Original Investigation / Orijinal Araştırma

DOI: 10.4274/tod.galenos.2024.03708
Turk J Osteoporos

Opportunistic Prediction of Osteoporosis with Machine Learning Models Based on Clivus-radiomic Features Obtained from CT Images

Klivusa Ait BT Tabanlı Radiomics Özelliklerinden Makine Öğrenme Algoritmaları Kullanılarak Elde Edilen Modellerin Osteoporoz Tanısındaki Yeri

Candan Güngör[1], Emrah Akay[2], Fatih Erdem[3], Erdoğan Bülbül[2], Gülen Demirpolat[2], Bahar Yanık[2]

[1]Gaziantep City Hospital, Clinic of Radiology, Gaziantep, Türkiye
[2]Balıkesir University Faculty of Medicine, Department of Radiology, Balıkesir, Türkiye
[3]Ankara University Faculty of Medicine, Department of Pediatric Radiology, Ankara, Türkiye

Abstract

Objective: Osteoporosis (OP) is a major public health problem that causes significant mortality and morbidity. Therefore, early diagnosis is essential. We aimed to predict OP by combining computed tomography (CT)-based radiomic data of the clivus with machine learning (ML) algorithms.

Materials and Methods: In this retrospective study, 140 cases that underwent dual energy X-ray absorptiometry (DEXA) and craniofacial CT within one year of each other between 2015 and 2021, were examined at our institution. According to DEXA T-scores, cases were divided into three groups: 30 OP, 33 osteopenia, and 77 normal. Trabecular components of the clivus were segmented, and 1023 radiomic features were extracted using 3D Slicer. Radiomic outputs consist of features from original, Laplacian of Gaussian, and wavelet transform filtered images. Voxel resampling was standardized as 1x1x1 mm^3. Orange Data Mining program was used for ML. Relief and fast correlation-based filter were used for feature reduction. K-nearest neighborhood, decision tree, random forest, logistic regression, support vector machine (SVM), Naive Bayes, and neural network were used as classifiers. Area under the curve (AUC), sensitivity, specificity, receiver operating characteristic curve, and confusion matrix were used for performance evaluation.

Results: In binary classification as OP and non-OP, neural network achieved the highest success in predicting OP (AUC 0.87). In the binary classification of BMD as low BMD and normal BMD, SVM was the best in predicting low BMD cases (AUC: 0.82). In the ternary classification of BMD as OP, osteopenia, and normal, Naive Bayes achieved the highest performance in distinguishing OP (AUC: 0.9) and osteopenia (AUC: 0.69). The Hounsfield Units values of the clivus were significantly different between low BMD and normal BMD cases ($p<0.001$).

Conclusion: ML algorithms using CT-based radiomic features of the clivus can predict OP and provide BMD information.

Keywords: Osteoporosis, clivus, machine learning, radiomics

Öz

Amaç: Bilgisayarlı tomografi (BT) görüntüleri üzerinden klivusun radiomics verilerini, makine öğrenme algoritmaları ile kombine ederek osteoporozu (OP) tahmin etmeyi amaçladık.

Gereç ve Yöntem: Retrospektif çalışmamızda, kurumumuzda çift enerjili X-ışını absorpsiyometrisi (DEXA) ve bir yıl içerisinde kraniofasial bölgeye BT tetkiki yapılmış olan 140 olgu incelemeye alındı. Hastalar DEXA T-skorlarına göre, 30'u OP, 33'ü osteopeni ve 77'si normal üç gruba ayrıldı. Segmentasyon işlemi ve radiomics özelliklerin çıkarımı "3D slicer" programı ile tek hekim tarafından yapıldı. Klivusun manuel çizilerek segmente edildi. Radiomics çıktıları, orijinal, ince-kaba Laplacian of Gaussian ve wavelet transform filtreli görüntülerden oluşmaktadır. Toplam 1023 adet radiomics özellik elde edildi. Voksel yeniden örnekleme 1x1x1 mm^3 olarak standardize edildi. Makine öğrenmesi (MÖ) için Orange Data Mining programı kullanıldı. Özellik azaltma için relieff ve fast correlation based filter metodları uygulandı. MÖ algoritmaları olarak k-nearest neighborhood, decision tree, random forest, logistic regression, support vector machine (SVM), Naive Bayes ve neural network sınıflandırmaları kullanıldı. Sınıflandırmaları karşılaştırmak için eğri altında kalan alan (EAA), duyarlılık (sensitivity, recall), özgüllük (spesifite), alıcı çalışma karakteristik eğri analizi, hata matriksi gibi parametreler kullanıldı. Tüm istatistiksel sonuçlar için $p<0,05$ değeri anlamlı kabul edildi.

Bulgular: OP ve OP olmayan (osteopeni + normal) ikili sınıflandırmada OP tahminin en yüksek başarıyı nöral network algoritması elde etti (EAA: 0,87). Düşük kemik mineral yoğunluğu (KMY) ile normal KMY'li olgulardan oluşan ikili sınıflandırmada, düşük KMY'yi en iyi tahmin eden

Corresponding Author/Sorumlu Yazar: Candan Güngör MD, Gaziantep City Hospital, Clinic of Radiology, Gaziantep, Türkiye
E-mail: candan_gungor@hotmail.com **ORCID ID:** orcid.org/0000-0003-1793-8823
Received/Geliş Tarihi: 12.10.2024 **Accepted/Kabul Tarihi:** 27.11.2024 **Epub:** 03.07.2025

Cite this article as/Atıf: Candan Güngör C, Akay E, Erdem F, Bülbül E, Demirpolat G, Yanık B. Opportunistic prediction of osteoporosis with machine learning models based on clivus-radiomic features obtained from CT images. Turk J Osteoporos. [Epub Ahead of Print]

Öz

SVM algoritması oldu (EAA: 0,82). Son olarak OP, osteopeni ve normal olmak üzere üçlü gruplandırma yapıldı. OP'yi ayırt etmede en yüksek performansı Naive Bayes algoritması elde etti (EAA: 0,9). Osteopeni grubunu tahmin etmede de Naive Bayes algoritması ön plana çıktı (EAA: 0,69). Düşük KMY ile normal KMY'li olgular arasında HU değerleri anlamlı olarak farklıydı (p<0,001).

Sonuç: Çalışmamızda klivusun BT tabanlı radiomics çıktılarını kullanarak elde edilen MÖ algoritmalarının OP tahmininde kullanılabileceğini ve KMY hakkında fikir verdiği gösterdi.

Anahtar kelimeler: Osteoporoz, klivus, makine öğrenmesi, radiomics

Introduction

Osteoporosis (OP) is a serious public health problem with the increasing elderly population worldwide. In developed countries, 30% of all postmenopausal women have OP, and 50% of these patients experience one or more osteoporotic fractures in their lifetime (1). Vertebral and femoral fractures are more common than other bone fractures and are a significant cause of morbidity and mortality. Therefore, early diagnosis and fracture risk prediction are important in OP diagnosis (2,3).

Dual energy X-ray absorptiometry (DEXA) is the gold standard diagnostic method for OP diagnosis. However, erroneous results may be obtained with this two-dimensional examination in cases with osteodegenerative bony changes, vertebral instrumentations, and aortic calcifications. In recent years, quantitative computed tomography (CT) has emerged as a new diagnostic method in OP diagnosis, successfully calculating bone density and mass (4). However, since it is a relatively expensive technique, researchers have searched for alternative methods to predict OP, such as detecting morphological changes in bone structures through conventional imaging techniques and analyzing histogram features of bone structures through software, without the need for new hardware. There are many studies in periodontology and implant dentistry with these purposes (5-9). Lespessailles et al. (10) reported that the combined evaluation of bone tissue analysis and bone mineral density (BMD) is superior to the evaluation of BMD alone in the diagnosis of OP. Kawashima et al. (11) retrospectively extracted the histogram features of the sphenoid triangle, mandibular condyle, and clivus from cranial CT images and reported significant results in the diagnosis of OP.

Radiomics, a new image-processing approach, has been developed in recent years. Hundreds of features from medical images that the human eye cannot distinguish are obtained quantitatively (12). Radiomics achieves successful results in the differential diagnosis of tumors, determining the prognosis, and evaluating the response to treatment (13-15). In recent years, the number of studies related to radiomics and artificial intelligence in OP has been steadily increasing. (16-25). Machine learning (ML) is a subset of artificial intelligence. It is used in the medical field to calculate large and complex data sets and assist in medical decision-making.

He et al. (26) showed that magnetic resonance imaging (MRI) of the lumbar spine and radiomics models could be used in the diagnosis of OP. Rastegar et al. (27) obtained radiomics data from DEXA images and created ML models that can be used in the classification of bone mineral loss.

We aim to investigate the usability of radiomics and ML algorithms in OP prediction. The reason why we chose clivus is that studies focusing on clivus for OP prediction are very rare. The only study we encountered was published by Kawashima et al. (11). Unlike this histogram analysis-based study, we used radiomic outputs and ML algorithms, which consist of a much larger number of high-level tissue features.

Materials and Methods

Cases with DEXA and craniofacial region CT (brain, neck, maxillofacial, and paranasal sinus CT) imaging within a maximum interval of one year between 2015 and 2021 were scanned retrospectively. Age and gender were not considered as exclusion criteria. CT images with motion artifacts, IV contrast, and slice thickness of more than 1 mm were excluded from the study. Finally, a study group with 140 cases was obtained.

DEXA scan was performed with Lunar Prodigy (model 8743, GE Lunar, Madison, WI, USA). The patient height and weight were recorded. Anterior-posterior lumbar vertebrae and femur BMD are routinely measured. Body regions with implants were excluded during imaging.

The DEXA scan used L1-4 and the femur as the basis for T-scores. The lowest T-score was used to group cases. The cases were classified as "osteoporosis" if the T-score was <-2.5, "osteopenia" if it was between -2.5 and -1, and normal if it was >-1.23 Binary classification was made as OP and non-OP (osteopenia + normal), low BMD (OP + osteopenia), and normal BMDs, and ternary classification was made as OP, osteopenia, and normal.

CT scans were performed with a 64-slice multidetector CT (Aquillon 64, Toshiba, Otawara, Japan). The parameters used in imaging are Pitch factor 0.6-0.9, rotation time 0.5-0.75 seconds, tube voltage 120 kV, tube current 150-250 mAs, and slice thickness 0.5-1 mm.

3D Slicer 4.11.2 (www.slicer.org) program was utilized for the segmentation process. After anonymization, CT images were obtained in DICOM format and imported into 3D Slicer. An experienced radiologist manually segmented trabecular bone components of the clivus. The petrooccipital fissure laterally and the hypoglossal canal inferiorly limited the segmentation borders. Dorsum sella and cortical bone were excluded from the segmentation (Figure 1).

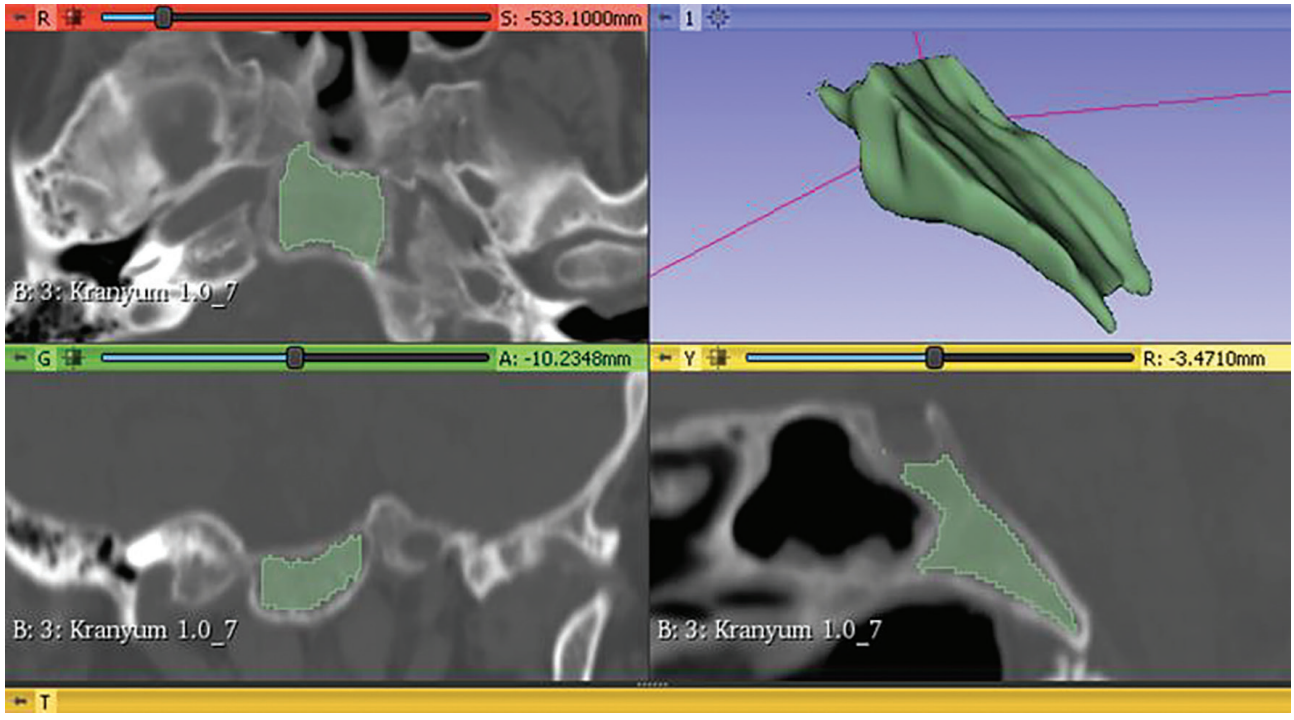

Figure 1. Clivus segmentation process

Laplacian of Gaussian image filters with two sigma values (0.5 mm and 2.5 mm) and wavelet transform filters were used for image filtering before radiomic feature extraction to create a high-throughput dataset. Voxel size for resampling was defined as 1x1x1 mm^3 for standardization.

A total of 1023 features were obtained, including 18 first-order features, 24 GLCM (Gray Level Co-Occurrence Matrix), 14 GLDM (Gray Level Dependence Matrix), 16 GLRLM (Gray Level Run Length Matrix), 16 GLSZM (Gray Level Size Zone Matrix), 5 NGTDM (Neighbouring Gray Tone Difference Matrix) based features, 93 features from Laplacian of Gaussian filtered images with sigma value of 0.5 mm, 93 features from Laplacian of Gaussian filtered images with sigma value of 2.5 mm, and 744 features from wavelet transformed images. Detailed mathematical descriptions of radiomic features are available in the pyRadiomics library (https://pyradiomics.readthedocs.io/en/latest/features.html).

Orange Data Mining Tool Version 3.27 (https://orange.biolab.si) was used for feature reduction and classification models. One scoring method among information gain, information gain ratio, Gini decrease, ANOVA, chi² (x^2), ReliefF, and fast correlation-based filter (FCBF) was used for feature selection. The best combination of the feature selection method and the number of features to be used was determined by the best-performing ML algorithm: The one with the highest area under the curve (AUC) after numerous tests. Stratified 10-fold cross-validation technique was used for validation.

K-nearest neighborhood, decision tree, random forest, logistic regression, support vector machine (SVM), Naive Bayes, and neural network were used as ML algorithms. AUC, classification accuracy (CA), sensitivity (recall), specificity, F1 score, precision, receiver operating characteristic (ROC) curve, and confusion matrix were used to evaluate ML model performances.

The Local Clinical Research Ethics Committee of Balıkesir University approved this study on 03.11.2021 with the decision number 2021/249.

Statistical Analysis

Statistical analysis was performed in the IBM SPSS 22.0 (SPSS Inc., Chicago, IL, USA) program. The Kolmogorov-Smirnov test was used to determine whether the data was normally distributed. Independent variables were shown as mean and standard deviation. The Tukey's HSD posthoc test was used to determine the relationship between BMD groups. Pearson and Spearman correlation tests evaluated the relationship between continuous independent variables. Dependent variables were evaluated with the chi-square test.

The Hounsfield Units (HU) values of the clivus were measured by drawing the largest region of interests (ROI) covering the trabecular bone from three consecutive axial CT slices, and their arithmetic mean was calculated for each case. Whether the mean HU values were discriminative in detecting the BMD group was evaluated with AUC, cut-off, sensitivity, and specificity parameters by performing ROC analysis. $P<0.05$ was considered significant in all statistical results.

The flow diagram is summarized in Figure 2.

Results

In our study, a total of 140 cases consisting of 124 women and 16 men aged between 33-91 years were included. Cases were

divided into three groups consisting of 30 OP, 33 osteopenia, and 77 normal cases according to T-scores. No statistically significant relationship was found between gender, age, and OP due to the low number of cases and the inhomogeneous age distribution. However, when compared according to T-scores, the mean T-scores of men (0.11) were significantly higher than the mean T-scores of women [(-1) (t(133)=-2.2, p=0.024]. BMI values were significantly lower in the OP group compared to the normal group (p=0.002) (Table 1). No statistically significant difference was found when the osteopenia vs. normal group and OP vs. osteopenia group comparisons were made.

First, cases were divided into two groups: OP and non-OP (osteopenia + normal). The feature selection method was chosen as ReliefF. 10 out of 1023 features were selected. In the classification process, the best-performing classifier predicting OP was neural network (AUC=0.87, CA=0.86) (Table 2). 102 of 110 non-OP cases were correctly identified, resulting in a very high specificity value (specificity 0.93). Some classifiers showed higher specificity values, such as SVM and logistic regression. However, these classifiers have lower reliability due to their lower sensitivity and F1 scores. The ROC curves of the ML algorithms are given in Figure 3.

The other binary classification was performed between cases with low BMD (osteopenia + OP) and normal BMD. We aimed to predict the decrease in BMD with ML algorithms. Sixteen features were selected from the database with ReliefF. In the classification process, SVM showed the most successful performance (AUC: 0.82, CA: 0.79) (Table 3), correctly predicting 46 of 63 patients with abnormal BMD and 65 of 77 patients with normal BMD. Other performance metrics of SVM were calculated as sensitivity 0.73, specificity 0.84, F1 score 0.76, and precision 0.79. All performance metrics of SVM to predict low BMD were higher

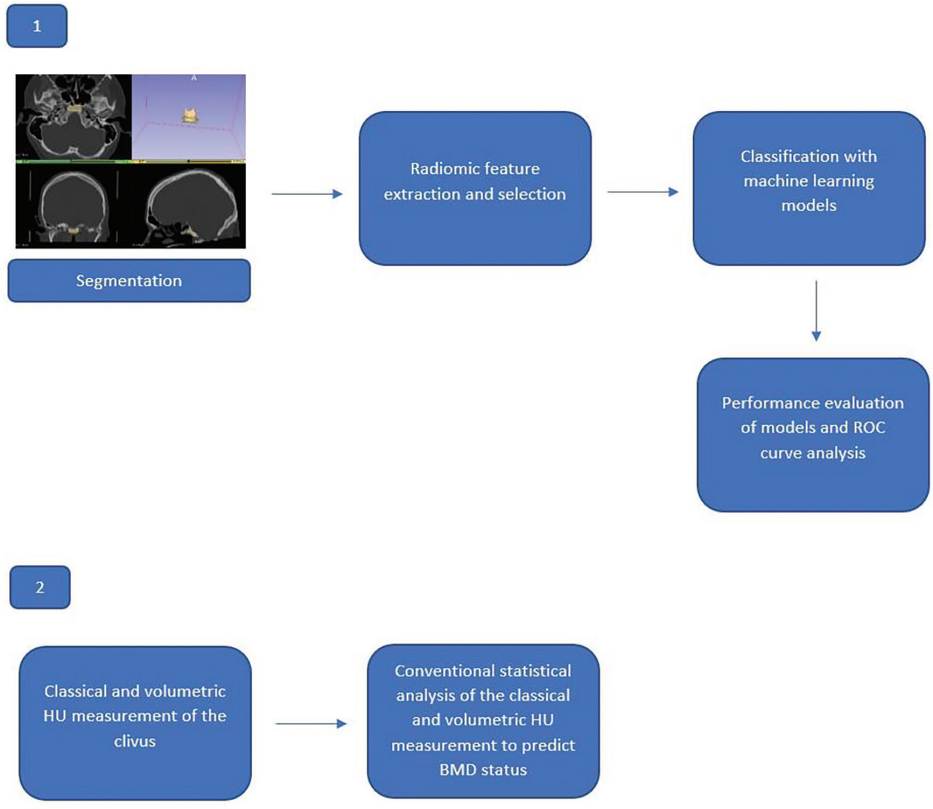

Figure 2. Flow diagram

BMD: Bone mineral density, HU: Hounsfield Units, ROC: Receiver operating characteristic

Table 1. Demographic data of BMD groups

	Osteoporosis n=30	Osteopenia n=33	Normal n=77	p-value
Gender (female/male)	27/3	32/1	65/12	0.149
Age (mean ± SD)	73±10	69±11	68±12	0.076
BMI	27±5	28±5	31±6	0.002

BMD: Bone mineral density, BMI: Body mass index, SD: Standard deviation

than the other algorithms. The ROC curves of the ML algorithms are given in Figure 4.

As a final ML classification step, cases were divided into three groups: Osteopenia, OP, and normal. FCBF method was applied, and the most optimal seven features were selected. In this ternary classification, the Naive Bayes algorithm was the best-performing classifier in distinguishing OP (AUC: 0.9, CA: 0.86) (Table 4), correctly predicting 22 of 30 cases with OP and 66 of 77 normal cases. Sensitivity was 0.73, and specificity was 0.89. Some classification methods, such as SVM, logistic regression, and random forest, reach higher specificity. However, these algorithms' sensitivities and F1 scores lag behind the Naive Bayes algorithm. The ROC curves of the algorithms for estimating OP in the ternary classification consisting of OP, osteopenia, and normal cases are given in Figure 5.

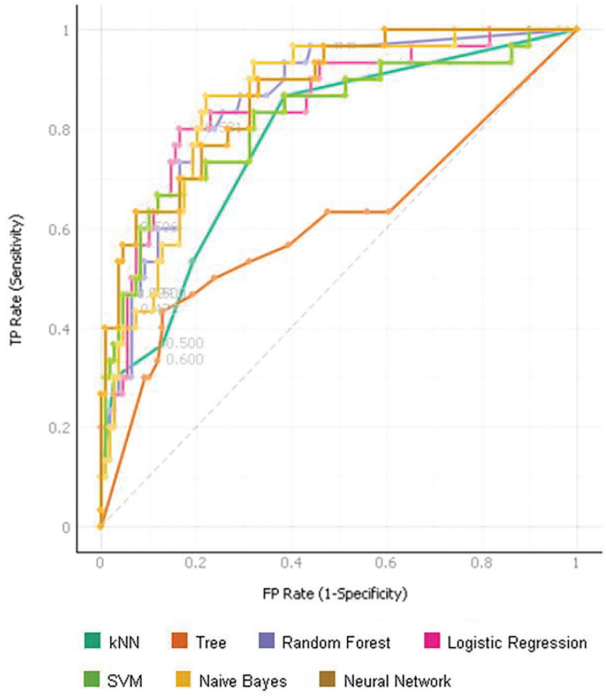

Figure 3. ROC curves of the ML algorithms to predict OP in the binary classification (OP and non-OP)

ML: Machine learning, OP: Osteoporosis, ROC: Receiver operating characteristic, SVM: Support vector machine, kNN: K-nearest neighborhood

In the ternary classification, the performance of ML algorithms in detecting cases with osteopenia is low (Table 5). The highest performance was obtained with the Naive Bayes algorithm (AUC: 0.69, CA: 0.76), predicting 11 of 33 osteopenic patients. 15 were misclassified as normal, and seven as OP.

When the mean HU values of clivus from three axial slices were calculated by the ROI method, a moderate positive correlation was found between HU values and T-scores (r^2=0.45 p<0.001). The cases were divided into three groups: OP, osteopenia, and normal. The mean HU value was 103 (74.9-131.1 with a 95% confidence interval) in the OP group, 113.8 (88.9-138.7 with a 95% confidence interval) in the osteopenia group, and 192 (168-215,9 with a 95% confidence interval) in the normal group (Table 6). Significant differences were found in the values measured between the low BMD (OP + osteopenia) and the normal group (p<0.001). No significant relationship was found between the mean HU values in the OP and osteopenia groups. ROC analysis was performed to determine the success of the classical HU measurement method in predicting the low BMD group (Figure 6). The AUC value was 0.75 (0.67-0.83 with a 95% confidence interval), the cut-off value was 137 HU, and the sensitivity and specificity values were 0.6 and 0.72, respectively. Finally, the volumetric mean HU values obtained from the segmentation of the clivus were examined. The original first-order mean values among the radiomic features, which express the volumetric mean HU value, were used without extra processing. There was no significant correlation between volumetric mean HU values and OP, osteopenia, and normal groups.

Discussion

High AUC values, such as 0.9 and 0.87, were obtained in the OP estimation using radiomics and ML algorithms. Osteopenia prediction performance was lower than OP prediction performance but at an acceptable level, at 0.82 (AUC). The combined use of radiomics and ML algorithms was significantly superior to HU values measured using the traditional ROI method in detecting OP and low BMD.

Table 2. Performance metrics of ML algorithms to predict OP in the binary classification (OP vs. non-OP)

Model	AUC	CA	F1	Precision	Sensitivity	Specificity
Neural network	0.87	0.9	0.7	0.7	0.63	0.93
Random forest	0.86	0.8	0.6	0.63	0.5	0.92
SVM	0.86	0.8	0.5	0.71	0.4	0.95
Naive Bayes	0.85	0.8	0.6	0.51	0.83	0.78
Logistic regression	0.84	0.8	0.6	0.67	0.47	0.94
kNN	0.78	0.8	0.4	0.44	0.37	0.87
Tree	0.6	0.8	0.4	0.43	0.33	0.88

ML: Machine learning, OP: Osteoporosis, AUC: Area under the curve, CA: Classification accuracy, SVM: Support vector machine, kNN: K-nearest neighborhood

Apart from being two-dimensional imaging and using ionizing radiation, the most significant disadvantage of DEXA is the possibility of superimposition of dense structures such as soft tissues, metallic instruments, osteodegenerative changes, and atherosclerotic calcifications, which may cause BMD to be miscalculated. It is mentioned in the literature that the use of CT imaging in such cases can help diagnose missed OP (28,29). We chose the clivus for this study because it is less prone to degeneration and is included in the field of view of common CT scans such as brain CT.

In the literature, there are efforts to develop an alternative diagnostic tool due to the limitations of DEXA. Many studies report a positive correlation between T-scores and HU values obtained from bone CT scans, such as lumbar and wrist CT scans (30-36). Alawi et al. (37) reported a positive correlation between DEXA T-scores and HU values of lumbar vertebrae from abdominal CT images. Their study measured mean attenuation values as 115 HU in osteoporotic cases, 120 HU in osteopenic cases, and 174 HU in normal cases. While the difference between the abnormal BMD and normal groups was statistically significant, there was no statistically significant difference between the OP and osteopenia groups (37). Similar mean attenuation values were measured in our study: 103 HU in the OP group, 113 HU in the osteopenia group, and 192 HU in the normal group. Decreases in mean HU values in the low BMD group were also statistically significant in our study. Our

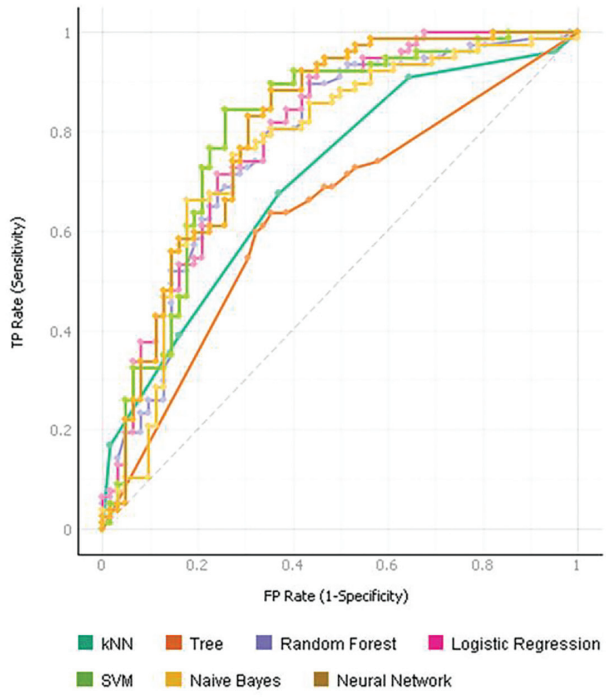

Figure 4. ROC curves of the ML algorithms to predict low BMD in the binary classification (low and normal BMD)

ML: Machine learning, BMD: Bone mineral density, SVM: Support vector machine, kNN: K-nearest neighborhood, ROC: Receiver operating characteristic

Table 3. Performance metrics of ML algorithms to predict low BMD in the binary classification (low and normal BMD)

Model	AUC	CA	F1	Precision	Sensitivity	Specificity
SVM	0.82	0.79	0.76	0.79	0.73	0.84
Neural network	0.80	0.76	0.72	0.77	0.68	0.83
Naive Bayes	0.79	0.74	0.71	0.71	0.71	0.77
Logistic regression	0.78	0.71	0.68	0.69	0.67	0.75
Random forest	0.77	0.74	0.69	0.73	0.65	0.81
kNN	0.72	0.66	0.63	0.62	0.63	0.68
Tree	0.63	0.62	0.56	0.59	0.54	0.69

ML: Machine learning, BMD: Bone mineral density, AUC: Area under the curve, CA: Classification accuracy, SVM: Support vector machine, kNN: K-nearest neighborhood

Table 4. Performance metrics of ML algorithms to predict OP in the ternary classification (OP, osteopenia, and normal)

Model	AUC	CA	F1	Precision	Sensitivity	Specificity
Naive Bayes	0.90	0.86	0.69	0.65	0.73	0.89
Logistic regression	0.87	0.79	0.00	0.00	0.00	1.00
SVM	0.84	0.85	0.62	0.68	0.57	0.93
Neural network	0.84	0.82	0.59	0.58	0.60	0.88
Random forest	0.82	0.84	0.58	0.64	0.53	0.92
kNN	0.78	0.77	0.41	0.46	0.37	0.88
Tree	0.75	0.81	0.57	0.57	0.57	0.88

ML: Machine learning, OP: Osteoporosis, AUC: Area under the curve, CA: Classification accuracy, SVM: Support vector machine, kNN: K-nearest neighborhood

Table 5. Performance metrics of ML algorithms to predict osteopenia in the ternary classification (OP, osteopenia, and normal)

Model	AUC	CA	F1	Precision	Sensitivity	Specificity
Naive Bayes	0.69	0.76	0.39	0.48	0.33	0.89
Tree	0.65	0.79	0.46	0.57	0.39	0.91
Neural network	0.61	0.76	0.33	0.50	0.24	0.93
SVM	0.59	0.78	0.34	0.57	0.24	0.94
Random forest	0.58	0.74	0.30	0.40	0.24	0.89
Logistic regression	0.54	0.76	0.00	0.00	0.00	1.00
kNN	0.39	0.71	0.20	0.28	0.15	0.88

ML: Machine learning, OP: Osteoporosis, AUC: Area under the curve, CA: Classification accuracy, SVM: Support vector machine, kNN: K-nearest neighborhood

Table 6. The means, standard deviations and 95% confidence intervals of HU values measured with the ROI method are shown

	Hounsfield Units (HU)	
	Mean ± SD	%95 confidence interval
Osteoporosis	103±13.7	74.9-131.1
Osteopenia	113.8±13.2	88.9-138.7
Normal	192±12	168-215.9

ROI: Region of interest, SD: Standard deviation

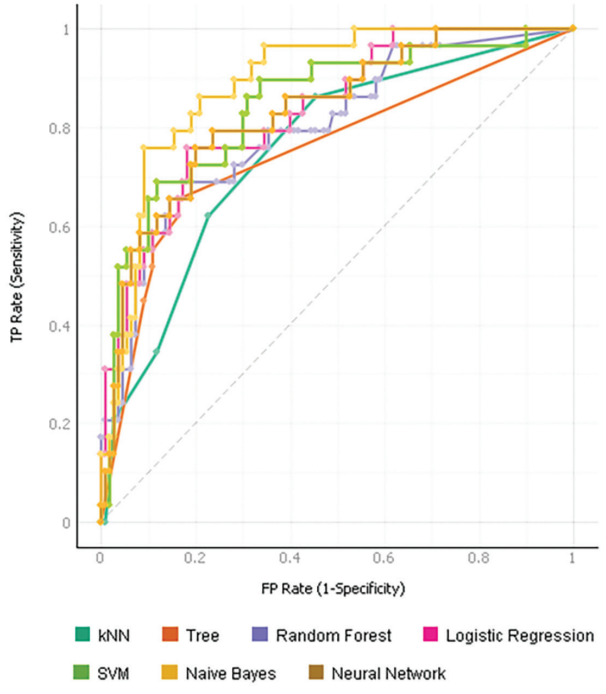

Figure 5. ROC curves of the ML algorithms to predict OP in the ternary classification (OP, osteopenia, and normal)

ML: Machine learning, OP: Osteoporosis, ROC: Receiver operating characteristic, SVM: Support vector machine, kNN: K-nearest neighborhood

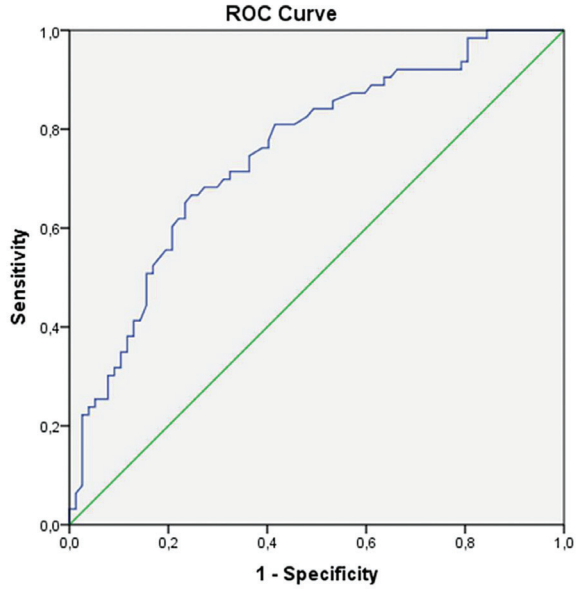

Figure 6. ROC curve of the classical HU measurement to predict low BMD

BMD: Bone mineral density HU: Hounsfield units, ROC: Receiver operating characteristic

study also had no significant difference between the OP and osteopenia groups. According to the ROC analysis, the group with low BMD was correctly diagnosed with a cut-off value of 137 HU with 72% specificity and 68% sensitivity. Considering that half of the insufficiency fractures in the population occur in osteopenic women (38), identifying patients with low BMD may be more important than distinguishing osteopenia and OP. In a study conducted with a small number of patients (29 normal, 29 OP), Kawashima et al. (11) extracted two-dimensional radiomic features from CT images of bilateral greater wings of sphenoid, bilateral mandibular condyles, and clivus using the ROI method. The various types of texture features extracted from craniofacial trabecular bones, such as histogram features, GLCM features, and GLRL features, were found to be associated with OP. It is also mentioned that the clivus, one of the three skull base structures examined in the study, stands out as being less affected by degenerative findings (11).

Rastegar et al. (27) extracted radiomic features from lumbar and femoral DEXA images with the ROI method and analyzed them with ML algorithms. Moderate diagnostic performance (AUC) values ranged from 0.5 to 0.78 in distinguishing OP, osteopenia, and normal groups (27).

In their retrospective study, Lim et al. (39) showed that the ML models using radiomic features obtained from abdominopelvic CT images can predict femoral OP. The proximal femur was automatically segmented, including the cortex. The number of radiomic features was limited to 41, consisting of semantic features, first-level tissue features, GLCM, and wavelet transform features. They used the Gini decrease for feature reduction and the random forest algorithm for classification. The cases were divided into two: 70% were used for the training dataset and 30% for the validation dataset. The random forest algorithm successfully predicted OP with 95% specificity and 80% sensitivity in the validation group. In addition, this study used 5-fold cross-validation. It is recommended to use 5 or 10 folds in the literature. We used 10-fold cross-validation technique in our study. Unlike this study, we did not divide the cases into training and validation datasets due to the limited number of cases.

In a recent article by Fang et al. (20), they mention that 2D transfer learning and 3D deep learning techniques have shown excellent performance in screening for OP in chest CT scans. In another recent article, it was found that in opportunistic OP screening using chest CT scans, the three-dimensional segmentation of the thoracic vertebral body and the subsequent radiomics outputs showed similar performance to ML models. The AUC values are similar to those in our article (AUC: 0.8-0.9) (21).

In another study regarding osteoporotic fracture estimation, using microstructural femoral MRI data and fracture risk assessment tool (FRAX) data together with ML algorithms was superior to using MRI data and FRAX data alone (40). A study conducted in India proposed that an automated diagnostic technique for low bone mass is possible using radiogrammetric measurements and texture features from radiography images together with a three-layer supervised artificial neural network (41).

Study Limitations

The main limitation of our study, apart from its retrospective nature, is the low number of patients. A larger patient group is needed for the use of training and external validation groups. In addition, the patient population was obtained from a specific region, and the findings may not be generalized worldwide. In our study, BMD was classified according to DEXA T-scores. Therefore, due to the nature of DEXA, erroneous BMD and T-scores may have been obtained, which may have misled the statistical results. In future studies, it will be possible to compare the performances of radiomics scores and ML algorithms with DEXA by grouping them as those with and without osteoporotic fractures. Using automatic segmentation can be beneficial in terms of standardization and saving time. Although the variety and number of algorithms we use are higher than most studies, it is a fact that there are more ML algorithms available to use. The DEXA and CT imaging time interval has been accepted as a maximum of one year, and this period can be kept shorter. In addition, the systemic diseases and the drugs used were not considered.

Conclusion

Our study showed that OP and osteopenia can be accurately detected using CT-based radiomic features of clivus and ML. We also found that clivus CT HU values correlated positively with DEXA T-scores.

Ethics

Ethics Committee Approval: The Local Clinical Research Ethics Committee of Balıkesir University approved this study on 03.11.2021 with the decision number 2021/249.

Informed Consent: Retrospective study.

Acknowledgments

Our study was uploaded to the 2022 thesis database with the reference number 10500307. It was also accepted and published as a poster at the Turkish Radiology Congress in 2022 and the European Radiology (ECR) Congress in 2024.

Footnotes

Authorship Contributions

Concept: C.G., E.A., G.D., Design: C.G., E.A., E.B., B.Y., Data Collection or Processing: C.G., Analysis or Interpretation: C.G., E.A., F.E., Literature Search: C.G., Writing: C.G., F.E.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Sambrook P, Cooper C. Osteoporosis. Lancet. 2006;367:2010-8.
2. Felsenberg D, Silman AJ, Lunt M, Armbrecht G, Ismail AA, Finn JD, et al. Incidence of vertebral fracture in europe: results from the European Prospective Osteoporosis Study (EPOS). J Bone Miner Res. 2002;17:716-24.
3. Johnel O, Gullberg B, Allander E, Kanis JA. The apparent incidence of hip fracture in Europe: a study of national register sources. Osteoporos Int. 1992;2:298-302.
4. Grampp S, Genant HK, Mathur A, Lang P, Jergas M, Takada M, et al. Comparisons of noninvasive bone mineral measurements in assessing age-related loss, fracture discrimination, and diagnostic classification. J Bone Miner Res. 1997;12:697-711.
5. Munhoz L, Gil Choi I, Miura D, Watanabe P, Arita E. Bone mineral density and mandibular osteoporotic alterations in panoramic radiographs: Correlation by peripheral bone densitometry in men. Indian J Dent Res. 2020;31:457.
6. Merheb J, Temmerman A, Coucke W, Rasmusson L, Kübler A, Thor A, et al. Relation between spongy bone density in the maxilla and skeletal bone density. Clin Implant Dent Relat Res. 2015;17:1180-7.
7. Kavitha MS, An SY, An CH, Huh KH, Yi WJ, Heo MS, et al. Texture analysis of mandibular cortical bone on digital dental panoramic

radiographs for the diagnosis of osteoporosis in Korean women. Oral Surg Oral Med Oral Pathol Oral Radiol. 2015;119:346-56.

8. Lin ZT, Wang TM, Ge JY, Lin H, Zhu XF. [Analysis of mandibular bone mineral density of senile osteoporosis patients]. Zhonghua Kou Qiang Yi Xue Za Zhi. 2010;45:214-8. Chinese.
9. Aliaga I, Vera V, Vera M, García E, Pedrera M, Pajares G. Automatic computation of mandibular indices in dental panoramic radiographs for early osteoporosis detection. Artif Intell Med. 2020;103:101816.
10. Lespessailles E, Gadois C, Kousignian I, Neveu JP, Fardellone P, Kolta S, et al. Clinical interest of bone texture analysis in osteoporosis: A case control multicenter study. Osteoporos Int. 2008;19:1019-28.
11. Kawashima Y, Fujita A, Buch K, Li B, Qureshi MM, Chapman MN, et al. Using texture analysis of head CT images to differentiate osteoporosis from normal bone density. Eur J Radiol. 2019;116:212-8.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278:563-77.
13. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RTH, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol. 2015;114:345-50.
14. Abdollahi H, Tanha K, Mofid B, Razzaghdoust A, Saadipoor A, Khalafi L, et al. MRI radiomic analysis of IMRT-induced bladder wall changes in prostate cancer patients: a relationship with radiation dose and toxicity. J Med Imaging Radiat Sci. 2019;50:252-60.
15. Nazari M, Shiri I, Hajianfar G, Oveisi N, Abdollahi H, Deevband MR, et al. Noninvasive Fuhrman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning. Radiol Med. 2020;125:754-62.
16. Fang K, Zheng X, Lin X, Dai Z. Unveiling osteoporosis through radiomics analysis of hip CT imaging. Acad Radiol. 2023:S1076-6332(23)00544-5.
17. Lai YH, Tsai YS, Su PF, Li CI, Chen HHW. A computed tomography radiomics-based model for predicting osteoporosis after breast cancer treatment. Phys Eng Sci Med. 2024;47:239-248.
18. Zhen T, Fang J, Hu D, Shen Q, Ruan M. Comparative evaluation of multiparametric lumbar MRI radiomic models for detecting osteoporosis. BMC Musculoskelet Disord. 2024;25:185.
19. Martel D, Monga A, Chang G. Radiomic analysis of the proximal femur in osteoporosis women using 3T MRI. Front Radiol. 2023;3:1293865.
20. Fang K, Zheng X, Lin X, Dai Z. A comprehensive approach for osteoporosis detection through chest CT analysis and bone turnover markers: harnessing radiomics and deep learning techniques. Front Endocrinol (Lausanne). 2024;15:1296047.
21. Lin X, Shen R, Zheng X, Shi S, Dai Z, Fang K. Utilizing radiomics techniques to isolate a single vertebral body from chest CT for opportunistic osteoporosis screening. BMC Musculoskelet Disord. 2024;25:785.
22. Tong X, Wang S, Zhang J, Fan Y, Liu Y, Wei W. Automatic osteoporosis screening system using radiomics and deep learning from low-dose chest CT images. Bioengineering (Basel). 2024;11:50.
23. Chen B, Cui J, Li C, Xu P, Xu G, Jiang J, et al. Application of radiomics model based on lumbar computed tomography in diagnosis of elderly osteoporosis. J Orthop Res. 2024;42:1356-68.
24. Nian S, Zhao Y, Li C, Zhu K, Li N, Li W, et al. Development and validation of a radiomics-based model for predicting osteoporosis in patients with lumbar compression fractures. Spine J. 2024;24:1625-34.
25. Cheng L, Cai F, Xu M, Liu P, Liao J, Zong S. A diagnostic approach integrated multimodal radiomics with machine learning models based on lumbar spine CT and X-ray for osteoporosis. J Bone Miner Metab. 2023;41:877-89.
26. He L, Liu Z, Liu C, Gao Z, Ren Q, Lei L, et al. Radiomics based on lumbar spine magnetic resonance imaging to detect osteoporosis. Acad Radiol. 2021;28:e165-71.
27. Rastegar S, Vaziri M, Qasempour Y, Akhash MR, Abdalvand N, Shiri I, et al. Radiomics for classification of bone mineral loss: A machine learning study. Diagn Interv Imaging. 2020;101:599-610.
28. Zou D, Li W, Deng C, Du G, Xu N. The use of CT Hounsfield unit values to identify the undiagnosed spinal osteoporosis in patients with lumbar degenerative diseases. Eur Spine J. 2019;28:1758-66.
29. Choi MK, Kim SM, Lim JK. Diagnostic efficacy of Hounsfield units in spine CT for the assessment of real bone mineral density of degenerative spine: correlation study between T-scores determined by DEXA scan and Hounsfield units from CT. Acta Neurochir (Wien). 2016;158:1421-7.
30. Marinova M, Edon B, Wolter K, Katsimbari B, Schild HH, Strunk HM. Use of routine thoracic and abdominal computed tomography scans for assessing bone mineral density and detecting osteoporosis. Curr Med Res Opin. 2015;31:1871-81.
31. Johnson CC, Gausden EB, Weiland AJ, Lane JM, Schreiber JJ. Using hounsfield units to assess osteoporotic status on wrist computed tomography scans: comparison with dual energy X-ray absorptiometry. J Hand Surg Am. 2016;41:767-74.
32. Alacreu E, Moratal D, Arana E. Opportunistic screening for osteoporosis by routine CT in Southern Europe. Osteoporos Int. 2017;28:983-90.
33. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N. Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. Ann Intern Med. 2013;158:588-95.
34. Schreiber JJ, Anderson PA, Rosas HG, Buchholz AL, Au AG. Hounsfield units for assessing bone mineral density and strength: A tool for osteoporosis management. J Bone Joint Surg Am. 2011;93:1057-63.
35. Wagner SC, Formby PM, Helgeson MD, Kang DG. Diagnosing the undiagnosed osteoporosis in patients undergoing lumbar fusion. Spine (Phila Pa 1976). 2016;41:E1279-83.
36. Lee S, Chung CK, Oh SH, Park SB. Correlation between bone mineral density measured by dual-energy X-ray absorptiometry and hounsfield units measured by diagnostic CT in lumbar spine. J Korean Neurosurg Soc. 2013;54:384-9.
37. Alawi M, Begum A, Harraz M, Alawi H, Bamagos S, Yaghmour A, et al. Dual-energy X-ray absorptiometry (DEXA) scan versus computed tomography for bone density assessment. Cureus. 2021;13:e13261.
38. Sanders KM, Nicholson GC, Watts JJ, Pasco JA, Henry MJ, Kotowicz MA, et al. Half the burden of fragility fractures in the community occur in women without osteoporosis. When is fracture prevention cost-effective? Bone. 2006;38:694-700.
39. Lim HK, Ha H Il, Park SY, Han J. Prediction of femoral osteoporosis using machine-learning analysis with radiomics features and abdomen-pelvic CT: a retrospective single center preliminary study. PLoS One. 2021;16:e0247330.
40. Ferizi U, Besser H, Hysi P, Jacobs J, Rajapakse CS, Chen C, et al. Artificial intelligence applied to osteoporosis: a performance comparison of machine learning algorithms in predicting fragility fractures from MRI data. J Magn Reson Imaging. 2019;49:1029-38.
41. Areeckal AS, Jayasheelan N, Kamath J, Zawadynski S, Kocher M, David S S. Early diagnosis of osteoporosis using radiogrammetry and texture analysis from hand and wrist radiographs in Indian population. Osteoporos Int. 2018;29:665-73.