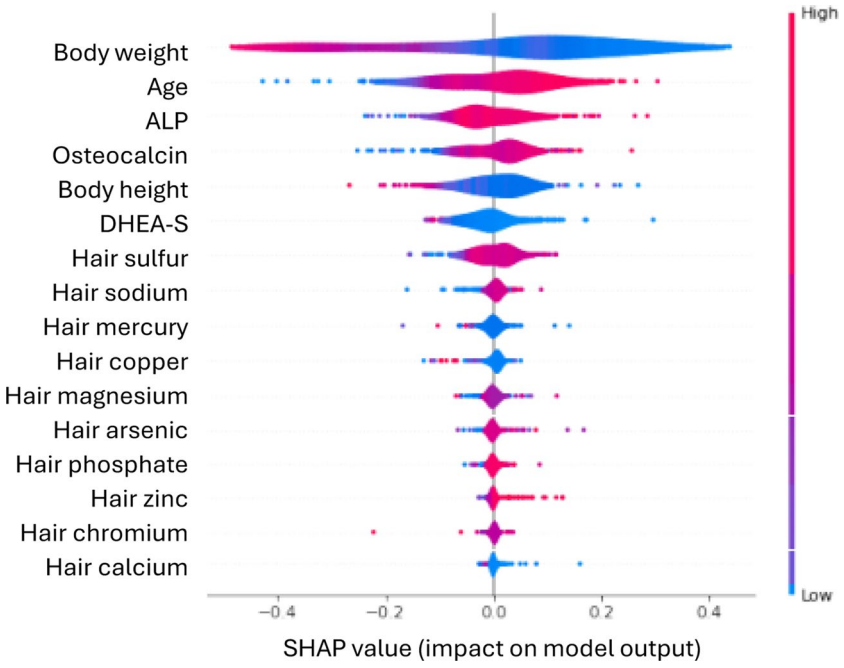

Figure 3. Summary plot of the Shapley additive explanation (SHAP) values for the top 7 significant features and the top 9 important hair minerals. A positive SHAP value indicates a positive impact on the model's prediction. Red indicates high values, and blue indicates low values. ALP alkaline phosphatase, DHEA-S dehydroepiandrosterone sulfate.

it is not shown in the table, considering that the SHAP value of CTX, which are known to be highly associated with osteoporosis, was 0.01389796, hair sodium, mercury, copper and magnesium may also have similar level of importance to this bone turnover marker (Supplementary Table S3)[37].

Maintaining precise levels of calcium and phosphate is essential for life. A complex mechanism involving many hormones is responsible for modulating these mineral levels[38]. Physicials typically check serum and urine calcium levels when diagnosing and evaluating patients with low bone mass. Serum calcium levels usually remain within the normal range due to this intricate homeostasis mechanism. Therefore, occasionally checking urine calcium, fasting calcium/creatinine urine levels and hypercalciuria (260 mg/24 h) can be beneficial[39]. In our study, hair calcium was not found to be more significant than hair cooper or sulfur. Similarly, like serum calcium, assessing only calcium levels does not reflect and individual's overall mineral stats or bone health. This emphasizes the need to consider the interaction of multiple minerals over time, given that hair mineral assay indicates long-term changes in mineral levels.

Our study faces several challenges. First, our study has a relatively small sample size. Basically, ML models require a sufficient sample size to achieve adequate performance. Nevertheless, a wide range of features including hair mineral analysis were encompassed in our models. Previous ML studies aimed to predict osteoporosis have utilized a similar number of samples to ours[31,32]. We are currently in the process of accumulating more data, and plan to conduct further analyses with a larger sample size. Second, the prediction target can vary. Other studies have categorized subjects into osteoporosis and non-osteoporosis groups[14,32]. Osteopenia poses a higher risk for fractures than normal group and shows higher number of fracture occurrences than in the osteoporosis group[40]. Hence, LBM holds significant clinical importance. Although BMD can be directly used as a continuous outcome variable, this type of ML model may be feasible when utilizing various image data[41]. Third, several data were not collected. Medication use data such as oral contraceptives[42], hormone replacement therapy[9], and use of steroids[43], in along with additional markers for bone metabolism such as parathyroid hormone and serum 25-hydroxyvitamin D[44], can contribute to building a more robust model.

Despite these limitations, our study has unique characteristics. Our study encompasses a wide range of features and is the first attempt to include hair mineral analysis in ML models. Moreover, integrating deep learning with supplementary imaging data, alongside training using clinical databases, may provide new insights into the early diagnosis of patients at risk of fractures.

Conclusions

The study findings emphasize the significance of employing ML algorithms to predict LBM using health checkup data and hair mineral analysis. While XGB exhibited superior performance, achieving the highest average AUROC of 0.744 and a maximum 0.802, all four algorithms demonstrated comparable performance. We compared the feature importance of traditional factors, known to be associated with osteoporosis, with that of hair minerals, and our findings identified several hair minerals exhibiting comparable levels of importance. While further research with larger datasets is necessary, this study demonstrates that integrating health checkup data and hair mineral analysis into these ML models may provide valuable insights into identifying individuals at risk of LBM.

Data availability

The datasets obtained and/or analyzed during the present study are available from the corresponding author on reasonable request.

Received: 12 January 2024; Accepted: 31 July 2024
Published online: 13 August 2024

References

1. Rachner, T. D., Khosla, S. & Hofbauer, L. C. Osteoporosis: Now and the future. *Lancet* **377**, 1276–1287 (2011).
2. Lane, N. E. Epidemiology, etiology, and diagnosis of osteoporosis. *Am. J. Obstetr. Gynecol.* **194**, S3–S11. https://doi.org/10.1016/j.ajog.2005.08.047 (2006).
3. Anam, A. K. & Insogna, K. Update on osteoporosis screening and management. *Med. Clin. North Am.* **105**, 1117–1134. https://doi.org/10.1016/j.mcna.2021.05.016 (2021).
4. Cummings, S. R. & Melton, L. J. Epidemiology and outcomes of osteoporotic fractures. *Lancet* **359**, 1761–1767. https://doi.org/10.1016/s0140-6736(02)08657-9 (2002).
5. Siris, E. S. *et al.* Bone mineral density thresholds for pharmacological intervention to prevent fractures. *Arch. Intern. Med.* **164**, 1108–1112. https://doi.org/10.1001/archinte.164.10.1108 (2004).
6. Rubin, K. H. *et al.* Comparison of different screening tools (FRAX(R), OST, ORAI, OSIRIS, SCORE and age alone) to identify women with increased risk of fracture: A population-based prospective study. *Bone* **56**, 16–22. https://doi.org/10.1016/j.bone.2013.05.002 (2013).
7. Koh, L. K. H. *et al.* A simple tool to identify asian women at increased risk of osteoporosis. *Osteop. Int.* **12**, 699–705. https://doi.org/10.1007/s001980170070 (2001).
8. Cadarette, S. M. *et al.* Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. *Cmaj* **162**, 1289–1294 (2000).
9. Cadarette, S. M., Jaglal, S. B. & Murray, T. M. Validation of the simple calculated osteoporosis risk estimation (SCORE) for patient selection for bone densitometry. *Osteoporos. Int.* **10**, 85–90. https://doi.org/10.1007/s001980050199 (1999).
10. Sedrine, W. B. *et al.* Development and assessment of the Osteoporosis Index of Risk (OSIRIS) to facilitate selection of women for bone densitometry. *Gynecol. Endocrinol.* **16**, 245–250 (2002).
11. Roth, J. A., Battegay, M., Juchler, F., Vogt, J. E. & Widmer, A. F. Introduction to machine learning in digital healthcare epidemiology. *Infect. Control Hosp. Epidemiol.* **39**, 1457–1462. https://doi.org/10.1017/ice.2018.265 (2018).
12. Chen, J. H. & Asch, S. M. Machine learning and prediction in medicine: Beyond the peak of inflated expectations. *N. Engl. J. Med.* **376**, 2507–2509. https://doi.org/10.1056/NEJMp1702071 (2017).
13. Deo, R. C. Machine learning in medicine. *Circulation* **132**, 1920–1930. https://doi.org/10.1161/circulationaha.115.001593 (2015).

14. Yoo, T. K. *et al.* Osteoporosis risk prediction for bone mineral density assessment of postmenopausal women using machine learning. *Yonsei Med. J.* **54**, 1321–1330. https://doi.org/10.3349/ymj.2013.54.6.1321 (2013).
15. Lee, Y. A., Kim, H. N. & Song, S.-W. Associations between hair mineral concentrations and skeletal muscle mass in Korean adults. *J. Nutr. Health Aging* **26**, 515–520. https://doi.org/10.1007/s12603-022-1789-5 (2022).
16. Sky-Peck, H. H. Distribution of trace elements in human hair. *Clin. Physiol. Biochem.* **8**, 70–80 (1990).
17. Song, C. H., Barrett-Connor, E., Chung, J. H., Kim, S. H. & Kim, K. S. Associations of calcium and magnesium in serum and hair with bone mineral density in premenopausal women. *Biol. Trace Elem. Res.* **118**, 1–9. https://doi.org/10.1007/s12011-007-0011-2 (2007).
18. Park, K. C. *et al.* Low selenium levels are associated with decreased bone mineral densities. *J. Trace Elem. Med. Biol.* **61**, 126534. https://doi.org/10.1016/j.jtemb.2020.126534 (2020).
19. Park, S. J. *et al.* Hair calcium concentration is associated with calcium intake and bone mineral density. *Int. J. Vitam. Nutr. Res.* **83**, 154–161. https://doi.org/10.1024/0300-9831/a000160 (2013).
20. Prevention and management of osteoporosis. World Health Organ Tech Rep Ser 921, 1–164, (2003).
21. Gunzerath, L., Faden, V., Zakhari, S. & Warren, K. National Institute on alcohol abuse and alcoholism report on moderate drinking. *Alcohol. Clin. Exp. Res.* **28**, 829–847. https://doi.org/10.1097/01.alc.0000128382.79375.b6 (2004).
22. Piercy, K. L. *et al.* The physical activity guidelines for Americans. *JAMA* **320**, 2020–2028. https://doi.org/10.1001/jama.2018.14854 (2018).
23. Khan, I. U. *et al.* A proactive attack detection for heating, ventilation, and air conditioning (HVAC) system using explainable extreme gradient boosting model (XGBoost). *Sensors* **22**, 9235 (2022).
24. Chen, T. & Guestrin, C. in Proceedings of the 22nd acm sigkdd International Conference on Knowledge Discovery and Data Mining. 785–794.
25. Ridgeway, G. The state of boosting. Comput. Sci. Stat. 172–181 (1999).
26. Zou, K. H., O'Malley, A. J. & Mauri, L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* **115**, 654–657 (2007).
27. Albahra, S. *et al.* Artificial intelligence and machine learning overview in pathology & laboratory medicine: A general review of data preprocessing and basic supervised concepts. *Semin. Diagn. Pathol.* **40**, 71–87. https://doi.org/10.1053/j.semdp.2023.02.002 (2023).
28. Lundberg, S. M., Erion, G. G. & Lee, S.-I. Consistent individualized feature attribution for tree ensembles. arXiv preprint arXiv:1802.03888 (2018).
29. Nohara, Y., Matsumoto, K., Soejima, H. & Nakashima, N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Comput. Methods Progr. Biomed.* **214**, 106584. https://doi.org/10.1016/j.cmpb.2021.106584 (2022).
30. Lee, I.-J. & Lee, J. Predictive of osteoporosis by tree-based machine learning model in post-menopause woman. *J. Radiol. Sci. Technol.* **43**, 495–502. https://doi.org/10.17946/jrst.2020.43.6.495 (2020).
31. Shim, J.-G. *et al.* Application of machine learning approaches for osteoporosis risk prediction in postmenopausal women. *Arch. Osteop.* **15**, 1–9 (2020).
32. Kwon, Y. *et al.* Osteoporosis pre-screening using ensemble machine learning in postmenopausal Korean women. *Healthcare* https://doi.org/10.3390/healthcare10061107 (2022).
33. Yang, O., Lai, W. Y., Tsou, M. T. & Hwang, L. C. Development of machine learning models for prediction of osteoporosis from clinical health examination data. *Int. J. Environ. Res. Public Health* https://doi.org/10.3390/ijerph18147635 (2021).
34. Inui, A. *et al.* Screening for osteoporosis from blood test data in elderly women using a machine learning approach. *Bioengineering* https://doi.org/10.3390/bioengineering10030277 (2023).
35. Kotkowiak, L. Behavior of selected bio-elements in women with osteoporosis. *Ann. Acad. Med. Stetin* **43**, 225–238 (1997).
36. Yamashiro, K., Ogata, F., Nakamura, T., Tanei, S. & Kawasaki, N. Relationship between self-reported osteoporosis and mineral concentrations in female hair. *J. Allied Health Sci.* **12**, 16–23 (2021).
37. Morris, H. *et al.* Clinical usefulness of bone turnover marker concentrations in osteoporosis. *Clinica chimica acta* **467**, 34–41 (2017).
38. Boden, S. D. & Kaplan, F. S. Calcium Homeostasis. *Orthopedic Clinics of North America* **21**, 31–42. https://doi.org/10.1016/S0030-5898(20)31563-7 (1990).
39. Rull, M. A., Cano-García Mdel, C., Arrabal-Martín, M. & Arrabal-Polo, M. A. The importance of urinary calcium in postmenopausal women with osteoporotic fracture. *Can. Urol. Assoc. J.* **9**, 183–186. https://doi.org/10.5489/cuaj.2695 (2015).
40. Khosla, S. & Melton, L. J. 3rd. Clinical practice. *Osteopenia. N. Engl. J. Med.* **356**, 2293–2300. https://doi.org/10.1056/NEJMcp070341 (2007).
41. He, Y., Lin, J., Zhu, S., Zhu, J. & Xu, Z. Deep learning in the radiologic diagnosis of osteoporosis: A literature review. *J. Int. Med. Res.* **52**, 3000605241244754. https://doi.org/10.1177/03000605241244754 (2024).
42. Gambacciani, M., Monteleone, P., Ciaponi, M., Sacco, A. & Genazzani, A. R. Effects of oral contraceptives on bone mineral density. *Treat. Endocrinol.* **3**, 191–196. https://doi.org/10.2165/00024677-200403030-00006 (2004).
43. Canalis, E., Mazziotti, G., Giustina, A. & Bilezikian, J. P. Glucocorticoid-induced osteoporosis: Pathophysiology and therapy. *Osteop. Int.* **18**, 1319–1328 (2007).
44. Sai, A. J., Walters, R., Fang, X. & Gallagher, J. Relationship between vitamin D, parathyroid hormone, and bone health. *J. Clin. Endocrinol. Metab.* **96**, E436–E446 (2011).

Acknowledgements

The authors thank all staff for their help with recruiting the subjects, conducting the measurements, and documenting the data in CHA Bundang Medical Center. The authors would like to thank Enago(www.enago.co.kr) for the English language review.

Author contributions

Study concept and design: Y.-S.K. and S.-J.K.; acquisition of data: Y.-S.K. and S.-J.K; analysis and interpretation of data: S.-J.K., J.-O.K., and Y-S.K.; software: S-J.K. and J.-O.K.; writing—original draft preparation: S.-J.K. and Y.-S.K.; review and editing: J.-O.K., M.-J.K., Y.-I.H., J.-H.H., K.-H.H.

Funding

This study was supported by the Korean Fund for Regenerative Medicine (KFRM) grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Health & Welfare, grant number 23C0102L1).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-69090-3.

Correspondence and requests for materials should be addressed to Y.-S.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.