JBM
J Bone Metab 2023;30(3):263-273
https://doi.org/10.11005/jbm.2023.30.3.263
pISSN 2287-6375 eISSN 2287-7029

Original Article

Osteoporosis Feature Selection and Risk Prediction Model by Machine Learning Using a Cross-Sectional Database

Yonghan Cha[1,*], Sung Hyo Seo[2,*], Jung-Taek Kim[3], Jin-Woo Kim[4], Sang-Yeob Lee[2], Jun-Il Yoo[5]

[1]Department of Orthopaedic Surgery, Daejeon Eulji Medical Center, Eulji University School of Medicine, Daejeon;
[2]Department of Biomedical Research Institute, Gyeongsang National University Hospital, Jinju;
[3]Department of Orthopedic Surgery, Ajou Medical Center, Ajou University School of Medicine, Suwon;
[4]Department of Orthopaedic Surgery, Nowon Eulji Medical Center, Eulji University School of Medicine, Seoul;
[5]Department of Orthopaedic Surgery, Inha University Hospital, Inha University School of Medicine, Incheon, Korea

Corresponding author
Jun-Il Yoo
Department of Orthopaedic Surgery, Inha University Hospital, Inha University School of Medicine, 27 Inhang-ro, Jung-gu, Incheon 22332, Korea
Tel: +82-32-890-3663
Fax: +82-55-754-0477
E-mail: furim@hanmail.net

*Yonghan Cha and Sung Hyo Seo contributed equally to this work and should be considered co-first authors.

Received: June 14, 2023
Revised: July 3, 2023
Accepted: July19, 2023

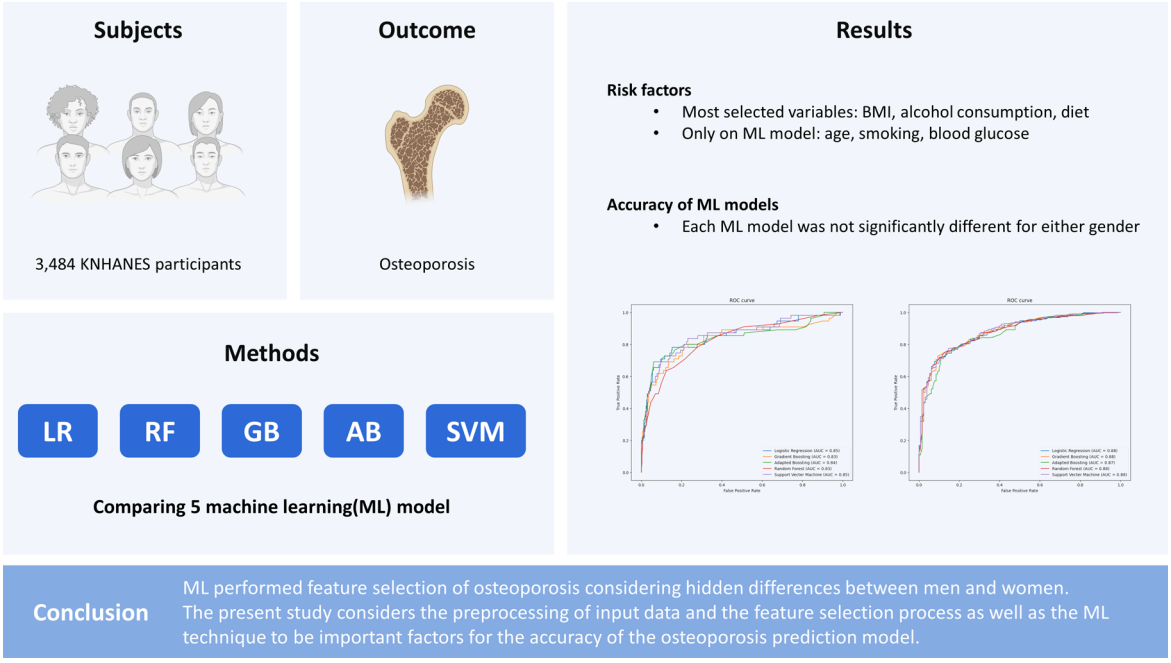
Background: The purpose of this study was to verify the accuracy and validity of using machine learning (ML) to select risk factors, to discriminate differences in feature selection by ML between men and women, and to develop predictive models for patients with osteoporosis in a big database. **Methods:** The data on 968 observed features from a total of 3,484 the Korea National Health and Nutrition Examination Survey participants were collected. To find preliminary features that were well-related to osteoporosis, logistic regression, random forest, gradient boosting, adaptive boosting, and support vector machine were used. **Results:** In osteoporosis feature selection by 5 ML models in this study, the most selected variables as risk factors in men and women were body mass index, monthly alcohol consumption, and dietary surveys. However, differences between men and women in osteoporosis feature selection by ML models were age, smoking, and blood glucose level. The receiver operating characteristic (ROC) analysis revealed that the area under the ROC curve for each ML model was not significantly different for either gender. **Conclusions:** ML performed a feature selection of osteoporosis, considering hidden differences between men and women. The present study considers the preprocessing of input data and the feature selection process as well as the ML technique to be important factors for the accuracy of the osteoporosis prediction model.

Key Words: Machine learning · Osteoporosis · Risk assessment · Risk factors

INTRODUCTION

As the elderly population increases, many social and economic burdens caused by osteoporosis and osteoporotic fractures have already been reported in developed countries.[1] Providing these patients with various services for evaluation and management of osteoporosis and fracture prevention not only increases patients' satisfaction with medical care but also has the effect of reducing the socioeconomic burden required to manage them.[2] However, although dual energy X-ray absorptiometry (DXA) is one of the preferred modalities for screening or diagnosis of osteoporosis and can predict the risk of osteoporotic fracture to some

Graphical Abstract



extent, it is difficult to screen osteoporosis unless the patient visits a hospital.[3,4] In addition, the purpose of visiting the hospital may affect whether or not DXA is implemented.[4] In busy outpatient clinics, it is easy for medical staff to miss out on the evaluation and management of osteoporosis.[5] Therefore, preventing or diagnosing osteoporosis early may be more effective than treating osteoporosis after disease progression or osteoporotic fractures occur.

Machine learning (ML) is a method of self-learning using data without special instructions to find complex patterns and create models.[6,7] A computational modeling tool using ML of artificial intelligence (AI) is widely used to predict the occurrence of disease or fracture and to estimate clinical outcomes in various clinical areas.[8,9] ML has made important improvements and discoveries in the medical field and is also being applied in rheumatism and osteoarthritis and has achieved many achievements.[10-12] The prediction model using ML can compensate for the shortcomings of conventional examination methods and handle large numbers of input variables simultaneously. Also, if an automated system of AI is constructed, there is an advantage that the hassle of checking examinations can be solved.[13] Thus, using ML and big data, it is possible to make a big change in the clinical pathway, which was previously used to screen and manage osteoporosis in postmenopausal women or old ages, to discover osteoporosis risk factors at an early stage and to implement interventions.

The purpose of this study was to verify the accuracy and validity of the use of ML to select risk factors, to discriminate differences in feature selection by ML between men and women, and to develop predictive models for patients with osteoporosis in a big database. The hypothesis of this study is that osteoporosis risk factors might be different between men and women in risk prediction models using ML.

METHODS

1. Ethics statement

Data from the 2008 to 2011 Korea National Health and Nutrition Examination Surveys (KNHANESs) were reviewed and approved by the Institutional Review Board of the Korea Centers for Disease Control and Prevention (Approval no. 2008–04EXP-01-C, 2009–01CON-03-C, 2010–02CON-

https://doi.org/10.11005/jbm.2023.30.3.263

21-C, and 2011–02CON-06-C). Informed consent was obtained from each participant when the 2008, 2009, 2010, and 2011 KNHANESs were conducted.

2. Participants

This study was based on data obtained from the 2008 to 2011 KNHANES conducted by the Korea Ministry of Health and Welfare. KNHANES is a nationwide representative cross-sectional survey of the Korean population; it uses a clustered, multistage, stratified, and rolling sampling design. It consists of 3 sections: a health interview, a health examination, and a dietary survey. More than 500 variables are examined each year the survey is conducted. These variables are included in a health questionnaire and in laboratory findings; data on nutritional factors are also collected. [14] Survey data are collected via household interviews and direct standardized physical examinations performed in specially-equipped mobile examination centers.

The data considered for use in this study were collected from a total of 37,753 the KNHANES participants (2008, 9,744 persons; 2009, 10,533 persons; 2010, 8,958 persons; and 2011, 8,518 persons). However, participants were excluded if they were non-menopausal if female, or less than 50 years old if male, or if the data required to evaluate skeletal muscle mass and dietary intake were unavailable. After these exclusions, data from a total of 3,484 participants (male 1,601, female 1,883) were included in the analysis (Fig. 1).

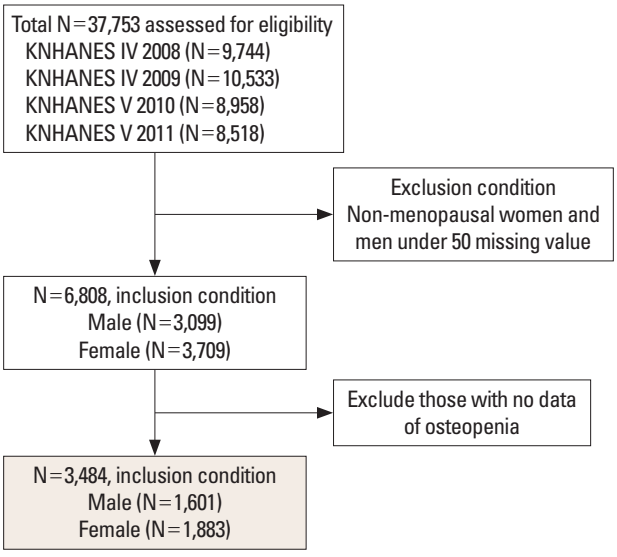


Fig. 1. Study subject selection process, the Korean National Health and Nutrition Examination Survey (KNHANES) IV and V (2008–2011).

3. Measurement of bone mineral density (BMD) and diagnosis for osteoporosis

BMD (g/cm^2) at the lumbar spine, femoral neck, and total proximal femur were measured by DXA (Hologic Inc., Bedford, MA, USA). According to the World Health Organization study group, the diagnosis of osteoporosis is based on T-score thresholds. T-scores at or above -1.0 are considered normal, those between -1.0 and -2.5 as osteopenia, and those at or below -2.5 as osteoporosis.

4. ML (random forest model) and knowledge-based feature selection

The data on 968 observed features of 1,601 male and 1,883 female participants from the KNHANES results were collected. During data curation, we manually excluded columns for unrelated features and features with missing values for more than 900 individuals. Then, the data from individuals without missing values were re-collected. A "osteoporosis" column was used as the classification label for the supervised learning. To find preliminary features that were well-related to osteoporosis, logistic regression (LR), random forest (RF), gradient boosting (GB), adaptive boosting (AB), and support vector machine (SVM) were used. A total of 53 candidate variables were used to train each of the 5 models, and features that were judged to be highly related to osteoporosis were selected. For hyperparameter tuning of 5 models, model optimization was performed using grid search. LR with least absolute shrinkage and selection operator (L1 norm) regularization was used to select features with coefficients other than 0, and for the remaining 4 models, permutation importance scores were calculated and features greater than 0 were selected. Among the features that appeared to be related to osteoporosis in each of the 5 models, features common to 3 or more models were selected. In the model using male data, 57 features were selected and, in the model, using female data, 54 features were selected, and selected risk variables that did not have clinically overlapping features.

These selected risk factors were used to build predictive models for osteoporosis based on 5 classification algorithms (i.e., LR, RF, GB, AB, SVM). To determine the hyperparameter, we implemented 5-fold cross-validation of the training set with 30 times train/validation set shuffling and using the hyperparameters corresponding to each model. Each of the 5 classification models was built on the training set

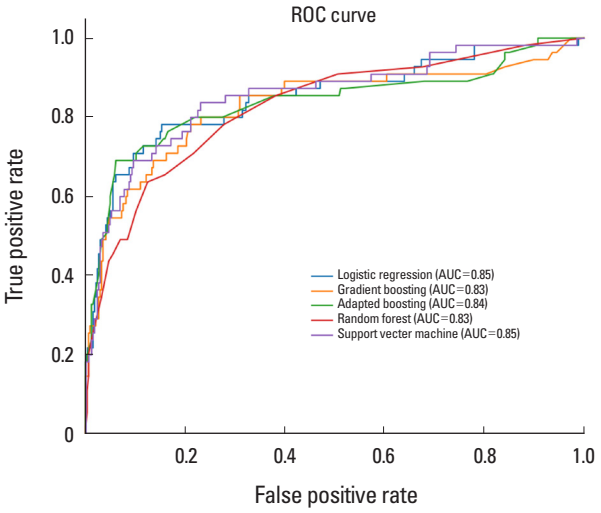


Fig. 2. Receiver operating characteristic (ROC) curve analysis for various machine learning models, data from elderly male participants. AUC, area under the curve.

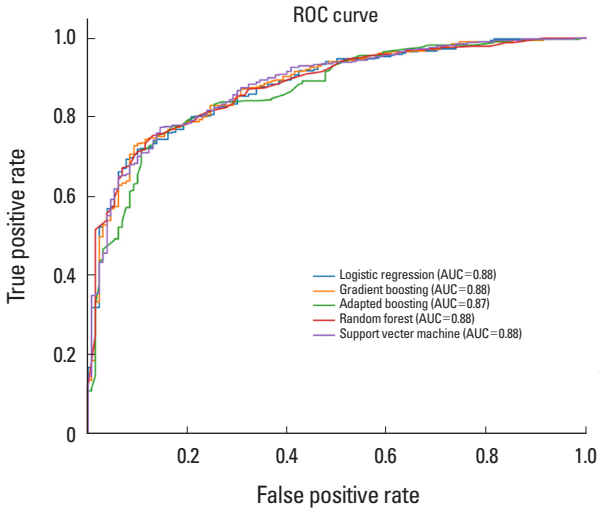


Fig. 3. Receiver operating characteristic (ROC) curve analysis for various machine learning models, data from elderly female participants. AUC, area under the curve.

with the selected hyperparameters. Each model's predicted probability of osteoporosis classification on the test set was subjected to receiver operating characteristic (ROC) curve analysis to obtain the reliability of the final model (Fig. 2, 3).

RESULTS

Among the 3,484 who were included in the study population, 1,601 (46.0%) were male and 1,883 (54.0%) were female. The features selected for men using the 5 methods are presented in Table 1. The significant feature in all 5 models was body mass index (BMI), and 23 features were selected from 4 models.

The features selected for women using the 5 methods are presented in Table 2. Significant features in all 5 models were monthly alcohol consumption, BMI, white blood cell, thrombocyte, food intake (g/day), water intake (g/day), calcium intake (mg/day), and niacin intake (mg/day). In osteoporosis feature selection by 5 ML models in this study, the most selected variables as risk factors in men and women were BMI, monthly alcohol consumption, and dietary surveys. However, differences between men and women in osteoporosis feature selection by ML models were age, smoking, and blood glucose level.

The ROC analysis revealed that the area under the ROC curve (AUC) for each ML model was not significantly different for either gender (LR: 0.85 in male, 0.88 in female; RF: 0.83 in men, 0.88 in women; GB: 0.83 in men, 0.88 in women; AB: 0.84 in men, 0.87 in women; SVM: 0.85 in men, 0.81 in women) (Table 3).

DISCUSSION

In osteoporosis feature selection by 5 ML models in the present study, the most selected variables as risk factors in men and women were BMI, monthly alcohol consumption, and dietary surveys. It is well known that chronic alcohol consumption is associated with osteoporosis and osteoporotic fractures.[15] Ethanol inhibits the differentiation and mineralization of osteoblasts, and it is reported that the risk increases as the amount of alcohol consumed increases.[16] Adequate nutritional status is also essential for maintaining proper skeletal structure.[17] The fact that adequate protein intake and low levels of serum albumin are related to the development of osteoporosis also means that nutritional status is an important factor influencing the development of osteoporosis.[18] Also, inadequate nutritional status can reduce daily activities and increase the hospitalization period and recovery time after an osteoporotic fracture.[17] Therefore, Huang et al. [19] reported that the geriatric nutritional risk index considering nutritional status and BMI is useful for assessing the risk of osteoporosis. The variables selected from ML models are traditionally well-

https://doi.org/10.11005/jbm.2023.30.3.263

Table 1. Scoring for common risk factors according to different machine learning models, data from male

Variables	Logistic regression	Random forest	Adaptive boosting	Gradient boosting	Support vector machine	Score
Health examination						
Smoking	o	o	x	x	o	3
High risk alcohol consumption	o	o	x	x	x	2
Monthly alcohol consumption	o	o	x	o	o	4
High-intensity physical activity	o	x	x	x	o	2
Middle-intensity physical activity	o	x	x	x	x	1
Walking participation	o	o	x	o	o	4
Body mass index	o	o	o	o	o	5
Waist circumference	o	o	x	x	o	3
Health laboratory tests						
Blood glucose	o	o	x	o	o	4
Cholesterol	x	o	x	o	x	2
High-density lipoprotein	o	o	x	o	o	4
Triglyceride	o	o	x	o	x	3
Aspartate transaminase	x	o	x	o	o	3
Alanine transferase	o	o	x	o	x	3
Hemoglobin	o	o	x	o	o	4
Hematocrit	x	o	x	o	o	3
Blood urea nitrogen	o	o	x	o	o	4
Blood creatinine	o	o	x	x	x	2
White blood cell	o	o	x	x	x	2
Red blood cell	x	o	x	o	x	2
Thrombocyte	o	o	x	o	o	4
Urine pH	o	o	x	o	x	3
Sodium nitrite	o	x	x	x	x	1
Urine specific gravity	x	x	x	x	x	0
Urine protein	o	x	x	x	o	2
Urine glucose	o	x	x	o	o	3
Urine ketone	o	o	x	x	o	3
Urine bilirubin	o	x	x	x	o	2
Urine blood	o	o	x	o	o	4
Urobilinogen	o	x	x	x	o	2
Urine sodium	o	o	x	o	o	4
Dietary survey						
Food intake	o	o	x	o	o	4
Energy intake	x	o	x	x	o	2
Water intake	o	o	x	o	x	3
Protein intake	o	o	o	o	x	4
Fat intake	o	o	x	o	x	3
Carbohydrate intake	o	o	x	o	o	4
Calcium intake	o	o	x	o	o	4
Phosphorus intake	o	o	x	x	o	3
Iron intake	o	o	x	o	o	4
Sodium intake	o	o	x	o	o	4
Potassium intake	o	o	x	o	o	4

(*Continued to the next page*)

Table 1. Continued

Variables	Logistic regression	Random forest	Adaptive boosting	Gradient boosting	Support vector machine	Score
β-carotene intake	o	o	x	o	o	4
Retinol intake	o	o	x	o	x	3
Thiamine intake	o	o	x	o	o	4
Riboflavin intake	o	o	x	o	x	3
Niacin intake	o	o	x	x	o	3
Vitamin-C intake	o	o	x	o	x	3
Health interview						
Age (30's year)	x	x	x	x	x	0
Age (40's year)	x	x	x	x	x	0
Age (50's year)	o	o	x	o	o	4
Age (60's year)	x	o	x	x	x	1
Age (≥70 year)	o	o	x	o	o	4
Living in a house	o	x	x	x	o	2
Living in apartment	o	x	x	x	o	2
Household income (low)	o	o	x	o	o	4
Household income (middle-low)	o	x	x	x	x	1
Household income (middle-high)	o	o	x	x	o	3
Household income (high)	o	x	x	o	o	3
Education (elementary school and below)	o	o	x	o	o	4
Education (middle school)	o	o	x	x	x	2
Education (high school)	o	o	x	x	o	3
Education (university graduate or higher)	o	o	x	x	o	3
Job (manager, expert)	o	x	x	x	o	2
Job (office workers)	x	x	x	x	o	1
Job (service, sales)	o	o	x	x	o	3
Job (forestry and fishery workers)	o	o	x	x	x	2
Job (technician, device/machine operator)	o	x	x	x	o	2
Job (simple labor workers)	o	o	x	x	x	2
Job (unemployed)	o	o	x	o	o	4

known risk factors of osteoporosis. However, among traditionally well-known risk factors of osteoporosis, there are differences in the extent to which men and women contribute to the development of osteoporosis. Men have a lower incidence of osteoporosis than women because they basically have more bone mass than women and have a larger physique.[20] In addition, age is an important factor, especially in women, as osteoporosis increases with age after menopause.[21] It has been reported that the smoking rate of men is higher than that of women, and this difference may indicate the extent to which smoking affects the incidence of osteoporosis depending on sex.[22] Secondary osteoporosis is more common in men than women, and diabetes is one of the main causes of secondary osteoporosis.[23] In the present study, differences between men and women in osteoporosis feature selection by several ML models were age, smoking, and blood glucose level. It is surprising that ML models automatically considered such substantial male-female differences and selected them as risk factors. ML is used in various medical fields because it can learn patterns from input data and predict results considering various hidden relationships.[24] We believe that perhaps this might be a higher-dimensional process than the selection of risk factors based on expert knowledge.[25] In the current ML model, food intake (g/day), water intake (g/day), calcium intake (mg/day), and Niacin intake (mg/day) were identified as risk factors related to osteoporosis. Similar to these results, the study of Park et al. [26]

https://doi.org/10.11005/jbm.2023.30.3.263

Table 2. Scoring for common risk factors according to different machine learning models, data from female

Variables	Logistic regression	Random forest	Adaptive boosting	Gradient boosting	Support vector machine	Score
Health examination						
Smoking	o	x	x	x	x	1
High risk alcohol consumption	x	x	x	x	x	0
Monthly alcohol consumption	o	o	o	o	o	5
High-intensity physical activity	o	x	x	o	x	2
Middle-intensity physical activity	o	x	x	x	o	2
Walking participation	o	x	x	x	x	1
Body mass index	o	o	o	o	o	5
Waist circumference	o	o	x	o	o	4
Health laboratory tests						
Blood glucose	o	x	x	o	x	2
Cholesterol	o	x	o	o	x	3
High-density lipoprotein	o	o	x	o	o	4
Triglyceride	o	x	o	o	o	4
Aspartate transaminase	o	o	x	x	o	3
Alanine transferase	o	o	o	o	x	4
Hemoglobin	o	x	x	x	o	2
Hematocrit	x	x	x	o	x	1
Blood urea nitrogen	o	o	x	o	o	4
Blood creatinine	o	x	x	x	o	2
White blood cell	o	o	o	o	o	5
Red blood cell	o	x	x	x	o	2
Thrombocyte	o	o	o	o	o	5
Urine pH	o	x	x	o	o	3
Sodium nitrite	o	x	x	x	o	2
Urine specific gravity	x	x	x	x	x	0
Urine protein	x	x	x	x	o	1
Urine glucose	o	o	o	x	o	4
Urine ketone	o	x	x	o	o	3
Urine bilirubin	o	x	x	x	o	2
Urine blood	o	x	x	o	o	3
Urobilinogen	o	x	x	x	o	2
Urine sodium	o	x	x	o	o	3
Dietary survey						
Food intake	o	o	o	o	o	5
Energy intake	x	x	o	o	o	3
Water intake	o	o	o	o	o	5
Protein intake	x	o	o	x	o	3
Fat intake	x	o	x	o	o	3
Carbohydrate intake	x	x	x	o	o	2
Calcium intake	o	o	o	o	o	5
Phosphorus intake	o	o	x	x	o	3
Iron intake	o	x	o	o	o	4
Sodium intake	o	x	o	o	o	4
Potassium intake	o	o	x	o	o	4

(*Continued to the next page*)

Table 2. Continued

Variables	Logistic regression	Random forest	Adaptive boosting	Gradient boosting	Support vector machine	Score
β-carotene intake	o	x	o	o	o	4
Retinol intake	o	o	x	o	x	3
Thiamine intake	o	o	x	o	o	4
Riboflavin intake	o	x	x	o	o	3
Niacin intake	o	o	o	o	o	5
Vitamin-C intake	o	o	x	o	o	4
Health interview						
Age (30's year)	o	x	x	x	o	2
Age (40's year)	o	o	o	o	o	5
Age (50's year)	o	o	o	o	o	5
Age (60's year)	x	o	x	x	o	2
Age (≥70 year)	o	o	o	o	o	5
Living in a house	o	x	x	x	x	1
Living in apartment	o	x	x	x	x	1
Household income (low)	o	x	x	x	x	1
Household income (middle-low)	o	x	x	x	x	1
Household income (middle-high)	o	x	x	x	o	2
Household income (high)	o	x	x	x	o	2
Education (elementary school and below)	o	o	o	o	x	4
Education (middle school)	x	x	o	o	x	2
Education (high school)	o	o	x	o	o	4
Education (university graduate or higher)	o	x	x	x	o	2
Job (manager, expert)	o	x	x	x	o	2
Job (office workers)	o	x	x	o	x	2
Job (service, sales)	o	o	x	x	o	3
Job (forestry and fishery workers)	x	x	x	o	o	2
Job (technician, device/machine operator)	o	x	x	x	o	2
Job (simple labor workers)	o	x	x	o	o	3
Job (unemployed)	o	x	o	o	o	4

Table 3. Accuracy of prediction models in various machine learning technique

Prediction model	Area under the curve	
	Male	Female
Machine learning		
Logistic regression	0.85	0.88
Random forest	0.83	0.88
Gradient boosting	0.83	0.88
Adaptive boosting	0.84	0.87
Support vector machine	0.85	0.88

showed that intake of calcium and niacin was associated with the risk of osteoporosis. Therefore, our study also suggests that careful attention to the intake of these nutritional factors will help prevent osteoporosis.

Shim et al. [27] performed osteoporosis risk prediction in postmenopausal women using 2010 and 2011 KNHNES data. A total of 7 ML models were used in their study, including the 5 used in this study. The AUC of osteoporosis prediction of their ML model varied from 0.685 to 0.743. This is lower than the AUC 0.83 to 0.88 reported in this study. On the other hand, Kwon et al. [25] performed osteoporosis risk prediction in postmenopausal women aged 40 to 69 years using the KNHNES data from 2008 to 2011, and used 3 techniques: RF, AdaBoost, and GB machine (GBM). Their reported AUCs were 0.919 for RF, 0.921 for AdaBoost, and 0.908 for GBM, which were higher than the measured values of the present study. Although there are some dif-

https://doi.org/10.11005/jbm.2023.30.3.263

ferences, the difference in measured AUC despite the similarity of input data and ML models is considered to be due to differences in the input data processing and selected variables. Kwon et al. [25] measured the accuracy in 3 models with different input data in ML, and different accuracy was measured for each. This means that the AUC of ML prediction can differ depending on the input data type. In particular, they argued that a small number of features is more effective than using all features in terms of model efficiency and stability in ML model. Therefore, it seems that the input of a lot of data does not necessarily guarantee the prediction of high osteoporosis. Kwon et al. [25] used medical domain knowledge alongside feature importance and recursive feature elimination techniques in the preprocessing of input data, and reported that this was a factor that improved ML training and obtained high AUC. Therefore, we believe that although the ML technique is important, the preprocessing of input data and the feature selection process are more important factors for the accuracy of the osteoporosis prediction model and will be the most important part for the success of the ML model to be developed.

This study had a few limitations. First, because the present study used the ML prediction model using cross-sectional data, it could not be confirmed whether similar results could be obtained with longitudinal data. Second, the present study could not analyze whether and how much the risk of osteoporosis could be reduced if risk factors were corrected in patients. It is considered that further research on this is needed in the future.

The clinical significance of this study is as follows. First, it is possible to provide customized algorithms for each country and race, which have been pointed out as limitations of existing fracture prediction models including fracture risk assessment tool. Through this, it will be able to serve as an important national database that enables customized osteoporosis treatment for each patient. Second, it is possible to present a customized management model necessary for osteoporosis risk management. The database, which reflects chronic diseases, nutrition, and exercise status, is representative of the population with lifestyles of a specific age group. Data on osteoporosis patients that match specific patients can be used to develop customized osteoporosis management models and programs.

In conclusion, ML performed a feature selection of osteoporosis considering hidden differences between men and women. The present study considers the preprocessing of input data and the feature selection process as well as the ML technique to be important factors for the accuracy of the osteoporosis prediction model.

DECLARATIONS

Funding

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant no. HI22C0494).

Ethics approval and consent to participate

The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Institutional Review Board of the Korea Centers for Disease Control and Prevention (Approval no. 2008–04EXP-01-C, 2009–01CON-03-C, 2010–02CON-21-C, and 2011–02CON-06-C). Informed consent was obtained from each participant when the 2008, 2009, 2010, and 2011 KNHANESs were conducted.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

ORCID

Yonghan Cha *https://orcid.org/0000-0002-7616-6694*
Sung Hyo Seo *https://orcid.org/0000-0003-3327-5425*
Jung-Taek Kim *https://orcid.org/0000-0003-4243-5793*
Jin-Woo Kim *https://orcid.org/0000-0003-0186-5834*
Sang-Yeob Lee *https://orcid.org/0000-0001-8024-1135*
Jun-Il Yoo *https://orcid.org/0000-0002-3575-4123*

REFERENCES

1. Cha YH, Ha YC, Lim JY, et al. Introduction of the cost-effectiveness studies of fracture liaison service in other countries. J Bone Metab 2020;27:79-83. https://doi.org/10.11005/jbm.2020.27.2.79.
2. Cha YH, Ha YC, Lim JY. Establishment of fracture liaison service in Korea: where is it stand and where Is it going? J Bone Metab 2019;26:207-11. https://doi.org/10.11005/jbm.2019.26.4.207.

3. LeBoff MS, Greenspan SL, Insogna KL, et al. The clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 2022;33:2049-102. https://doi.org/10.1007/s00198-021-05900-y.
4. Cheng CT, Wang Y, Chen HW, et al. A scalable physician-level deep learning algorithm detects universal trauma on pelvic radiographs. Nat Commun 2021;12:1066. https://doi.org/10.1038/s41467-021-21311-3.
5. Ok HS, Kim WS, Ha YC, et al. Alarm services as a useful tool for diagnosis and management of osteoporosis in patients with hip fractures: a prospective observational multicenter study. J Bone Metab 2020;27:65-70. https://doi.org/10.11005/jbm.2020.27.1.65.
6. Goecks J, Jalili V, Heiser LM, et al. How machine learning will transform biomedicine. Cell 2020;181:92-101. https://doi.org/10.1016/j.cell.2020.03.022.
7. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artif Intell Health 2020:25-60. https://doi.org/10.1016/B978-0-12-818438-7.00002-2.
8. Basheer IA, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Methods 2000;43:3-31. https://doi.org/10.1016/s0167-7012(00)00201-3.
9. Patel JL, Goyal RK. Applications of artificial neural networks in medical science. Curr Clin Pharmacol 2007;2:217-26. https://doi.org/10.2174/157488407781668811.
10. Pandit A, Radstake T. Machine learning in rheumatology approaches the clinic. Nat Rev Rheumatol 2020;16:69-70. https://doi.org/10.1038/s41584-019-0361-0.
11. Kokkotis C, Moustakidis S, Papageorgiou E, et al. Machine learning in knee osteoarthritis: a review. Osteoarthr Cartil Open 2020;2:100069. https://doi.org/10.1016/j.ocarto.2020.100069.
12. Jamshidi A, Pelletier JP, Martel-Pelletier J. Machine-learning-based patient-specific prediction models for knee osteoarthritis. Nat Rev Rheumatol 2019;15:49-60. https://doi.org/10.1038/s41584-018-0130-5.
13. de Vries BCS, Hegeman JH, Nijmeijer W, et al. Comparing three machine learning approaches to design a risk assessment tool for future fractures: predicting a subsequent major osteoporotic fracture in fracture patients with osteopenia and osteoporosis. Osteoporos Int 2021;32:437-49. https://doi.org/10.1007/s00198-020-05735-z.
14. Kweon S, Kim Y, Jang MJ, et al. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). Int J Epidemiol 2014;43:69-77. https://doi.org/10.1093/ije/dyt228.
15. Maurel DB, Boisseau N, Benhamou CL, et al. Alcohol and bone: review of dose effects and mechanisms. Osteoporos Int 2012;23:1-16. https://doi.org/10.1007/s00198-011-1787-7.
16. Cheraghi Z, Doosti-Irani A, Almasi-Hashiani A, et al. The effect of alcohol on osteoporosis: a systematic review and meta-analysis. Drug Alcohol Depend 2019;197:197-202. https://doi.org/10.1016/j.drugalcdep.2019.01.025.
17. Bakker MH, Vissink A, Spoorenberg SLW, et al. Are edentulousness, oral health problems and poor health-related quality of life associated with malnutrition in community-dwelling elderly (aged 75 years and over)? A cross-sectional study. Nutrients 2018;10:1965. https://doi.org/10.3390/nu10121965.
18. Bonjour JP. Protein intake and bone health. Int J Vitam Nutr Res 2011;81:134-42. https://doi.org/10.1024/0300-9831/a000063.
19. Huang W, Xiao Y, Wang H, et al. Association of geriatric nutritional risk index with the risk of osteoporosis in the elderly population in the NHANES. Front Endocrinol (Lausanne) 2022;13:965487. https://doi.org/10.3389/fendo.2022.965487.
20. Rinonapoli G, Ruggiero C, Meccariello L, et al. Osteoporosis in men: a review of an underestimated bone condition. Int J Mol Sci 2021;22:2105. https://doi.org/10.3390/ijms22042105.
21. Kelsey JL. Risk factors for osteoporosis and associated fractures. Public Health Rep 1989;104 Suppl:14-20.
22. Li H, Wallin M, Barregard L, et al. Smoking-induced risk of osteoporosis is partly mediated by cadmium from tobacco smoke: the MrOS Sweden study. J Bone Miner Res 2020;35:1424-9. https://doi.org/10.1002/jbmr.4014.
23. Vilaca T, Eastell R, Schini M. Osteoporosis in men. Lancet Diabetes Endocrinol 2022;10:273-83. https://doi.org/10.1016/s2213-8587(22)00012-2.
24. Roth JA, Battegay M, Juchler F, et al. Introduction to machine learning in digital healthcare epidemiology. Infect Control Hosp Epidemiol 2018;39:1457-62. https://doi.org/10.1017/ice.2018.265.
25. Kwon Y, Lee J, Park JH, et al. Osteoporosis pre-screening using ensemble machine learning in postmenopausal Korean women. Healthcare (Basel) 2022;10:1107. https://doi.org/10.3390/healthcare10061107.

https://doi.org/10.11005/jbm.2023.30.3.263

26. Park HM, Heo J, Park Y. Calcium from plant sources is beneficial to lowering the risk of osteoporosis in postmenopausal Korean women. Nutr Res 2011;31:27-32. https://doi.org/10.1016/j.nutres.2010.12.005.

27. Shim JG, Kim DW, Ryu KH, et al. Application of machine learning approaches for osteoporosis risk prediction in postmenopausal women. Arch Osteoporos 2020;15:169. https://doi.org/10.1007/s11657-020-00802-8.