

A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data

Hassan Zohrevand¹, Mahdi Ghorbani², Sajjad Yazdanpanah², Fahimeh Faghihi Moghadam², Ahmad Shalbaf², Mohammad Reza Asgari², Hasan Langari³, Faraj Tabei⁴

¹Student Research Committee, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

²Biomedical Engineering and Medical Physics Department, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

³Radiology Department, Taleghani Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁴Nuclear Medicine Department, Taleghani Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Corresponding Author:

Dr. Mahdi Ghorbani

Email: ghorbanim@sbmu.ac.ir

Dr. Ahmad Shalbaf

Email: shalbaf@sbmu.ac.ir

Address: Biomedical Engineering and Medical Physics Department, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Velenjak, Iran.

Running title: Radiomics prediction-osteoporosis

Article History:

Received: 29 March 2026

Revised: 06 August 2026

Accepted: 07 August 2026

Published Online: 17 August 2026

ABSTRACT

Introduction: Osteoporosis is a major public health problem worldwide. Patients with osteoporosis are at higher risk of bone fractures, pain, and immobilization. The aim of the present study was to predict osteopenia/osteoporosis from demographic, clinical and laboratory data using machine learning (ML) algorithms.

Methods: Demographic, clinical, and laboratory data from 239 cases (101 normal and 138 with osteopenia/osteoporosis) were collected. Thirteen features were included: age, sex, weight, height, smoking status, glucocorticoids use, chemotherapy, insulin use, levothyroxine usage, history of bone fracture, premature menopause, calcium usage, and vitamin D usage. The Python programming language was used for preprocessing and modeling. Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), MultiLayer Perceptron (MLP), and eXtreme Gradient Boosting (XGBoost) models were trained and evaluated. Accuracy, precision, recall, F1-score, and Area under the curve (AUC) were calculated.

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

Results: The XGBoost model showed better performance in terms of accuracy, precision and AUC. For the test group, the accuracy, precision, and AUC achieved by the XGBoost model were 0.68, 0.71, and 0.73, respectively.

Conclusion: The results of this study indicate that using the 13 selected features, it is possible to perform an initial prediction to distinguish between normal individuals and those with osteopenia or osteoporosis. Among the evaluated features, sex, age, and smoking had the highest importance. The proposed model may serve as a low-cost preliminary risk prediction model for identifying individuals at increased risk of low bone mineral density, particularly in settings with limited access to DEXA scanning.

Keywords: Machine learning; Osteopenia; Osteoporosis; Demographic data

INTRODUCTION

Osteoporosis is a skeletal disorder characterized by reduced bone strength, leading to an increased risk of fractures, particularly in older adults [1]. In the United States, 1.5 million cases of osteoporosis-related fractures are observed each year [2]. Unlike many chronic diseases that have multiple symptoms, osteoporosis is asymptomatic until a fracture occurs [3]. The most important uncontrollable factors of osteoporosis are aging, being female, and decreasing estrogen levels after menopause [4]. Osteoporosis is typically diagnosed using bone densitometry [5]. Osteoporosis management strategies are diverse, including hormone therapy, lifestyle modifications (exercise and diet), and the use of pharmacologic agents [6].

Currently, dual-energy X-ray absorptiometry (DEXA) is considered the gold standard for measuring bone mineral density (BMD) [7]. Using this, X-rays from two sources with different energies are emitted to the bone, and the amount of radiation that has passed or absorbed is measured [8]. The output values of these two rays are measured by a sensor in the DEXA device. The denser the bone, more radiation is absorbed, and as a result, less output radiation reaches the device's sensor [9]. Although DEXA is considered the gold standard for measuring bone mineral density, access to DEXA scanners may be limited in certain regions because of equipment costs, availability, and the need for specialized facilities. Therefore, alternative prediction models that use routinely available along with clinical information may help identify individuals who would benefit from further diagnostic evaluation [10]. Additionally, the people which are undergone a DEXA examination are exposed to extra ionizing radiation [11, 12].

In addition to BMD measurements, several clinical tools have been developed to estimate osteoporosis and fracture risk. One of the most widely used approaches is the Fracture Risk Assessment Tool (FRAX), which estimates the 10-year probability of major osteoporotic fractures using demographic and clinical risk factors, with or without bone mineral density measurements. While FRAX has demonstrated substantial clinical utility, its underlying statistical framework may not fully capture complex nonlinear relationships among risk factors. Machine learning (ML) approaches may provide an alternative strategy by modeling interactions among multiple variables and potentially improving risk stratification [13]. By decoding information from medical images, both ML and deep learning (DL) provide predictive and prognostic models for disease diagnosis and treatment. It is expected that ML will become an important component in the diagnosis and treatment of diseases in the near future [14-19].

There are different studies on determination of bone density based on computed tomography (CT) or radiography images using ML or DL methods [20-23]. In a study by Yasaka et al. [20], bone mineral was predicted from CT images using DL technique. Hu et al. [21] extracted preoperative

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

vertebral quantitative CT radiomic features and developed models to predict changes in BMD after sleeve gastrectomy. Kalavakonda et al. [22] automated Singh Index identification from hip radiographs using machine learning algorithms. Sato et al. [23] used DL for prediction of bone mineral density and *T*-score from chest X-rays.

There are also other studies on diagnosis of osteoporosis based on either MRI or CT images using ML or DL methods [24-30]. In a study conducted by Küçükçiloğlu et al. [24], CT and MRI images of lumbar vertebrae were used in to diagnose osteoporosis. In a study by Jang et al. [25], radiographic images of the hip joint and the use of artificial intelligence algorithms were used to diagnose osteoporosis. Klontzas et al. [26] used radiomics and ML for differentiation of transient osteoporosis from avascular necrosis of hip. He et al. [27] used radiomics and lumbar MRI images for detection of osteoporosis. Zhao et al. [28] applied an automated radiomics screening pipeline for osteoporosis and abnormal bone density detection. Nam et al. [29] utilized lumbar CT images and ML model for prediction of osteoporotic spine. Xie et al. [30] developed a ML-derived radiomics model for diagnosis of osteoporosis and osteopenia from CT images.

While the ML and DL methods reported in previous studies are promising they are in some extent complex and require access to patients images and often involves exposure to ionizing radiation. In better words, Previous ML studies have primarily focused on imaging-based approaches using CT, MRI, radiographs features. Although these methods often achieve high predictive performance, they require access to imaging data and specialized image-processing workflows. Comparatively fewer studies have investigated whether osteoporosis can be predicted using only routinely available demographic and clinical variables. Therefore, there remains a need for simplified prediction models that can be implemented in primary care settings without the need for imaging data.

Huang et al. [31] introduced a simple and user-friendly ML model for osteoporosis detection in healthy populations. Their model utilized variables such as age, height, weight, and body mass index (BMI). The key advantage of the study was the prediction of osteoporosis using general demographic data, however, its applicability was limited to individuals older than 50 years. To the best of our knowledge, few studies have focused on detecting osteoporosis using routine general data rather than imaging modalities, which cause complexity and radiation burden to the patients. The objective of this study was to develop and evaluate machine learning models to distinguish normal individual from those with osteopenia/osteoporosis using routinely available demographic, clinical, and laboratory variables and to identify the most influential predictors associated with low bone mineral density.

METHODS

Study population

A total of 239 cases (not necessarily consecutive) were included in this study, comprising 101 normal, 104 osteopenia and 34 osteoporosis. For modeling purposes, it was defined as a binary classification task, considering two classes: normal versus osteopenia/osteoporosis. In other words, the data were structured so that the algorithm only recognizes two classes: healthy subjects labeled as 0 ($n=101$), and patients labeled as 1 ($n=138$), which combines both osteopenia and osteoporosis. Ethical guidelines of our University Hospital were followed, and no additional intervention was applied to the patients. This study utilized original, de-identified patient data collected at Taleghani Hospital (Tehran, Iran) and Akhtar Orthopedic Hospital (Tehran, Iran). Patient confidentiality was strictly maintained in accordance with ethical guidelines of Shahid Beheshti University of Medical Sciences. The dataset has not been previously published and

maybe made available upon official request from the corresponding author, subject to approval by the relevant ethics committee. The standards for reporting diagnostic accuracy (STARD) guideline was not followed.

Demographic and clinical data

Thirteen demographic, clinical and laboratory features were recorded for each participant, including age, sex, weight, height, smoking, glucocorticoids usage, chemotherapy, using insulin, levothyroxine usage, bone fracture history, premature menopause, calcium usage and vitamin D usage. These data were obtained from standardized questionnaire forms completed by the participants. Feature selection was based on established osteoporosis risk factors reported in the literature and availability of variables in routine clinical practice. There are also other features which are related to osteopenia/osteoporosis (physical activity, dietary calcium intake, hormonal status) which were not recorded in the clinics and could not be selected. The questionnaire forms have been completed by the Nuclear Medicine Department of Taleghani Hospital (Tehran, Iran) and DEXA Imaging Department of Akhtar Orthopedic Hospital (Tehran, Iran). The demographic data were extracted and recorded based on these forms. Continuous features were analyzed using their numerical values, while categorical variables were encoded as binary numbers (0 or 1).

Bone densitometry

Bone mineral densitometry was obtained using DEXA, which was considered the gold standard for the diagnosis of osteopenia and osteoporosis in this study. The DEXA reports which have been recorded before in the Nuclear Medicine Department of Taleghani Hospital (Tehran, Iran) and DEXA Imaging Department of Akhtar Orthopedic Hospital (Tehran, Iran) were included in the modeling. The DEXA scans were performed using a BMD QDR Explorer (Hologic, USA) with 100, and 140 kVp voltages setting and a Siemens 2D Fan Beam (Siemens, Germany) with 43, and 70 kVp setting at Taleghani and Akhtar hospitals, respectively.

Data preprocessing

The percentage of missing values varied among features. Six variables including age, weight, height, levothyroxine, bone fracture history and premature menopause had only one missing value (0.42%), whereas calcium usage and vitamin D usage variables had 49 (20.5%) and 47 (19.7%) missing values, respectively. Missing values were handled using median imputation, which was selected because of its robustness to outliers and suitability for relatively small datasets.

Data preprocessing and analysis were implemented in Python (Version 3.12.3, Python Software Foundation). All features and labels were defined as variables in the software, and missing values were filled using median imputation. The dataset was randomly divided into training and test sets using the scikit-learn library in Python, including 75% of the data in the training set (179 cases) and 25% of the data in the test set (60 cases). The samples were randomly assigned to the test set while preserving the proportions of healthy and patient classes. In other words, for the osteopenia/osteoporosis breakdown within the patient class, but it was not tried to keep it consistent with the original population for clinical diversity. In the overall dataset, healthy cases The data were split using a 75/25 ratio (75% for training and the remaining 25% for testing). Since different features had different scales, feature standardization was applied using the StandardScaler function from scikit-learn in the Python software environment, fitted on the training data [32]. After the standardization on training set, the standard parameters were utilized for the test group. All features were standardized to have zero mean and unit standard deviation.

Following the initial description of the dataset, the subsequent step involved data preprocessing to prepare the dataset for model training. Several key steps were performed to ensure data quality and suitability for analysis. First, features recorded as text format turned into numerical variables. Second, missing values were addressed using median imputation. Subsequent, the dataset randomly divided into training and test sets. After that, feature scaling using standardization technique were conducted to ensure the comparable range of features.

ML models

Following data preprocessing and standardization, various ML models were evaluated for their performance. The selected models included: Support Vector Machine (SVM) [33], Logistic Regression (LR) [34], Random Forest (RF) [35], Multi Layer Perceptron (MLP), and eXtreme Gradient Boosting (XGBoost) [36]. Logistic Regression was included as a conventional baseline model. Random Forest and XGBoost were selected because of their ability to capture nonlinear relationships, while SVM and MLP represented kernel-based and neural network approaches, respectively [37]. Model training was performed using sikit-learnlibrary in Python. Each model consisted of an initial set of hyperparameters, which were subsequently optimized. The final model training was conducted using these optimized hyperparameters. In the following, more details on optimization of hyperparameters are presented.

Model performance was evaluated using a comprehensive set of metrics: Accuracy, Precision, Recall, F1-Score, Mathews Correlation Coefficient (MCC) and AUC. As it was mentioned before, the classification task included two classes: normal and osteopenia/osteoporosis.

Model hyperparameter optimization

From the total of 239 subjects, 179 (75%) were allocated to the training set and 60 (25%) to the test set. Hyperparameter optimization and model training were performed on the training set, employing cross-validation techniques. Hyperparameter optimization were conducted using Grid Search Cross-Validation (GridSearchCV function in sikit-learn) with 10-fold cross validation. The optimal hyperparameters for each model were selected based on their performance on the validation folds within the training set.

For the SVM model, three hyperparameters were evaluated, the first and second hyperparameters had three different states, the third one had five different states. Totally 45 different sets were evaluated for SVM to determine that which state is better. 10-fold cross-validation technique was used for each set-up in the SVM algorithm. The optimized set-up was determined using the accuracy criteria. Mean accuracy between 45 set-ups were compared.

For LR model three hyperparameters were considered. The first hyperparameter had three states, the second one had four states, and the third one had three states. Totally 36 states were examined and the optimum was determined. For RF algorithm seven hyperparameters were evaluated and totally 288 states were evaluated to find the optimum set of hyperparameters. For MLP model 360 different states were evaluated to find the optimum one. XGBoost was evaluated based on 729 states. The other procedures for evaluation of LR, RF, MLP, and XGBoost were the same as described before for the SVM model. The list of hyperparameters for SVM, LR, RF, MLP, and XGBoost are presented in Table 1.

Model training and performance evaluation

Accuracy, precision, recall, F1-score, and area under curve (AUC) of the receiver-operating characteristic (ROC) curve were those criteria which were utilized for evaluation of the models.

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

The selected optimized hyperparameters from cross-validation in the training step were used to retrain models with the total training data and evaluated for the test data (the remaining 20% of the data). The best model was selected based on these criteria for the test set. The numbers of true positive, false positive, true negative and false negative was determined and used for calculation of these quantities.

Feature importance analysis was conducted for the best-performing model to identify the most effective predictors of osteopenia/osteoporosis.

RESULTS

Study population characteristics

As mentioned before, the demographic, clinical and laboratory data of the 239 participants were recorded. Table 2 encompasses thirteen features, providing descriptive statistical measures including mean and standard deviation for continuous variables, whereas the number of positive and negative samples for categorical features. These thirteen features of 239 samples form the foundation for subsequent analyses and the development of predictive models in this study. The rationale for selection of these features was that they were normally recorded in the clinics and they are relevant to osteopenia/osteoporosis.

These thirteen features of 239 samples form the foundation for subsequent analyses and the development of predictive models in this study. Among the 239 participants, 101 (42.3%) were classified as normal, 104 (43.5%) as osteopenic, and 34 (14.2%) as osteoporotic according to DEXA measurements. The study population was predominantly female (84.9%), which is consistent with the higher prevalence of osteoporosis among women. The mean age of participants was 57.29 ± 11.60 years, reflecting a population at increased risk of bone mineral density loss.

Median imputation successfully addressed missing values without excluding subjects from the analysis. Standardization ensured that all continuous variables contributed equally to model training and prevented features with larger numerical ranges from dominating the learning process. These preprocessing steps improved data consistency and facilitated fair comparison among machine learning algorithms.

Model performance evaluation

Following the data preprocessing stage, the five selected ML models were trained and evaluated. As detailed in method section, a 10-fold cross-validation approach was employed to optimize the best set of hyperparameters. The best set of hyperparameters in each model determined based on mean accuracy on 10 validation folds. The performance metrics of classifiers for cross-validation technique are presented in Table 3. Subsequently, the finalized models, selected based on their cross-validation performance, were applied to the independent test set. This test data provides an unbiased assessment of how the models are expected to perform on unseen data. Table 3 also shows the performance of each model on test data in tabular format, whereas Figure 1 demonstrates the results of all classifiers on test set using barplot to simply compare all classifiers in terms of each metric.

Across all evaluated models, XGBoost achieved the highest performance on the independent test set. Compared with Logistic Regression, XGBoost improved AUC from 0.65 to 0.73 and increased recall from 0.68 to 0.77. The XGBoost model also yielded a specificity of 0.56. These findings suggest that nonlinear ensemble-learning approaches may better capture complex relationships among demographic and clinical risk factors associated with reduced bone mineral density.

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

ROC curve for the trained models compared to the random classifier is presented in Figure 2. Results on test data shows that XGBoost model outperforms other classifiers. The accuracy, precision and recall of this model achieved 0.68, 0.71 and 0.77 respectively. The F1-score, MCC and AUC evaluated on test data was 0.74, 0.34 and 0.73.

Although all machine learning models performed better than random classification, the ROC curve of XGBoost consistently remained above those of the other algorithms throughout most decision thresholds, indicating superior discriminative capability. The observed AUC of 0.73 suggests moderate predictive performance when using demographic and clinical variables alone.

Importance of features

The importance of all thirteen features used in the XGBoost model was evaluated to identify which predictors contributed most to the classification performance. The importance analysis generated non-zero importance scores for 11 out of 13 features. A ranked horizontal bar plot summarizing these importance values is presented in Figure 3.

Sex and age demonstrated the highest feature importance scores, followed by smoking status. Several additional variables, including calcium supplementation, insulin use, and thyroid hormone replacement therapy, also contributed to model performance. These findings indicate that osteoporosis risk prediction benefits from a combination of demographic, behavioral, and clinical factors rather than reliance on a single predictor.

False-negative (FN) error pattern analysis

To further investigate the performance of the XGBoost model on the test set, a detailed analysis of false-negative cases was conducted, as presented in Table 4. This table shows the number and percentage of patients within the test group who were real positive cases, and the number and percentage of patients that misclassified to healthy (false-negative cases). The results shows that the frequency of osteopenia is much more than osteoporosis (26 versus 9). Eight subjects out of 35 (total of real patients) classified as healthy (7 osteopenia out of 26 and 1 osteoporosis out of 9). This means that 87.50% and 12.50% of FN cases belongs to osteopenia and osteoporosis, respectively. If we consider the ratio of real patients, 35 samples (26: osteopenia, 9: osteoporosis), the expected percentage of FN is 74.29% and 25.71% for osteopenia and osteoporosis. In other words, the majority of false-negative predictions occurred among osteopenic individuals. Specifically, 7 of the 8 false-negative cases belonged to the osteopenia group, whereas only one false-negative case was observed among patients with osteoporosis. This observation suggests that osteopenia may represent an intermediate biological state that shares characteristics with both normal bone density and osteoporosis, making accurate classification more challenging. In contrast, osteoporosis is typically associated with more pronounced clinical and physiological alterations that are easier for machine learning models to identify. Confusion matrix for the patients and healthy samples which were calculated using XGBoost model for the test set is presented in Table 5.

DISCUSSION

In this study, ML models were developed and evaluated for the prediction of osteopenia and osteoporosis using routinely available demographic, clinical, and laboratory variables. Among the evaluated algorithms, XGBoost demonstrated the best overall performance, achieving an accuracy of 0.68, precision of 0.71, recall of 0.77, F1-score of 0.74, and AUC of 0.73 on the independent

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

test dataset. These findings suggest that meaningful prediction of low bone mineral density can be achieved without the use of imaging-derived features.

The superior performance of XGBoost is consistent with previous studies demonstrating the effectiveness of gradient boosting methods in healthcare prediction tasks. Unlike traditional statistical approaches, XGBoost can model nonlinear interactions among variables while simultaneously reducing overfitting through regularization techniques. These characteristics likely contributed to its improved predictive capability compared with Logistic Regression, Random Forest, SVM, and MLP models. This trend is in agreement with the AUC data presented in Table 3.

Based on the results in Table 4, analysis of false-negative cases demonstrated that most classification errors occurred among osteopenic patients rather than osteoporotic patients. This finding is clinically reasonable because osteopenia represents an intermediate condition between normal bone density and osteoporosis. Consequently, the demographic and clinical characteristics of osteopenic individuals may overlap substantially with those of normal individuals, making classification more difficult. In contrast, osteoporosis is generally associated with more advanced biological changes that are easier to distinguish from normal bone status.

Based on the data presented in Figure 3, feature importance analysis identified sex, age, and smoking status as the strongest predictors of osteopenia and osteoporosis. These findings are biologically plausible and consistent with previous epidemiological studies. Sex emerged as the most influential predictor, likely reflecting the increased prevalence of osteoporosis among women following menopause. Estrogen deficiency accelerates bone resorption and contributes substantially to bone mineral loss. Age was also highly important, which is expected given the progressive decline in bone formation and remodeling capacity that accompanies aging. Smoking ranked among the top predictors, consistent with evidence that tobacco exposure negatively affects osteoblast activity, calcium metabolism, and hormonal regulation, thereby increasing osteoporosis risk.

The performance achieved in the present study was lower than that reported by many imaging-based machine learning studies. Previous investigations using CT, MRI, radiomics, and deep learning approaches frequently report AUC values exceeding 0.85. This difference is expected because imaging features directly capture structural information related to bone quality and mineral density. Nevertheless, the present study demonstrates that acceptable predictive performance can be achieved using readily available demographic and clinical variables alone, eliminating the need for image acquisition and image-processing workflows.

There are different studies on determination of bone density using ML or DL methods [20-31]. A list of different studies on determination of bone density or detection of osteoporosis with radiomics/ML or DL is presented in Table 6. As it is obvious from this table the AUC performance quantities in this study is lower than the other studies. The reason for this trend is that in the other studies image features were used as alone features or in combination with the clinical data. It is clear that by the use of image features the performance will be increased. There is also another study by Huang et al. [31] in which only general population data were used for modeling. The AUC in that study is relatively the same as that in our study. The advantage of the other studies, in which imaging features were used, is higher AUC, accuracy, and precision in detection of osteoporosis. On the other hand, the advantage of the present study is the lack of the need for imaging tools. In other words, based on the results of the present study, by the use of a simple questionnaire and general data (including sex, age, smoking, etc.) and the presented optimum model, it is feasible to predict osteopenia/osteoporosis with a relatively good performance. Another

advantage of this study over other studies is that in most of the other studies only osteoporosis abnormality was predicted, while in this study osteopenia was also predicted.

From a clinical perspective, the proposed model is not intended to replace DEXA-based diagnosis. Rather, it may serve as a preliminary prediction model for identifying individuals at elevated risk of osteopenia or osteoporosis. Such a strategy may be particularly valuable in primary-care settings, community prediction programs, and regions with limited access to bone densitometry services. Individuals identified as high-risk could subsequently be referred for confirmatory DEXA evaluation.

There is a tool for calculation of ten-year probability of fracture with or without BMD (FRAX [38]) which can be found in the following webpage: <https://www.fraxplus.org/calculation-tool>. It is capable of online calculation of fracture risk by introducing age, sex, weight, height, previous fracture, parent fractured hip, current smoking, glucocorticoids, rheumatoid arthritis, secondary osteoporosis, alcohol 3 or more units per day, and or femoral neck BMD. There are similarities and differences between this study and the FRAX. The similarity is that both are using relatively the same features while their scopes are different (one predicts being normal or osteopenia/osteoporosis and the other predicts bone fracture risk). Using ML for prediction of bone fracture risk and comparison with the values which can be calculated by FRAX can be a subject for a future study. Furthermore, superiority of this model over osteoporosis self-assessment tool (OST), osteoporosis risk assessment instrument (ORAI), simple calculated osteoporosis risk estimation (SCORE), osteoporosis index of risk (OSIRIS), QFracture, and Garvan fracture risk calculator need to be evaluated in the future studies.

Several limitations should be considered when interpreting the findings of this study. First, the sample size and the predictive performances of the models were relatively modest, which may limit model generalizability. Second, the study population was derived from a limited geographic region without performing external validation using independent cohorts. Third, several potentially important predictors of osteoporosis, including body mass index, physical activity level, dietary calcium intake, family history of osteoporosis, alcohol consumption, and history of diseases related to secondary osteoporosis (examples: hypogonadism, type 1 diabetes, untreated hyperthyroidism) were not available. Fourth, osteopenia and osteoporosis were combined into a single abnormal category to improve statistical power, which prevented separate evaluation of the two conditions. Finally, imaging biomarkers derived from CT, MRI, or radiographic studies were not incorporated into the models. Future studies involving larger multicenter datasets, external validation cohorts, and additional clinical and imaging variables may further improve the predictive performance.

In this study only the features which are related to normal and osteopenia/osteoporosis cases were introduced, while no software was developed. Therefore, developing a software or application which includes these features and XGBoost model as input and prediction of being normal or osteopenia/osteoporosis as the output will be useful.

Future research should focus on the integration of demographic, clinical, laboratory, and imaging features within unified machine learning frameworks. Additionally, external validation using multicenter cohorts and prospective datasets will be necessary to determine the clinical applicability and generalizability of the proposed model. The incorporation of explainable artificial intelligence techniques may further enhance clinician confidence and facilitate adoption in clinical practice.

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

CONCLUSION

The findings of this study demonstrate that machine learning models trained on routinely available demographic, clinical, and laboratory variables can provide meaningful prediction of osteopenia and osteoporosis. Among the evaluated algorithms, XGBoost achieved the highest overall performance, with an AUC of 0.73 on the independent test dataset. Feature importance analysis identified sex, age, and smoking status as the most influential predictors of low bone mineral density. Although the predictive performance was lower than that reported in many imaging-based studies, the proposed approach offers a simple and accessible prediction strategy that does not require imaging data. Therefore, this model may serve as a preliminary risk assessment tool to identify individuals who would benefit from further evaluation using DEXA and a comprehensive clinical assessment. Developing a software application which incorporate these features and the XGBoost model as input, with prediction of being normal or osteopenia/osteoporosis as the output, has a useful potential for future clinical implications.

Acknowledgements

The authors acknowledge Shahid Beheshti University of Medical Sciences for financial support of this study. We are also thankful to the staff of Nuclear Medicine and DEXA imaging departments of Taleghani and Akhtar Hospitals (Tehran, Iran) for providing patient data and DEXA information.

REFERENCES

1. Clynes MA, Harvey NC, Curtis EM, Fuggle NR, Dennison EM, Cooper C. The epidemiology of osteoporosis. *Br Med Bull.* 2020 May;133(1):105-17.
2. Ford MA, Bass M, Zhao Y, Bai JB, Zhao Y. Osteoporosis knowledge, self-efficacy, and beliefs among college students in the USA and China. *J Osteoporos.* 2011;2011: 729219.
3. Kanis JA, Melton LJ 3rd, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. *J Bone Miner Res.* 1994 Aug;9(8):1137-41.
4. Melton LJ. Epidemiology worldwide. *Endocrinol Metab Clin North Am.* 2003 Mar;32(1):1-13.
5. Kasai S, Li F, Shiraishi J, Li Q, Doi K. Computerized detection of vertebral compression fractures on lateral chest radiographs: preliminary results with a tool for early detection of osteoporosis. *Med Phys.* 2006 Dec;33(12):4664-74.
6. Gregson CL, Armstrong DJ, Bowden J, Cooper C, Edwards J, Gittoes NJL, Harvey N, Kanis J, Leyland S, Low R, McCloskey E, Moss K, Parker J, Paskins Z, Poole K, Reid DM, Stone M, Thomson J, Vine N, Compston J. UK clinical guideline for the prevention and treatment of osteoporosis. *Arch Osteoporos.* 2022 Apr;17(1):58.
7. Park YS. Diagnosis and treatment of osteoporosis. *J Korean Med Assoc.* 2012;55(11):1083-1094.
8. Sebro R, De la Garza-Ramos C. Opportunistic screening for osteoporosis and osteopenia from CT scans of the abdomen and pelvis using machine learning. *Eur Radiol.* 2023 Mar;33(3):1812-1823.
9. Sangondimath G, Sen RK, Rehman TF. DEXA and Imaging in Osteoporosis. *Indian J Orthop.* 2023 Dec;57(Suppl 1):82-93.

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

10. Amani F, Tarighatnia A, Amanzadeh M, Asgarzadeh SA, Jalali S, Hamedan M, Nader N. Predicting osteoporosis with various machine learning algorithms using dual energy x-ray absorptiometry: A comparative analysis. *Front Biomed Technol*. 2026. (In press).
11. Ahmad A, Crawford CH 3rd, Glassman SD, Dimar JR 2nd, Gum JL, Carreon LY. Correlation between bone density measurements on CT or MRI versus DEXA scan: A systematic review. *N Am Spine Soc J*. 2023 Mar;14:100204.
12. Harmer H, Radford S, Smail M. Feasibility study of using PCXMC 2.0 to estimate patient dose arising from DEXA scans. *Radiat Prot Dosimetry* 2018 Nov;181(4):418-422.
13. Schini M, Johansson H, Harvey NC, Lorentzon M, Kanis JA, McCloskey EV. An overview of the use of the fracture risk assessment tool (FRAX) in osteoporosis. *J Endocrinol Invest*. 2024 Mar;47(3):501-511.
14. Scapicchio C, Gabelloni M, Barucci A, Cioni D, Saba L, Neri E. A deep look into radiomics. *Radiol Med*. 2021 Oct;126(10):1296-1311.
15. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, Cook G. Introduction to Radiomics. *J Nucl Med*. 2020 Apr;61(4):488-495.
16. Tamai K, Imanishi K, Terakawa M, Uematsu M, Kato M, Toyoda H, Suzuki A, Takahashi S, Yabu A, Sawada Y, Iwamae M, Kobayashi Y, Okamura Y, Taniwaki H, Kinoshita Y, Hoshino M, Tabuchi H, Nakamura H, Terai H. Deep learning algorithm for identifying osteopenia/osteoporosis using cervical radiography. *Sci Rep*. 2025 Jul;15(1):25274.
17. Wang J, He Y, Yan L, Chen S, Zhang K. Predicting osteoporosis and osteopenia by fusing deep transfer learning features and classical radiomics features based on single-source dual-energy CT imaging. *Acad Radiol*. 2024 Oct;31(10):4159-4170.
18. Ghazizadeh M, Amoui M, Deevband MR, Aryana K, Karamian F, Divband A, Mahboubi-Fooladi Z, Montazerabadi A, Jafari E, Zirak A, Tavakoli M. Machine learning based radiomics analysis of SPECT images in predicting neuroendocrine tumors. *Biomed Eng Adv*. 2025;10:100197.
19. Khajetash B, Mahdavi SR, Nikoofar A, Johnson L, Tavakoli M. Ensemble learning approach for prediction of early complications after radiotherapy for head and neck cancer using CT and MRI radiomic features. *Sci Rep*. 2025 Apr 24;15(1):14229.
20. Yasaka K, Akai H, Kunitatsu A, Kiryu S, Abe O. Prediction of bone mineral density from computed tomography: application of deep learning with a convolutional neural network. *Eur Radiol*. 2020 Jun;30(6):3549-3557.
21. Hu D, Ruan J, Liu C, Liang Z, Mu X. Quantitative CT imaging radiomics-based prediction of bone mineral density changes after sleeve gastrectomy. *Obes Surg*. 2025 Aug;35(8):3047-3056.
22. Kalavakonda V, Madhivanan V, Lal A, Rithika S, Subramaniam SK, Sameer M. Unsupervised machine learning for osteoporosis diagnosis using singh index clustering on hip radiographs. *arXiv preprint arXiv* 2024:2411.15253. <https://doi.org/10.48550/arXiv.2411.15253>

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

23. Sato Y, Yamamoto N, Inagaki N, Iesaki Y, Asamoto T, Suzuki T, Takahara S. Deep Learning for bone mineral density and T-score prediction from chest X-rays: A multicenter study. *Biomedicines*. 2022 Sep;10(9):2323.
24. Küçükçiloğlu Y, Şekeroğlu B, Adalı T, Şentürk N. Prediction of osteoporosis using MRI and CT scans with unimodal and multimodal deep-learning models. *Diagn Interv Radiol*. 2024 Jan;30(1):9-20.
25. Jang R, Choi JH, Kim N, Chang JS, Yoon PW, Kim CH. Prediction of osteoporosis from simple hip radiography using deep learning algorithm. *Sci Rep*. 2021 Oct;11(1):19997.
26. Klontzas ME, Manikis GC, Nikiforaki K, Vassalou EE, Spanakis K, Stathis I, Kakkos GA, Matthaïou N, Zibis AH, Marias K, Karantanis AH. Radiomics and Machine Learning Can Differentiate Transient Osteoporosis from Avascular Necrosis of the Hip. *Diagnostics (Basel)*. 2021 Sep;11(9):1686.
27. He L, Liu Z, Liu C, Gao Z, Ren Q, Lei L, Ren J. Radiomics based on lumbar spine magnetic resonance imaging to detect osteoporosis. *Acad Radiol*. 2021 Jun;28(6):e165-e171.
28. Zhao Y, Zhao T, Chen S, Zhang X, Serrano Sosa M, Liu J, Mo X, Chen X, Huang M, Li S, Zhang X, Huang C. Fully automated radiomic screening pipeline for osteoporosis and abnormal bone density with a deep learning-based segmentation using a short lumbar mDixon sequence. *Quant Imaging Med Surg*. 2022 Feb;12(2):1198-1213.
29. Nam KH, Seo I, Kim DH, Lee JI, Choi BK, Han IH. Machine learning model to predict osteoporotic spine with hounsfield units on lumbar computed tomography. *J Korean Neurosurg Soc*. 2019 Jul;62(4):442-449.
30. Xie Q, Chen Y, Hu Y, Zeng F, Wang P, Xu L, Wu J, Li J, Zhu J, Xiang M, Zeng F. Development and validation of a machine learning-derived radiomics model for diagnosis of osteoporosis and osteopenia using quantitative computed tomography. *BMC Med Imaging*. 2022 Aug;22(1):140.
31. Huang WC, Chen IS, Yu HC, Chen CS, Wu FZ, Hsu CL, Wu PC. A simple and user-friendly machine learning model to detect osteoporosis in health examination populations in Southern Taiwan. *Bone Rep*. 2025 Jan;24:101826.
32. Scikit Learn. Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>. Accessed on: 6 October 2026.
33. Talebi A, Bitarafan-Rajabi A, Alizadeh-Asl A, Seilani P, Khajetash B, Hajianfar G, Tavakoli M. Machine learning based radiomics model to predict radiotherapy induced cardiotoxicity in breast cancer. *J Appl Clin Med Phys*. 2025 Apr;26(4):e14614.
34. Li XY, Duan GY, Li L, Gao YY, Sun LW, Zhu S, Diao X, Wang N, Huang H. Logistic regression and machine learning algorithms for the risk prediction of perioperative adverse cardiovascular events in elderly patients. *Aging Med (Milton)*. 2025 Oct 3;8(5):398-411.
35. Rigatti SJ. Random Forest. *J Insur Med*. 2017;47(1):31-39.

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

36. Wang J, Wu H, Wang F, Wang Y, Gu J. A Machine learning model integrating preoperative blood-based indices for early and noninvasive detection of endometrial cancer. J Inflamm Res. 2025 Aug;18:10873-10884.
37. Rahmani AM, Yousefpoor E, Yousefpoor MS, Mehmood Z, Haider A, Hosseinzadeh M, Ali Naqvi R. Machine learning (ML) in medicine: review, applications, and challenges. Mathematics 2021;9:2970.
38. Kanis JA, Harvey NC, Johansson H, Odén A, Leslie WD, McCloskey EV. FRAX Update. J Clin Densitom. 2017 Jul-Sep;20(3):360-367.

Table 1. List of hyperparameters for SVM, LR, RF, MLP, and XGBoost models

Model	Hyperparameter
SVM	C, Kernel, Gamma
LR	C, penalty, solver
RF	max_depth, min_samples_split, min_samples_leaf, max_features, max_leaf_nodes, bootstrap, n_estimators
MLP	hidden_layer_sizes, activation, solver, alpha, learning_rate, max_iter, early_stopping, batch_size
XGBoost	n_estimators, max_depth, learning_rate, subsample, colsample_bytree, gamma

SVM: Support Vector Machine, LR: Logistic Regression, RF: Random Forest, MLP: MultiLayer Perceptron, XGBoost : eXtreme Gradient Boosting

Table 2. Study population characteristics related to the 13 features

Feature	Descriptive Stat
Age	Healthy(Mean=52.15, SD=12.36)/ Sick(Mean=60.64, SD=10.75)
Sex	Healthy(Male=21, Female =80)/ Sick(Male=15, Female =123)
Weight	Healthy(Mean=75.30, SD=14.54)/ Sick(Mean=72.61, SD=11.68)
Height	Healthy(Mean=163.21, SD=20.20)/ Sick(Mean=161.22, SD=7.03)
Smoking	Healthy(Pos=13, Neg =88)/ Sick(Pos=14, Neg =124)
Glucocorticoids usage	Healthy(Pos=10, Neg =91)/ Sick(Pos=8, Neg =130)
Chemotherapy	Healthy(Pos=6, Neg =95)/ Sick(Pos=11, Neg =127)
Insulin usage	Healthy(Pos=11, Neg =90)/ Sick(Pos=6, Neg =132)
Levothyroxine usage	Healthy(Pos=25, Neg =76)/ Sick(Pos=32, Neg =105)
Bone fracture history	Healthy(Pos=27, Neg =74)/ Sick(Pos=29, Neg =108)
Premature menopause	Healthy(Pos=15, Neg =86)/ Sick(Pos=38, Neg =99)
Calciumusage	Healthy(Pos=57, Neg =24)/ Sick(Pos=85, Neg =24)
Vitamin dosage	Healthy(Pos=65, Neg =16)/ Sick(Pos=94, Neg =17)

Table 3. The performance of cross-validation and independent test set for all classifiers

Model	Training group (validation folds)					
	Accuracy	Precision	Recall	F1Score	MCC	AUC
SVM	0.69±0.10	0.71±0.09	0.77±0.14	0.74±0.10	0.37±0.21	0.71±0.11
LR	0.71±0.10	0.77±0.10	0.73±0.12	0.74±0.10	0.43±0.20	0.75±0.12
RF	0.70±0.10	0.74±0.10	0.76±0.12	0.74±0.09	0.40±0.22	0.74±0.12
MLP	0.69±0.09	0.69±0.07	0.83±0.14	0.75±0.09	0.35±0.21	0.68±0.12
XGBoost	0.74±0.05	0.76±0.13	0.82±0.10	0.77±0.06	0.48±0.10	0.75±0.08
Test group						
SVM	0.67(0.55-0.78)	0.69(0.54-0.83)	0.77(0.64-0.90)	0.73(0.61-0.83)	0.30(0.03-0.54)	0.68(0.53-0.82)
LR	0.62(0.50-0.73)	0.67(0.51-0.82)	0.69(0.53-0.83)	0.68(0.54-0.79)	0.21(-0.05-0.44)	0.65(0.49-0.80)
RF	0.65(0.53-0.77)	0.68(0.53-0.83)	0.74(0.60-0.88)	0.71(0.59-0.82)	0.27(0.02-0.52)	0.72(0.58-0.85)
MLP	0.63(0.52-0.75)	0.65(0.51-0.79)	0.80(0.67-0.93)	0.72(0.59-0.81)	0.22(-0.05-0.45)	0.68(0.54-0.81)
XGBoost	0.68(0.55-0.80)	0.71(0.55-0.85)	0.77(0.63-0.91)	0.74(0.61-0.85)	0.34(0.09-0.58)	0.73(0.59-0.85)

SVM: Support Vector Machine, LR: Logistic Regression, RF: Random Forest, MLP: MultiLayer Perceptron, XGBoost : eXtreme Gradient Boosting

Table 4. False-negative (FN) error analysis for the osteopenia and osteoporosis patients which were calculated using XGBoost model for the test set

	Number of real patients	Number of FN	FN (%)	Expected FN (%)
Osteopenia	26	7	87.50	74.29
Osteoporosis	9	1	12.50	25.71
Total of diseases	35	8	100	100

Table 5. Confusion matrix for the osteopenia and osteoporosis patients which were calculated using XGBoost model for the test set

		Prediction	
		Positive	Negative
True value	Positive	27	8
	Negative	11	14

Table 6. List of a number of different studies on detection of osteoporosis using radiomics or ML techniques

Reference	Aim	Method	AUC
Sato et al. [23] used for from	Prediction of bone mineral density and T-score	Chest X-rays, DL	0.89, 0.70, 0.84
Küçükçiloğlu et al. [24]	Diagnose of osteoporosis	CT and MRI images of lumbar vertebrae, DL models	0.965, 0.988
Jang et al. [25]	Diagnosis of osteoporosis	Radiographic images of the hip joint, artificial intelligence algorithms	0.867
Zhao et al. [28]	Osteoporosis and abnormal bone density detection	Short lumbar mDixon sequence, automated radiomics screening pipeline	0.899
Nam et al. [29]	Prediction of osteoporotic spine	Lumbar CT images, ML model	0.900
Xie et al. [30]	Diagnosis of osteoporosis and osteopenia	CT images, ML-derived radiomics model	0.96
Huang et al. [31]	Detection of osteoporosis in health populations	General data, simple and user-friendly ML model	0.67, 0.77
Present study	Prediction of osteopenia/osteoporosis	Demographic data, ML	0.73

AUC: Area under the curve, ML: Machine learning, DL: Deep learning

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

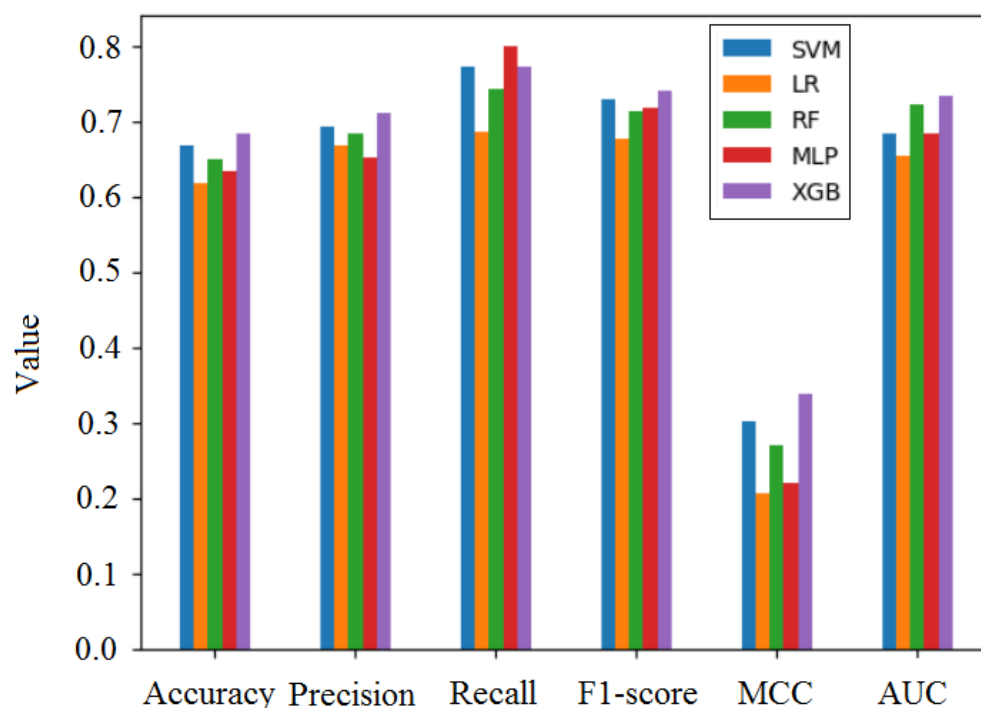


Figure 1. Barplot of Accuracy, precision, recall, F1-score, MCC and AUC values for SVM, LR, RF, MLP and XGBoost classifiers. (MCC: Mathews Correlation Coefficient, AUC: Area under the curve, SVM: Support Vector Machine, LR: Logistic Regression, RF: Random Forest, MLP: MultiLayer Perceptron, XGBoost : eXtreme Gradient Boosting)

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

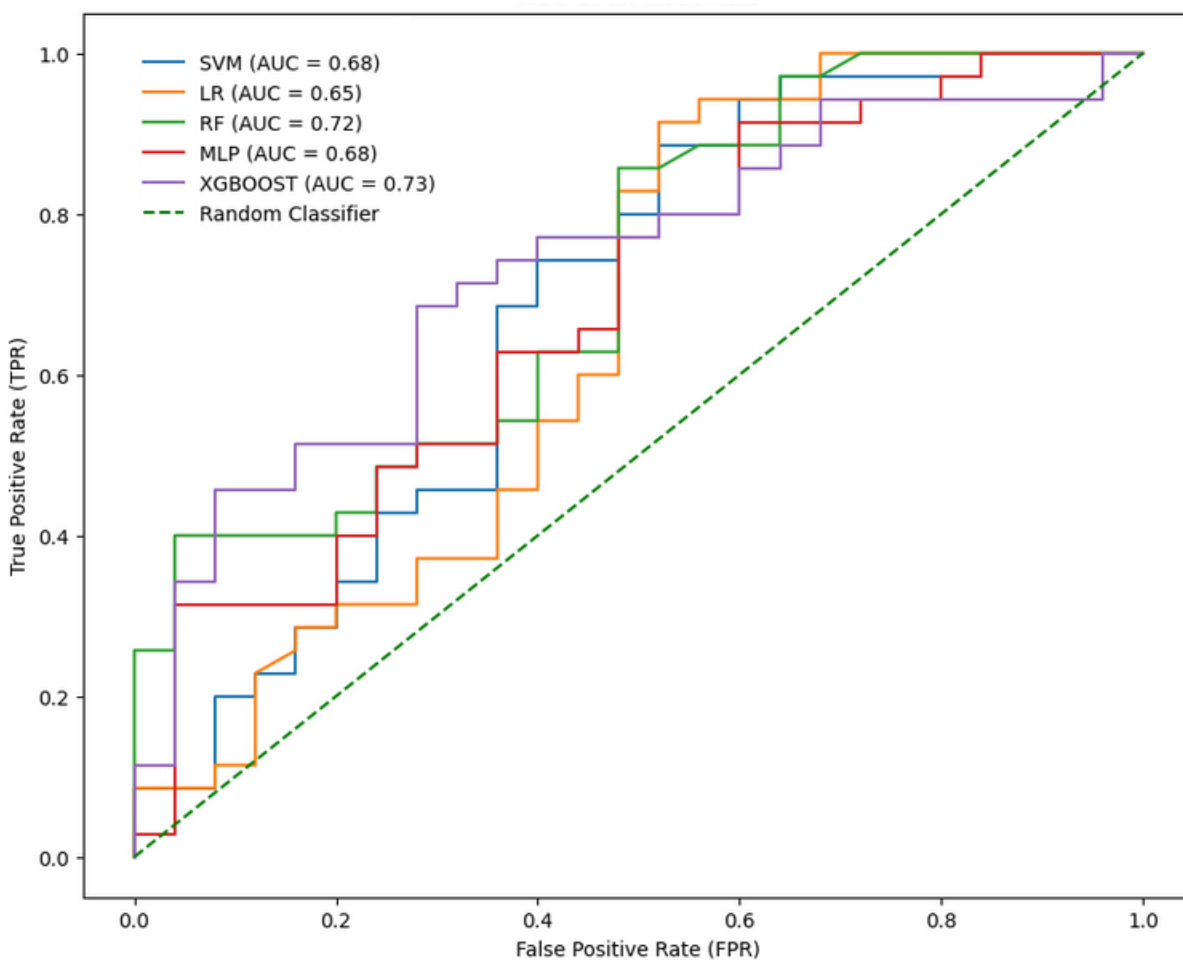


Figure 2. A receiver-operating characteristic (ROC) curve for SVM, LR, RF, MLP and XGBoost classifiers compared to the random classifier. (SVM: Support Vector Machine, LR: Logistic Regression, figfigRF: Random Forest, MLP: MultiLayer Perceptron, XGBoost : eXtreme Gradient Boosting)

How to cite this article:

Zohrevand H, Ghorbani M, Yazdanpanah S, Faghihi Moghadam F, Shalbaf A, Asgari MR, Langari H, Tabei F. A machine learning study on prediction of osteopenia/osteoporosis from demographic and clinical data. Iran J Nucl Med. doi: 10.22034/irjnm.2026.130533.1763

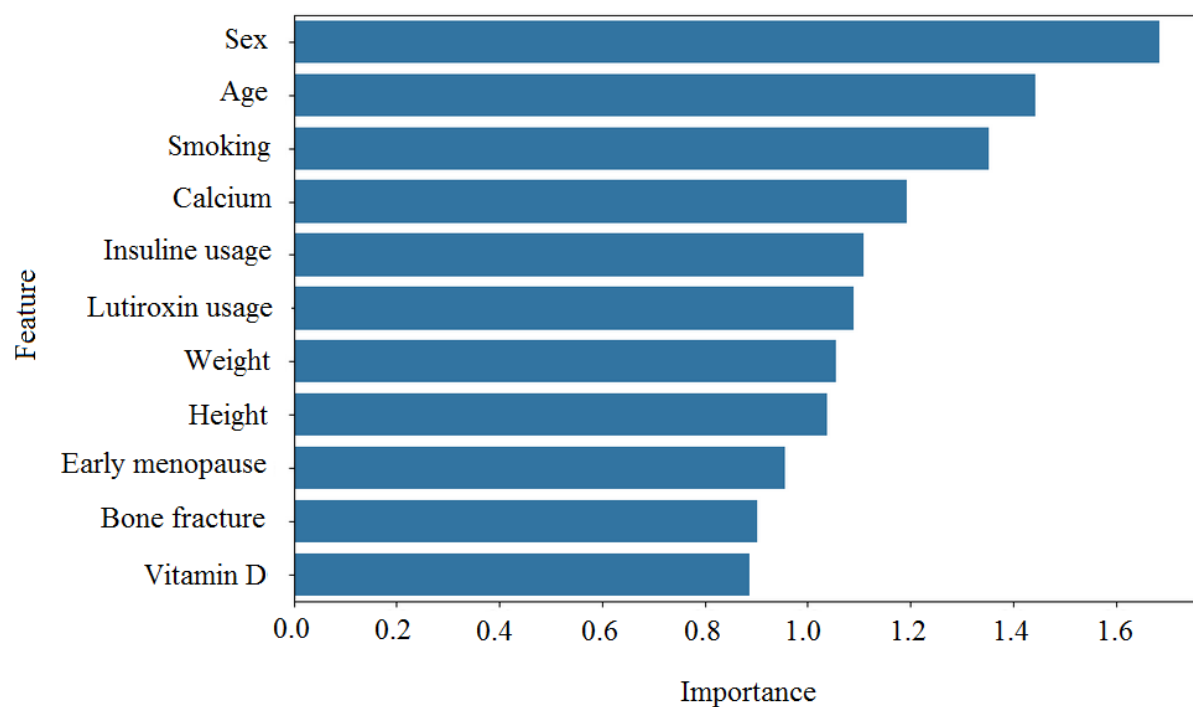


Figure 3. Importance for different demographic features of the cases, obtained using eXtreme Gradient Boosting (XGBoost) model