Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multi-site Radiomics and Machine Learning for Bone Status Classification in Panoramic Radiographs: A STARD-Compliant Diagnostic Accuracy Study

Kelly Wivian de Souza Leitão Viana
Federal University of Rio de Janeiro
Carla Barros de Oliveira
Federal University of Rio de Janeiro
Sâmila Barra
Federal University of Rio de Janeiro
Cláudia Borges Brasileiro
Universidade Federal de Minas Gerais
Miguel Madeira
Federal University of Rio de Janeiro
Maria Augusta Visconti
gutavisconti@odonto.ufrj.br
Federal University of Rio de Janeiro

Research Article

Keywords: Radiomics, Texture analysis, Panoramic Radiography, Bone Density

Posted Date: July 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9971493/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Objective

To evaluate texture analysis (TA) in panoramic radiographs for bone status classification (normal, osteopenia, osteoporosis) in women and its correlation with clinical variables.

Materials and methods

Panoramic radiographs of 47 women (Control, n = 11; Osteopenia, n = 25; Osteoporosis, n = 11) were retrospectively analyzed. ROIs were defined in the maxillary tuberosity and posterior mandible. Thirty-six features selected via Fisher and POE + ACC were extracted using MaZda software. Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM) models were developed with SMOTE for class balancing.

Results

The 10th Percentile showed significant intergroup differences in both ROIs ($p < 0.05$). RF achieved acceptable discriminative power (AUC > 0.70), whereas LR performed poorly (AUC < 0.60). Although SVM showed high sensitivity for osteopenia (0.875), it failed for the control group (Sens: 0.0) and showed low accuracy for osteoporosis (Sens: 0.333). Strong correlations were confirmed between GLRLM features and age at menopause onset ($p < 0.05$, FDR-adjusted); no associations were found with BMD or fracture risk ($p > 0.05$).

Conclusions

Panoramic radiograph TA shows potential for bone screening, particularly in detecting early postmenopausal changes. However, reliability is highly model-dependent; RF provided the most stable global classification, whereas SVM was limited to osteopenia detection. Small sample size warrants further large-scale validation.

Clinical relevance: Conventional panoramic radiographs serve as a powerful opportunistic screening tool for early systemic bone loss. By integrating multi-site radiomics and machine learning, clinicians can detect sub-visual trabecular microarchitectural degradation, particularly early-stage osteopenia, during routine dental evaluations, bridging the diagnostic gap left by densitometric examinations (DEXA) that fail to capture structural bone quality.

Introduction

Skeletal integrity is maintained through continuous bone modelling and remodelling; however, aging triggers a progressive decline in bone mass and critical alterations in bone microarchitecture, leading to osteopenia or osteoporosis [1, 2]. While osteoporosis is a systemic, often asymptomatic metabolic disease prevalent in postmenopausal women[1, 3], its impact extends beyond the axial skeleton to the

gnathic bones. Although official diagnosis relies on bone mineral density (BMD) T-scores [4], these systemic conditions significantly compromise the overall bone quality and trabecular patterns in the maxillofacial region, necessitating precise local assessment [1, 2].

While Dual-energy X-ray absorptiometry (DEXA) remains the clinical gold standard for quantifying BMD, it fails to capture the essential structural complexity and microarchitectural quality of bone [2, 5]. Beyond its cost and limited accessibility for large-scale screening, DEXA's focus on mineral mass overlooks early trabecular degradation, which typically precedes cortical bone loss [5]. Consequently, dental radiographs emerge as a powerful opportunistic screening tool; they allow for the early detection of these critical structural alterations, enabling a more comprehensive correlation with disease severity and progression at a fraction of the cost and systemic exposure[2, 6].

Beyond simple pixel intensity (PI) analysis, which has shown promising results in initial osteoporosis detection [7], texture analysis (TA) provides a superior quantitative imaging framework. By converting grey-level distributions into measurable features, TA reveals critical trabecular bone alterations that are not visually discernible, significantly enhancing diagnostic accuracy and prognostic assessment[8–10]. Characterizing bone structure through complex first- and second-order statistical methods is now a standard for objective skeletal quality evaluation [11]. Furthermore, the integration of artificial intelligence techniques allows these radiomic patterns to be leveraged for superior discriminative potential and predictive performance, bridging the gap between routine dental imaging and systemic fracture risk assessment[12, 13].

Despite growing interest, the clinical application of TA in dental imaging remains significantly limited for systemic bone assessment. Current literature focuses predominantly on the inferior mandibular border and fractal dimension analysis, often failing to integrate multi-site data, such as the maxillary tuberosity, or establish systematic correlations with clinically relevant variables. Based on this critical gap, the present study hypothesizes that radiomic features extracted from both the maxilla and mandible can objectively discriminate between normal bone status, osteopenia, and osteoporosis, providing essential diagnostic information when correlated with demographic and clinical variables. Accordingly, this study aims to evaluate the discriminative and predictive performance of TA applied to panoramic radiographs as an advanced auxiliary tool for comprehensive osteoporosis assessment.

Materials and Methods

This retrospective diagnostic accuracy study was approved by the local Research Ethics Committee (Protocol No. 7.330.010) and conducted in strict accordance with the Standards for Reporting Diagnostic Accuracy Studies (STARD) guidelines[14]. A convenience sample of panoramic radiographs from 123 female patients was initially screened from the institutional database. Eligibility criteria included adult women with established diagnoses of osteopenia, osteoporosis, or normal bone status (control). Patients with a history of necrotic bone removal, extensive maxillofacial surgery, medication-related osteonecrosis of the jaws (MRONJ), or other primary bone lesions (e.g., cysts or tumors) were

excluded. Furthermore, cases lacking accessible BMD records or those with insufficient trabecular area for region of interest (ROI) delineation in the maxillary tuberosity were removed from the final analysis.

After applying the eligibility criteria, 16 examinations were selected for the pilot study and 47 examinations for the main study, which were used for TA feature extraction. For the pilot study, the requirement of an available BMD record in the clinical chart was waived. All panoramic radiographs were acquired using a Kodak 9000C 3D™ device (Kodak Dental Systems, Carestream Dental LLC, Atlanta, GA, USA) with predefined acquisition parameters (70 kVp and 10 mA). All images were obtained for clinical purposes in accordance with the recommendations of the American Dental Association [15] and the OSTEODENT project guidelines [16].

Image assessment was performed in a dimly lit environment using a 15-inch monitor with a resolution of 1920 × 1080 pixels by a single examiner, an oral and maxillofacial radiologist experienced in radiographic image interpretation. An initial pilot phase (n = 16) was conducted to provide examiner training and calibration, refine methodological aspects, and validate the proposed workflow. This pilot sample was not included in the main analysis; however, all procedures strictly followed the same protocol applied in the main study. To ensure intra-examiner reliability, 20% of the sample was re-evaluated after 15 days, and the intraclass correlation coefficient (ICC) or Kappa value was calculated.

For the main study, 47 panoramic radiographs were allocated into three groups: a control group (CG) comprising 11 patients without bone disease; an osteopenia group (G1) including 25 patients with osteopenia; and an osteoporosis group (G2) including 11 patients with suspected osteoporosis. As this was a convenience sample, descriptive analysis and statistical power calculation were performed exclusively for the main study. Descriptive statistics were obtained using Jamovi software (version 2.3.8), yielding mean ± standard deviation values of 6.00 ± 3.32 for CG and G2, and 13.00 ± 7.36 for G1. Statistical power analysis was conducted using G*Power software (version 3.1; Heinrich Heine University, Düsseldorf, Germany), applying a two-sample t-test with a statistical power of 80% and a significance level (α) of 0.05. This analysis indicated a large effect size (Cohen's d = 1.226) and an estimated statistical power of 0.8267 for comparisons between the G1 and G2 groups.

After initial selection, the images were converted to bitmap format (.bmp) and exported to MaZda software (version 4.6, Technical University of Lodz, Poland; available at https://qmazda.p.lodz.pl/). During this process, images were downsampled to 8 bits/pixel to ensure full compatibility with the MaZda™ algorithmic framework for feature extraction. While this conversion may reduce the native dynamic range, image normalization was performed using the ± 3σ (sigma) method (mean ± three standard deviations) to minimize the influence of non-significant variations in brightness and contrast.

After initial selection, the images were converted to bitmap format (.bmp) and exported to MaZda™. During this process, images were downsampled to 8 bits/pixel to ensure full compatibility with the MaZda™ algorithmic framework for feature extraction. In each panoramic radiograph, regions of interest (ROIs) were semi-automatically delineated in a square shape within areas with a predominance of trabecular bone. A calibrated oral and maxillofacial radiologist manually adjusted the ROI positions to

ensure the strict exclusion of cortical plates, dental roots, and vascular canals, thereby enhancing textural purity. ROI 1 (1.932 pixels) was positioned in the maxillary tuberosity, whereas ROI 2 (2.300 pixels) was placed in the posterior mandibular region, distal to the last molar, corresponding to the retromolar trigone and excluding the external oblique line (Fig. 1).

For each included examination, an individual texture report was generated using MaZda™ software. Texture feature extraction was performed by selecting first-order and second-order statistical parameters, including histogram-based features and gradient features. Gray-level co-occurrence matrix (GLCM) and gray-level run-length matrix (GLRLM) features were computed across four directions (0°, 45°, 90°, and 135°) and four pixel distances, maintaining consistency with the 8 bits/pixel depth. Image normalization was performed using the ± 3σ method to minimize non-significant variations in brightness and contrast, compensating for the dynamic range reduction during downsampling. Following ROI delineation, quantitative data were extracted from the texture analysis reports, yielding a total of 230 texture features. Dimensionality reduction was subsequently performed using the Fisher coefficient (F), the probability of error and the average correlation coefficients (POE + ACC)[17–19], enabling the selection of the most relevant features and the identification of texture parameters with the highest discriminative potential among the analyzed groups.

All parameters identified by the Fisher coefficient and POE + ACC reports were considered study variables for both ROIs (ROI 1 and ROI 2) across all groups, resulting in a total of 36 TA features. First-order features, derived from histogram-based analysis, assess individual pixel gray-level values and represent global image characteristics[20]. Second-order features included those derived from the GLRLM and the GLCM. GLRLM features analyze consecutive runs of pixels with identical gray-level values in a specific direction and are sensitive to repetition patterns and the uniformity of homogeneous regions [21, 22]. In contrast, GLCM features describe the spatial relationship between pairs of pixels with defined gray-level values, enabling the assessment of local contrast, repetition, and textural heterogeneity [23]. The texture features evaluated in this study are detailed in Fig. 2 and Table 1.

Table 1
Definition of the parameters obtained by the texture analysis (TA) technique.

Category	Parameters	Descriptions
1st order	Percentis (Perc)	Describes the value of the highest gray level under which a certain percentage of pixels in the image is contained.
2nd order (GLCM)	Sum of squares (Sumofsqs)	Evaluates the dispersion of the distribution of gray levels.
2nd order (GLCM)	Sum of variance (Sumofvarnc)	Evaluates the dispersion in the distribution of the sum of gray levels in the image.
2nd order (GLCM)	Sum of the average (Sum averg)	Evaluates the mean distribution of the sum of gray levels.
2nd order (GLRLM)	Non-uniformity of running length (RLNonUni)	Measures the similarity in the length of the pixel runs.
2nd order (GLRLM)	Non-uniformity of the gray level (GLevNonU)	Measures the similarity of gray level values.

Source

Data obtained from texture analysis (TA).

Legend

GLCM = *Gray Level Co-occurrence Matrix* ; GLRLM = *Gray Level Run Length Matrix* .

Data processing and statistical analyses were performed using R software (version 4.5.1). The final sample comprised 47 patients (CG, G1, and G2), evaluated across a total of 42 variables, including one grouping variable, one identification variable (ID), four clinical variables—BMD, fracture history, menopausal status at diagnosis, and age category—and 36 texture features, of which 18 corresponded to ROI 1 and 18 to ROI 2.

Data normality was assessed for all variables. Group comparisons involving TA features were performed using one-way analysis of variance (ANOVA) for normally distributed data and the Kruskal–Wallis test when normality assumptions were violated, followed by appropriate post hoc and pairwise comparisons using the Mann–Whitney U test. Statistical significance was set at $p < 0.05$. Associations between TA features and clinical variables, like BMD, fracture status, menopausal status at diagnosis, and age category, were evaluated using Pearson's or Spearman's correlation coefficients, as appropriate. Given the binary nature of fracture status, comparisons involving this variable were conducted using the Mann–Whitney U test. To control for multiple testing and reduce false-positive findings, p-values were adjusted using the false discovery rate (FDR) method.

Machine learning (ML) models, including Random Forest, Logistic Regression, and Support Vector Machine (SVM), were applied to explore the discriminative and predictive performance of texture features. The dataset (n = 47) was randomly split into training (n = 32) and testing (n = 15) sets, with class imbalance in the training data addressed using the Synthetic Minority Over-sampling Technique (SMOTE). Model performance was evaluated using standard ML metrics, and multiclass receiver operating characteristic (ROC) analysis was performed for the Random Forest and Logistic Regression models.

Results

In the intergroup comparison, two variables met the assumptions for normality and were analyzed using one-way ANOVA, revealing significant differences for the 10th Percentile in both ROI 1 (p = 0.0129) and ROI 2 (p = 0.0058). The non-parametric Kruskal-Wallis test identified significant differences for the Sum Average (SumAverg) parameter in ROI 1 across four spatial positions. Subsequent post-hoc pairwise comparisons (Table 2) further clarified these findings: the SumAverg feature in ROI 1 significantly discriminated between the Control Group (CG) and the Osteopenia Group (G1), whereas the 10th Percentile effectively distinguished G1 from the Osteoporosis Group (G2) in both anatomical regions (ROI 1 and ROI 2).

Tabel 2. Results of post-hoc tests between the groups and texture analysis (TA) variables.

Variable	**Comparison**	***p*-value**
S (1,0) SumAverg_ROI1	GC vs G1	0,0148
S (2,0) SumAverg_ROI1	GC vs G1	0,0110
S (3,0) SumAverg_ROI1	GC vs G1	0,0100
S (4,0) SumAverg_ROI1	GC vs G1	**0,0054***
Perc 10% ROI1	G1 vs G2	0.0110
Perc 10% ROI2	G1 vs G2	**0,0037***

Legend

S = Position; SumAverg = Sum of the average; ROI1 = region of interest 1; ROI2 = region of interest 2; Perc = percentile; CG = Control Group; G1 = Osteopenia Group; G2 = Osteoporosis Group.

Note

Post-hoc tests performed by paired comparison and Mann-Whitney U. Values of **p < 0.05*** were considered statistically significant.

Regarding the correlation between texture parameters and clinical variables, Spearman and Pearson coefficients were calculated to assess the relationship between Radiomic features and BMD, age, and time since menopause onset. No statistically significant correlations were observed between the 36 selected texture variables and BMD ($p > 0.05$). Conversely, Spearman correlation analysis identified significant associations ($p < 0.05$) between specific Gray-Level Run-Length Matrix (GLRLM) parameters—RLNonUni and GLevNonU—and patient age, with these findings predominantly observed in ROI 2 (Table 3).

Table 3
Spearman's correlation between texture analysis (TA) parameters and the age variable.

Texture Variable	**Spearman's coefficient (ρ)**	***p-value***
Vertl_RLNonUni_ROI1	0,308	0,035
135dr_GLevNonU_ROI2	–0,360	0,013
45dgr_GLevNonU_ROI2	–0,363	0,012
Horzl_GLevNonU_ROI2	–0,347	0,017
Vertl_GLevNonU_ROI2	–0,360	0,013
Legend: ρ = Spearman's correlation coefficient; RLNonUni = non-uniformity of execution length and GLevNonU = non-uniformity of the level and gray extracted from the GLRLM (Gray-Level Run Length Matrix) matrix in the respective directions (horizontal, vertical, 45° and 135°). ROI1 = region of interest 1; ROI2 = region of interest 2;		
Note: Values of $p < 0.05$ were considered statistically significant. Positive correlations indicate a direct association between the variables, while negative correlations indicate an inverse association.		

Spearman correlation coefficients revealed statistically significant associations ($p < 0.05$) between Radiomic features and the age at menopause onset. These correlations were observed for parameters derived from the GLRLM across both anatomical regions (ROI 1 and ROI 2), as detailed in Table 4.

Table 4
Spearman's correlation between texture analysis (TA) variables and the patient's age at the time of initial diagnosis of menopause.

Matrix/Parameter	Texture Variable (ROI 1)	ρ	*p*-value	Textura Variable (ROI 2)	ρ	*p*-value
GLRLM – RLNonUni	135dr_RLNonUni	0,5070	0,0003	135dgr_RLNonUni	0,4041	0,0048
	45dgr_RLNonUni	0,4839	0,0006	45dgr_RLNonUni	0,3992	0,0054
	Horzl_RLNonUni	0,4762	0,0007	Horzl_RLNonUni	0,4030	0,0050
	Vertl_RLNonUni_	0,5018	0,0003	Vertl_RLNonUni	0,4264	0,0028
GLRLM – GLevNonU	135dr_GLevNonU	0,4228	0,0031	135dr_GLevNonU	0,4892	0,0005
	45dgr_GLevNonU	0,4128	0,0039	45dgr_GLevNonU	0,4849	0,0006
	Horzl_GLevNonU	0,4141	0,0038	Horzl_GLevNonU	0,4861	0,0005
	Vertl_GLevNonU	0,4062	0,0046	Vertl_GLevNonU	0,4739	0,0008

Legend: ρ = Spearman's correlation coefficient. GLRLM – RLNonUni = non-uniformity of execution length; GLRLM – GLevNonU = non-uniformity of the gray level;

Note: Values of $p < 0.05$ were considered statistically significant. Positive correlations indicate a direct association between the variables, while negative correlations indicate an inverse association.

The association between Radiomic features and the outcome variable 'fracture risk' was assessed using the Mann-Whitney U test, with no statistically significant differences observed ($p > 0.05$). Subsequently, the False Discovery Rate (FDR) adjustment was applied to control for multiple testing across all clinical variables (BMD, age, menopause onset, and fracture risk) and the 36 texture features. Following FDR correction, correlations involving BMD, fracture risk, and patient age no longer reached statistical significance. Conversely, the relationship between age at menopause onset and specific GLRLM parameters was confirmed as statistically significant ($p < 0.05$), validating these features as robust indicators of hormonal-related trabecular changes (Table 5).

Table 5
Results of the correction by the *False Discovery Rate* (FDR) method between the patient's age variable at the time of the initial diagnosis of menopause and the texture variables.

Texture Variable	**ROI 1**	**ROI 1**	**ROI 2**	**ROI 2**
	***p*-value gross**	***p*-value adjusted**	***p*-value gross**	***p*-value adjusted**
RLNonUni (Non-uniformity of running length 135°)	0,000	0,003	0,005	0,012
RLNonUni (45°)	0,001	0,003	0,005	0,012
RLNonUni (Horizontal)	0,001	0,003	0,005	0,012
RLNonUni (Vertical)	0,000	0,003	0,003	0,011
GLevNonU (Non-uniformity of the gray level 135°)	0,003	0,011	0,000	0,003
GLevNonU (45°)	0,004	0,012	0,001	0,003
GLevNonU (Horizontal)	0,004	0,012	0,001	0,003
GLevNonU (Vertical)	0,005	0,012	0,001	0,003
Legend: RLNonUni = Non-uniformity of execution length and GLevNonU = Non-uniformity of the level and gray extracted from the GLRLM (Gray-Level Run Length Matrix) matrix in the respective directions (horizontal, vertical, 45° and 135°). ROI1 = region of interest 1; ROI2 = region of interest 2;				
Note: Values with $p < 0.05$ were considered statistically significant after adjustment by the FDR method.				

Following the training and testing phases, multiclass Receiver Operating Characteristic (ROC) analysis was performed to evaluate the diagnostic performance of the Machine Learning (ML) classifiers. The Random Forest (RF) model demonstrated superior diagnostic capability, achieving Area Under the Curve (AUC) values exceeding 0.70 for all diagnostic categories (Control, Osteopenia, and Osteoporosis), indicating consistent discriminative power. Conversely, the Logistic Regression model exhibited limited performance, with AUC values falling below 0.60, suggesting an inability to effectively differentiate between the bone status groups (Figs. 3 and 4).

Legend

GC = Controle; G1 = Osteopenia; G2 = Osteoporosis; AUC = Area under the ROC Curve (*Receiver Operating Characteristic*).

Note

AUC values closer to 1 indicate greater discriminative ability. Values 0.5 = no discriminative power (equal to chance); from 0.6–0.7 = Poor discriminative power; 0.71–0.8 = Acceptable discriminative power; from 0.81–0.9 = Excellent discriminative power; > 0.91 = Exceptional discriminative power.

Legend

GC = Controle; G1 = Osteopenia; G2 = Osteoporosis; AUC = Area under the ROC Curve (*Receiver Operating Characteristic*).

Note

AUC values closer to 1 indicate greater discriminative ability. Values 0.5 = no discriminative power (equal to chance); from 0.6–0.7 = Poor discriminative power; 0.71–0.8 = Acceptable discriminative power; from 0.81–0.9 = Excellent discriminative power; > 0.91 = Exceptional discriminative power.

Machine Learning (ML) classifiers, specifically Random Forest (RF), Logistic Regression, and Support Vector Machine (SVM), were deployed to identify distinct trabecular patterns across the study groups. Model performance was rigorously benchmarked using specificity, sensitivity, F1-score, and confusion matrix analysis to determine classification accuracy. The SVM model demonstrated superior overall performance, particularly in the Osteopenia Group (G1), achieving a sensitivity of 0.875 and an F1-score of 0.700. This high sensitivity indicates a robust capability for the early detection of osteopenic changes. Confusion matrix analysis further validated the SVM's proficiency in G1, correctly classifying seven out of eight samples, although minor misclassifications persisted between the Control and Osteoporosis groups.

Discussion

The initial hypothesis of this study proposed that texture analysis (TA), associated with machine learning algorithms, would be capable of recognizing specific patterns of bone trabecular microarchitecture and, based on them, distinguishing between control, osteopenia, and osteoporosis groups. Based on this assumption, it was expected that Random Forest, Logistic Regression, and Support Vector Machine (SVM) models, trained with parameters extracted from texture analysis in panoramic radiographs, would present satisfactory performance in image classification, reflecting the degree of bone commitment of each group. Thus, the results were discussed in light of this hypothesis, considering not only the performance metrics of the models (specificity, sensitivity, and F1-score) but also the biological aspects related to bone remodeling and the methodological factors that may have influenced the discriminative and predictive capacity of the applied algorithms.

In the present study, the texture parameters obtained from panoramic radiographs were similar among the control, osteopenia, and osteoporosis groups, with no significant correlation with BMD or fracture. In contrast, an association was observed between texture variables and the age at menopause onset, suggesting the influence of hormonal status on the trabecular pattern. The differences in correlations with age reinforce the multifactorial nature of bone remodeling. Among the Machine Learning models, Random Forest presented better performance in group classification, while SVM detected changes only in patients with osteopenia, indicating that algorithm performance may have been influenced by data distribution and overlap.

In the correlation between the group variables and the 36 TA parameters, a statistically significant p-value was observed for the parameters perc10% and Sum Averg. The Percentile parameter is a first-order statistic. In the present study, the Percentile was defined at 10%, representing the fraction of pixels with higher luminosity, which possibly reflects the global brightness of the radiograph[24]. Percentile showed a statistically significant difference in both regions of interest (ROIs), indicating the ability to distinguish texture variations—or brightness—between the maxilla and the mandible in patients of groups G1 and G2. It is important to highlight that the standardization of ROIs, regarding both size and shape, can significantly interfere with TA results[25]. Given this finding, it is possible to infer that Percentile may be related to bone density values, even though it is not traditionally used to quantify texture specifically[20].

A statistically significant p-value was observed for SumAverage parameter in four positions only in ROI 1, indicating discriminative capacity only between patients in the GC and G1 groups. Thus, the SumAverg parameter may be more sensitive for detecting trabecular differences in the maxilla due to its specific structural characteristics. In a previous study, authors reported that in individuals with osteoporosis, there is greater loss of trabecular bone, with the mandible being structurally denser than the maxilla [5]. The findings of this study suggest, therefore, that the Sum Average parameter was useful for distinguishing the GC and G1 groups, possibly due to the greater heterogeneity of microtrabeculae in these patients, in contrast to the relative homogeneity observed in individuals in the G2 group. However, additional studies are necessary to confirm and deepen these observations.

In the correlation between the variables "Age at menopause onset" and the 36 TA parameters, a statistically significant p-value was observed for the parameters RLNonUni and GLevNonU, both from the GLRLM matrix, in all directions and regions (ROI 1 and 2), indicating a possible hormonal effect on the trabecular pattern. The Run Length Non-Uniformity (RLNonUni) parameter presented consistent variation for group G2, and the Grey Level Non-Uniformity (GLevNonU) parameter presented consistent variation between groups G1 and G2. These findings suggest that these parameters are sensitive to hormonal and microstructural changes, since patients with osteoporosis tend to exhibit loss of connectivity and thickness in their trabeculae, as well as thinners and more porous cortices. [26, 27]. The confirmation of robustness by the FDR technique reinforces the reliability of this analysis.

In the correlation between the age variable and the 36 TA parameters, a statistically significant p-value was observed for the parameters RLNonUni and GLevNonU. The RLNonUni parameter was significant only in the vertical direction and in ROI 1, suggesting less age influence, while GLevNonU was present in all directions and restricted to ROI 2, suggesting the influence of aging on the mandibular structure. The predominance of GLevNonU suggests sensitivity to bone microstructural changes, since elderly patients tend to have poorer bone quality, with bone loss occurring first in the trabecular bone, followed by the cortical bone. [28]CBCT studies corroborate these findings, showing high values of GLevNonU and RLNonUni in osteoporotic patients[26].

The analysis between the BMD variable and the 36 texture analysis (TA) variables revealed no statistically significant differences. In a study on bone density assessment, it was reported that the

DEXA exam evaluates only bone mineral density without providing information about bone microstructure [29]. On the other hand, the literature highlights that TA in digital images can be a promising tool to identify patterns of trabecular alteration[30], a finding corroborated in this study. In view of this, further investigations are recommended, with larger samples and different imaging modalities, to confirm and deepen these results.

In the stage where machine learning (ML) models were applied, superior performance of the Random Forest and SVM algorithms was observed. The Random Forest model proved capable of distinguishing and classifying the GC, G1, and G2 groups, demonstrating good discrimination capacity between classes. On the other hand, the SVM model presented better performance only in the G1 group (osteopenia), with lower specificity and sensitivity for the other groups. These results indicate that although texture analysis (TA) contributes to the identification of trabecular patterns, its isolated application may not be sufficient to consistently differentiate different degrees of bone commitment, and its association with other clinical and radiographic variables is recommended to improve the accuracy of predictive models.

Among the limitations of this study, the low repeatability of GLCM matrix texture parameters between the pilot and main studies stands out, reinforcing the need for additional investigations evaluating the applicability of this matrix in different contexts and its correlation with the pathophysiology of osteoporosis[31]. This divergence may be related to sample characteristics, such as sample size, acquisition protocols, and image reconstruction algorithms—factors that directly influence the reproducibility and repeatability of TA[32–34]. The sample used was a convenience sample, composed exclusively of female participants, and the sample power calculation was performed previously to ensure the statistical representativeness of the groups. Even so, future studies with a larger number of participants are recommended to improve the generalization of machine learning models and explore possible sexual and morphological differences in bone microarchitecture, since men and women present distinct trabecular patterns in cases of osteoporosis [32, 35].

Finally, considering the sensitivity of texture parameters to digital image processing, it is recommended that this study be reproduced using the same parameters in other analysis software to verify if the parameters that showed statistical significance maintain consistency and repeatability of results. This step is fundamental to validate the robustness and clinical applicability of texture analysis as a complementary tool in the screening and early prediction of osteoporosis, as suggested by [30]. Such efforts may contribute to the establishment of standardized protocols, expanding the reliability and diagnostic potential of the technique in future investigations.

Texture analysis (TA) proved to be a sensitive tool for identifying trabecular changes in panoramic radiographs of patients with osteopenia and osteoporosis, evidencing discriminative potential between different bone conditions. Although no significant correlation was observed with bone mineral density (BMD), TA allowed for the identification of structural differences associated with clinical variables, such as age and age at menopause onset. Regarding the "fracture risk" outcome, no association was

observed between texture parameters and this risk, indicating the need for additional studies that integrate clinical, biological, and structural variables to improve the understanding and predictive capacity of this outcome in populations with bone loss.

Declarations

This retrospective diagnostic accuracy study involving human participants was conducted in strict accordance with the ethical standards of the institutional and national research committees, the 1964 Helsinki Declaration and its later amendments or comparable ethical standards, and the Standards for Reporting Diagnostic Accuracy Studies (STARD) guidelines. The local Research Ethics Committee approved this study protocol (Approval No. 7.330.010).

Funding:

This study was supported by the project “Jovem Cientista do Nosso Estado” - Research Support Foundation of the State of Rio de Janeiro – FAPERJ – E-26/200.299/2023 (283749). The authors confirm independence from the sponsors.

Conflict of interest:

The authors declare that they have no conflict of interest.

References

1. Brandi ML. (2009) Microarchitecture, the key to bone quality. Rheumatology (United Kingdom) 48:. https://doi.org/10.1093/rheumatology/kep273
2. Alzubaidi MA, Otoom M. A comprehensive study on feature types for osteoporosis classification in dental panoramic radiographs. Comput Methods Programs Biomed. 2020;188. https://doi.org/10.1016/j.cmpb.2019.105301.
3. Lane NE. Epidemiology, etiology, and diagnosis of osteoporosis. Am J Obstet Gynecol. 2006;194. https://doi.org/10.1016/j.ajog.2005.08.047.
4. Amarnath SS, Kumar V, Das SL. Classification of Osteoporosis. Indian J Orthop. 2023;57:49–54.
5. White SC, Rudolph DJ. Alterations of the trabecular pattern of the jaws in patients with osteoporosis. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 1999;88:628–35. https://doi.org/10.1016/S1079-2104(99)70097-1.
6. Byung DL, White SC. Age and trabecular features of alveolar bone associated with osteoporosis. Oral Surgery, Oral Medicine, Oral Pathology. Oral Radiol Endodontology. 2005;100:92–8. https://doi.org/10.1016/j.tripleo.2004.11.020.
7. Anandan R, Ganesan CLK, Aniyan A K. Y. Strut and radio-morphometric analysis of mandibular trabecular structure in pre-and post-menopausal women to aid in the diagnosis of osteoporosis. J Oral Biol Craniofac Res. 2024;14:273–9. https://doi.org/10.1016/j.jobcr.2024.03.002.

8. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48:441–6. https://doi.org/10.1016/j.ejca.2011.11.036.
9. Ghasemi RMA M. Mandibular Trabecular Bone Analysis. Using Local Binary Pattern for Osteoporosis Diagnosis; 2019.
10. Taguchi A, Tanaka R, Kakimoto N, et al. Clinical guidelines for the application of panoramic radiographs in screening for osteoporosis. Oral Radiol. 2021;37:189–208.
11. Mayerhoefer ME, Materka A, Langs G, et al. Introduction to radiomics. J Nucl Med. 2020;61:488–95. https://doi.org/10.2967/JNUMED.118.222893.
12. Forghani R, Chatterjee A, Reinhold C, et al. Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT texture analysis with machine learning. Eur Radiol. 2019;29:6172–81. https://doi.org/10.1007/s00330-019-06159-y.
13. Muraoka H, Kaneda T, Ito K, et al. Machine learning approach using radiomics features to distinguish odontogenic cysts and tumours. Int J Oral Maxillofac Surg. 2025. https://doi.org/10.1016/j.ijom.2025.06.020.
14. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. Clin Chem. 2015;61:1446–52.
15. Hellstein JW, Adler RA, Edwards B, et al. Managing the care of patients receiving antiresorptive therapy for prevention and treatment of osteoporosis: Executive summary of recommendations from the American Dental Association Council on Scientific Affairs. J Am Dent Assoc. 2011;142:1243–51. https://doi.org/10.14219/jada.archive.2011.0108.
16. Devlin H, Karayianni K, Mitsea A, et al. Diagnosing osteoporosis by using dental panoramic radiographs: The OSTEODENT project. Oral Surgery, Oral Medicine, Oral Pathology. Oral Radiol Endodontology. 2007;104:821–8. https://doi.org/10.1016/j.tripleo.2006.12.027.
17. Materka A. (2004) Texture analysis methodologies for magnetic resonance imaging.
18. Szczypiński PM, Strzelecki M, Materka A, Klepaczko A. MaZda-A software package for image texture analysis. Comput Methods Programs Biomed. 2009;94:66–76. https://doi.org/10.1016/j.cmpb.2008.08.005.
19. Muraoka H, Ito K, Hirahara N, et al. Magnetic resonance imaging texture analysis in the quantitative evaluation of acute osteomyelitis of the mandibular bone. Dentomaxillofacial Radiol. 2022;51. https://doi.org/10.1259/DMFR.20210321.
20. Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. J Nucl Med. 2011;52:369–78. https://doi.org/10.2967/jnumed.110.082404.
21. Galloway MM. SHORT NOTE Texture Analysis. Using Gray Level Run Lengths; 1975.
22. Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: Review and opportunities. Med. Image Anal. 2014;18:176–96.
23. Haralick RM. (1979) Statistical and Structural Approaches to Texture.

24. Muraoka H, Kaneda T, Ito K, et al. Magnetic resonance imaging texture analysis for discrimination between malignant and benign lesions in the floor of the mouth. Oral Surg Oral Med Oral Pathol Oral Radiol. 2025;140:125–33. https://doi.org/10.1016/j.oooo.2025.02.009.
25. Otani T, Yoshida H, Sugawara D, et al. Prone position magnetic resonance imaging for the mandibular bone: enhancing image quality to perform texture analysis for medication-related osteonecrosis of the jaw and carcinoma of the lower gingiva. Oral Radiol. 2024;40:468–9.
26. Kawashima Y, Fujita A, Buch K, et al. Using texture analysis of head CT images to differentiate osteoporosis from normal bone density. Eur J Radiol. 2019;116:212–8. https://doi.org/10.1016/j.ejrad.2019.05.009.
27. Payne JB, Reinhardt RA, Nummikoski PV, Patil KD. Longitudinal Alveolar Bone Loss in Postmenopausal Osteoporotic/ Osteopenic Women.
28. Gulsahi A. Osteoporosis and jawbones in women. J Int Soc Prev Community Dent. 2015;5:263. https://doi.org/10.4103/2231-0762.161753.
29. Najafi M, Yousefi Rezaii T, Danishvar S, Razavi SN. (2023) Qualitative Classification of Proximal Femoral Bone Using Geometric Features and Texture Analysis in Collected MRI Images for Bone Density Evaluation. Sensors 23:. https://doi.org/10.3390/s23177612
30. Ranjanomennahary P, Ghalila SS, Malouche D, et al. Comparison of radiograph-based texture analysis and bone mineral density with three-dimensional microarchitecture of trabecular bone. Med Phys. 2011;38:420–8. https://doi.org/10.1118/1.3528125.
31. Meyer HJ, Hamerla G, Höhn AK, Surov A. CT texture analysis-correlations with histopathology parameters in head and neck squamous cell carcinomas. Front Oncol. 2019;9. https://doi.org/10.3389/fonc.2019.00444.
32. Humeau-Heurtier A. Texture feature extraction methods: A survey. IEEE Access. 2019;7:8975–9000.
33. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. Int J Radiat Oncol Biol Phys. 2018;102:1143–58. https://doi.org/10.1016/j.ijrobp.2018.05.053.
34. van Timmeren JE, Cester D, Tanadini-Lang S et al. (2020) Radiomics in medical imaging—how-to guide and critical reflection. Insights Imaging 11.
35. Nussi AD, de Castro Lopes SLP, De Rosa CS, et al. In vivo study of cone beam computed tomography texture analysis of mandibular condyle and its correlation with gender and age. Oral Radiol. 2023;39:191–7. https://doi.org/10.1007/s11282-022-00620-3.

Figures

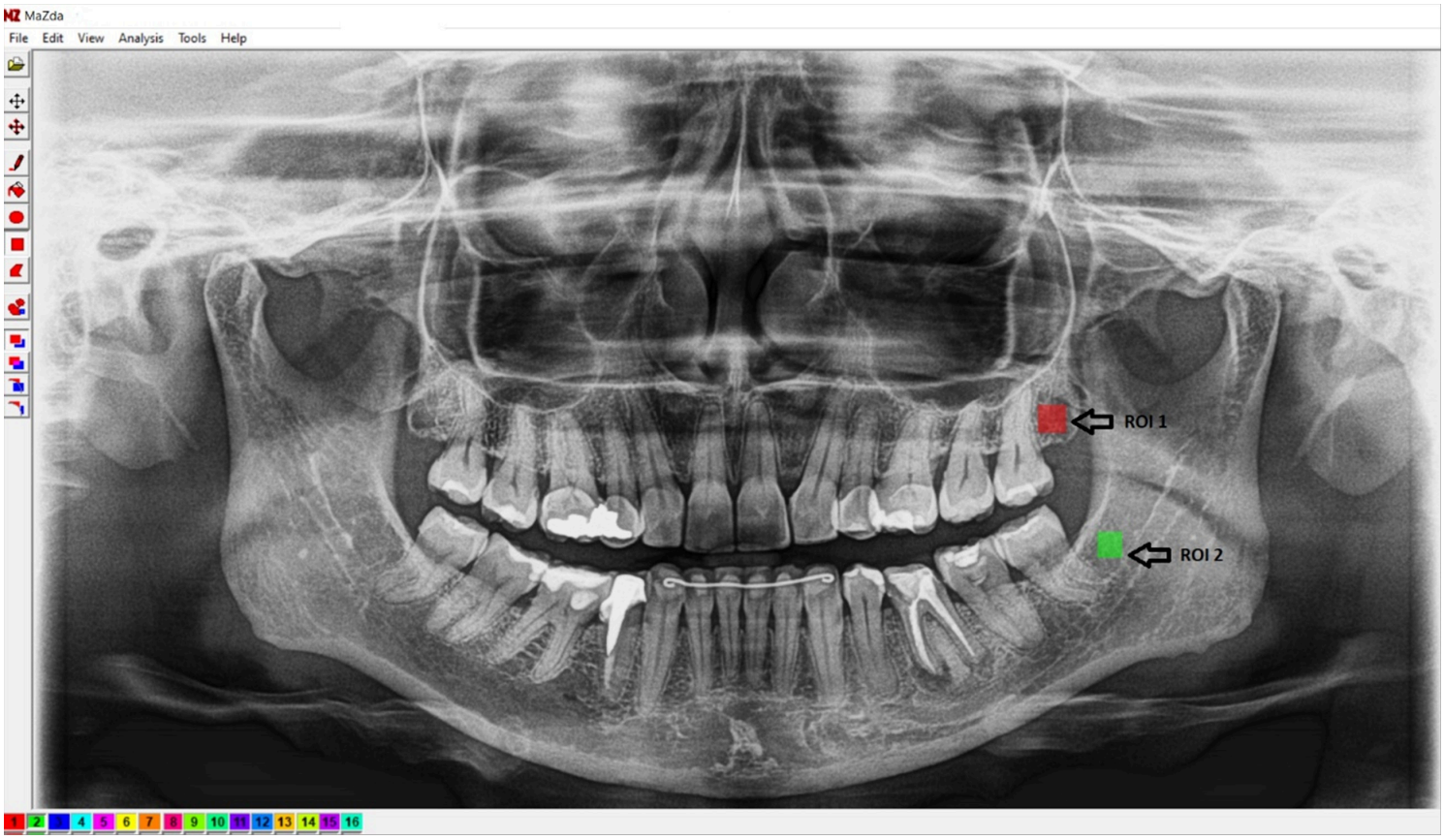


Figure 1

The MaZda™ software interface illustrates the selection of regions of interest (ROIs) in a panoramic radiograph. ROI 1 (in red) was positioned in the maxillary tuberosity region, and ROI 2 (in green) in the mandibular retromolar trigone region

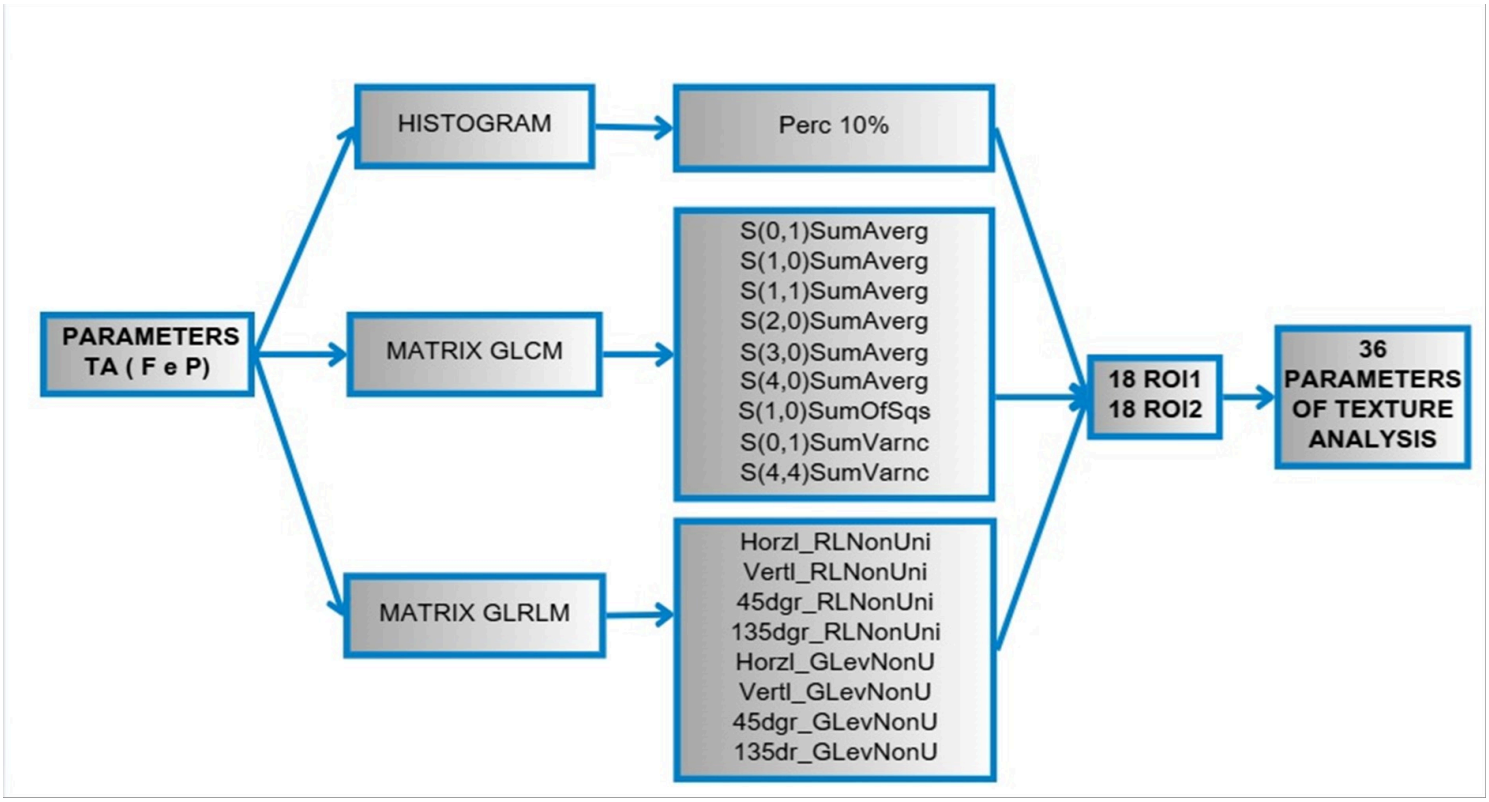


Figure 2

Flowchart illustrating the selection of variables obtained after analysis of Fisher's (F) and POE+ACC (P) reports

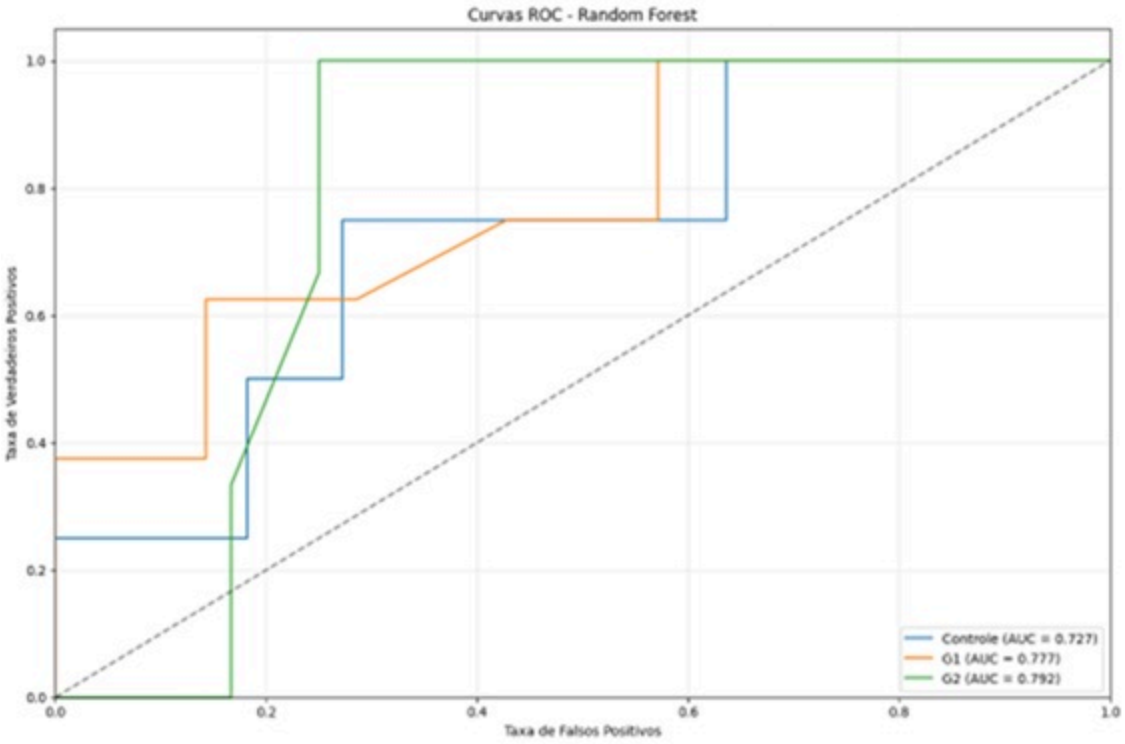


Figure 3

Multiclass ROC Curve for Random Forest Model

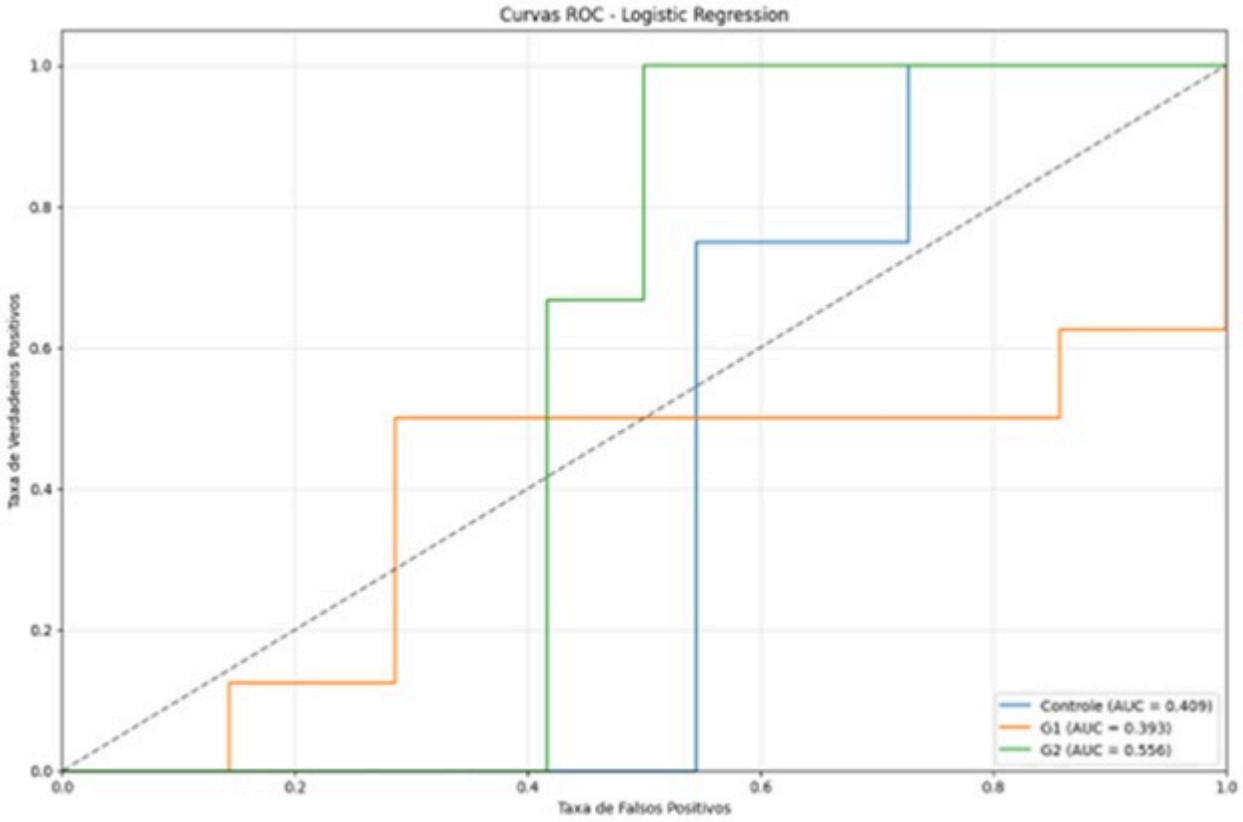


Figure 4

Multiclass ROC Curve for Logistic Regression