Online ISSN 2663-8290

Research Article

Academia Journal of Medicine
Vol. 9, No. 2, July–December 2026

Development of a Personalized Mandibular Bone Health Score using Explainable Artificial Intelligence and Cone-beam Computed Tomography Radiomics for Opportunistic Osteoporosis Screening

Neeraj Gupta[1*], Jyotirmayee Mishra[2], Siddharth Saurabh[3] and Rashmi Kiran Ekka[4]

[1]*PG student Bhabha University Bhopal(MP)Bhabha college of Dental Sciences,NH-12,Hoshangabad Road,Jatkhedi,Bhopal-462026(MP)*
[2]*PG student dept Of OMR,BCDS*
[3]*HOD Dept Of OMR,BCDS*
[4]*Prof Dept Of OMR,BCDS*

* Corresponding author. E-mail: dentistneerajgupta@gmail.com

https://doi.org/10.48165/ajm.2026.9.02.23

ARTICLE INFO

Keywords:
osteoporosis
bone mineral density
mandible
cone-beam computed tomography
CBCT
radiomics
artificial intelligence
explainable artificial intelligence
machine learning
opportunistic screening
bone health score

ABSTRACT

Background Osteoporosis is a systemic skeletal disorder characterized by reduced bone strength and increased susceptibility to fragility fracture. Although dual-energy X-ray absorptiometry (DXA) remains the principal method for assessing bone mineral density (BMD), a substantial proportion of individuals at risk are not identified before clinical presentation. Dental cone-beam computed tomography (CBCT) examinations are routinely acquired for several dentomaxillofacial indications and contain three-dimensional information about mandibular cortical and trabecular architecture. Such examinations may therefore provide an opportunity for opportunistic assessment of skeletal health without additional radiation exposure. Objective To develop a personalized Mandibular Bone Health Score (MBHS) using conventional mandibular measurements, CBCT-derived radiomic features and selected clinical variables within an explainable artificial intelligence framework for opportunistic identification of individuals at increased probability of low BMD. Methods This study proposes a retrospective diagnostic prediction-model design involving adults who have undergone clinically indicated mandibular CBCT and DXA examination within a predefined temporal interval. A standardized mandibular region of interest will be segmented from CBCT volumes. Conventional imaging variables, including mandibular cortical measurements and morphological characteristics, will be combined with standardized radiomic descriptors of mandibular bone architecture. Prespecified clinical variables will also be incorporated. Radiomic features will undergo reproducibility assessment, preprocessing, dimensionality reduction and nested feature selection. Clinical-only, conventional-imaging, radiomics-only and integrated models will be developed and compared. Nested cross-validation will be used to reduce optimism and prevent information leakage. The final model will estimate the probability of low BMD and transform this probability into a continuous MBHS ranging from 0 to 100. SHapley Additive exPlanations (SHAP) will be used to provide global and individual-level explanations. Expected outcome The proposed framework is expected to determine whether quantitative mandibular CBCT information provides incremental predictive value beyond conventional mandibular measurements and clinical variables. Model performance will be evaluated using discrimination, calibration and decision-curve analysis. Conclusion The MBHS represents a proposed explainable, opportunistic screening framework that could transform routinely acquired dental CBCT examinations into an additional source of information regarding systemic skeletal health. It is intended to identify patients who may benefit from formal osteoporosis assessment and is not intended to replace DXA. Independent external validation across populations, institutions and CBCT systems will be essential before clinical implementation.

INTRODUCTION

Osteoporosis is a major skeletal disorder characterized by compromised bone strength and increased risk of fragility fracture. Its clinical importance extends beyond bone mineral loss because fractures associated with osteoporosis can result in substantial morbidity, reduced quality of life and increased healthcare utilization. The disease may remain clinically silent for years, making identification of individuals at increased risk before the occurrence of a fragility fracture an important component of preventive healthcare.

Dual-energy X-ray absorptiometry is currently the principal imaging method used to quantify BMD and remains central to the clinical classification of osteoporosis. However, DXA is not routinely performed in every individual who may be at risk. Consequently, opportunities for earlier identification may be missed, particularly among individuals undergoing healthcare investigations for unrelated conditions.

Opportunistic screening has emerged as an approach to address this problem. Rather than acquiring an additional examination solely for screening, information contained within an imaging study performed for another clinical indication can be retrospectively or prospectively analyzed to identify unsuspected disease. This principle has been investigated extensively using computed tomography and other imaging modalities.

Dentistry provides a particularly interesting setting for opportunistic skeletal assessment. CBCT is now widely used for implant planning, assessment of impacted teeth, dentoalveolar surgery, orthodontic treatment planning, endodontic assessment and other maxillofacial applications. Consequently, a considerable number of individuals undergo three-dimensional mandibular imaging without any specific indication for skeletal evaluation.

The mandible may provide an indirect window into systemic skeletal status. Changes in mandibular cortical thickness, cortical morphology and trabecular organization have been investigated as potential indicators of reduced BMD. A systematic review and meta-analysis published in *Osteoporosis International* evaluated 64 studies examining mandibular radiomorphometric indices against DXA-derived BMD. Mandibular cortical width, mandibular cortical index and panoramic mandibular index were among the most frequently studied parameters. The review found potentially useful screening performance for some indices, while also emphasizing methodological limitations and insufficient accuracy for these measures to function as stand-alone diagnostic tests. [1] (Springer)

Conventional radiomorphometric analysis, however, represents only a small fraction of the information contained within a CBCT volume. Cortical thickness is a linear measurement, whereas bone quality is a multidimensional property involving density, architecture, cortical integrity, trabecular organization and spatial heterogeneity.

CBCT provides three-dimensional visualization of the mandibular cortex and trabecular compartment. Nevertheless, quantitative interpretation of CBCT gray values requires caution. Unlike conventional multidetector CT, CBCT gray values are influenced by scanner characteristics, field of view, scattered radiation, acquisition parameters and reconstruction algorithms. Consequently, CBCT gray values should not automatically be treated as equivalent to calibrated Hounsfield units. [2] (PubMed)

This limitation makes a purely density-based CBCT model problematic. A structural approach that evaluates spatial patterns rather than relying exclusively on absolute gray values may be more suitable for cross-platform investigation.

Radiomics provides such an approach. Radiomic analysis transforms medical images into quantitative descriptors representing intensity distributions, texture and spatial relationships. Standardized radiomic definitions have been developed through the Image Biomarker Standardisation Initiative (IBSI), improving reproducibility and facilitating comparison between studies. [3] (PMC)

In mandibular bone assessment, radiomics may capture subtle variations in trabecular organization and cortical architecture that are difficult to quantify using conventional manual measurements. The combination of radiomic information with established mandibular measurements may therefore provide a more comprehensive representation of mandibular bone characteristics.

Artificial intelligence and machine-learning methods further enable multiple imaging and clinical variables to be combined into an individualized prediction. However, model performance alone is insufficient for clinical translation. Models must also be appropriately calibrated, validated and interpretable.

Explainable artificial intelligence offers methods for understanding how individual predictors contribute to model predictions. SHAP provides an additive framework for assigning feature-attribution values to individual predictions and can be used to examine both global

model behavior and patient-specific explanations. [4] (NeurIPS Papers)

The proposed study therefore introduces a **Mandibular Bone Health Score (MBHS)** that integrates three information domains:

conventional mandibular structural measurements;

CBCT-derived radiomic features; and

relevant clinical variables.

The central premise is that the mandible may contain information that can be used to identify individuals at increased probability of low systemic BMD, even when the original CBCT examination was performed for a dental indication.

Importantly, the proposed score is not intended to diagnose osteoporosis. Instead, it is designed as an opportunistic screening and referral-support tool. A patient identified as being at increased risk would subsequently undergo appropriate clinical assessment and, where indicated, formal DXA evaluation.

STUDY AIM

To develop and internally validate an explainable artificial intelligence-based Mandibular Bone Health Score using mandibular CBCT radiomics, conventional mandibular imaging characteristics and selected clinical variables for opportunistic identification of individuals at increased probability of low BMD.

OBJECTIVES

Primary objective

To determine whether an integrated model combining clinical variables, conventional mandibular measurements and CBCT radiomic features can discriminate individuals with DXA-defined low BMD from those with normal BMD.

Secondary objectives

To quantify conventional mandibular cortical and trabecular characteristics.

To identify reproducible CBCT radiomic features associated with low BMD.

To compare clinical-only, conventional-imaging, radiomics-only and integrated models.

To assess the incremental predictive contribution of radiomics beyond conventional mandibular measurements.

To develop a continuous MBHS ranging from 0 to 100.

To evaluate discrimination and calibration of the proposed score.

To assess potential clinical utility using decision-curve analysis.

To identify the most influential predictors using SHAP analysis.

To establish a reproducible framework suitable for subsequent external validation.

HYPOTHESIS

The primary hypothesis is that an integrated model incorporating clinical variables, conventional mandibular imaging characteristics and CBCT radiomic features will demonstrate superior predictive performance for DXA-defined low BMD compared with conventional mandibular measurements alone.

The secondary hypothesis is that radiomic features will provide incremental information regarding mandibular bone architecture that is not captured by cortical thickness and conventional radiomorphometric assessment.

MATERIALS AND METHODS

Study design

The proposed investigation is a retrospective observational diagnostic prediction-model study.

The study will use existing clinically acquired CBCT examinations and corresponding DXA measurements.

The protocol is designed according to contemporary prediction-model methodology and will be reported according to TRIPOD+AI principles. TRIPOD+AI provides updated guidance applicable to prediction models developed using conventional regression or machine-learning methods and supersedes the original TRIPOD 2015 checklist. [5] (BMJ)

AI-specific diagnostic-accuracy elements will be addressed in accordance with STARD-AI where applicable. STARD-AI contains 40 reporting items and specifically emphasizes dataset construction, AI index-test description, evaluation procedures, bias, fairness and generalizability. [6] (Nature)

Potential sources of bias and applicability concerns will be assessed using PROBAST. [7] (DOI)

Study setting

The study will be conducted at [institution] using CBCT examinations acquired between [date] and [date].

Participants will be identified from institutional radiology or CBCT records.

Only CBCT examinations originally obtained for a clinically justified dental or maxillofacial indication will be eligible.

No additional radiation exposure will be required specifically for the study.

STUDY POPULATION

Inclusion criteria

Participants will be eligible if they meet all of the following:

age ≥18 years;

available mandibular CBCT examination;

CBCT examination obtained for a clinically indicated dental/maxillofacial purpose;

adequate visualization of the predefined mandibular ROI;

available DXA examination performed within a prespecified interval;

available BMD measurement from an anatomically appropriate DXA site;

adequate demographic and clinical information for the prespecified analysis.

Exclusion criteria

Participants will be excluded if they have:

previous mandibular segmental resection or reconstruction involving the ROI;

extensive mandibular pathology;

active destructive bone disease involving the ROI;

substantial metal artefact;

severe motion artefact;

inadequate CBCT field of view;

incomplete DICOM data;

technically inadequate DXA examination;

unavailable reference-standard information;

an interval between CBCT and DXA exceeding the prespecified limit.

REFERENCE STANDARD

DXA-derived BMD will serve as the reference standard.

For skeletal sites where T-score classification is clinically appropriate, participants will be classified according to accepted criteria:

Normal BMD: T-score ≥ −1.0

Low bone mass: T-score −2.5

Osteoporosis: T-score ≤ −2.5

The primary binary outcome will be:

Normal BMD versus low BMD.

Low BMD will include both osteopenia and osteoporosis.

A secondary analysis will evaluate:

Osteoporosis versus non-osteoporosis.

The exact skeletal site used as the reference measurement will be prespecified before model development.

CBCT IMAGE ACQUISITION

The following acquisition characteristics will be extracted from DICOM metadata whenever available:

CBCT manufacturer;

scanner model;

voxel size;

field of view;

tube voltage;

tube current;

exposure parameters;

reconstruction protocol;

acquisition date.

Scanner and acquisition variables will be retained as potential sources of technical variation.

Because CBCT gray values are not necessarily standardized across systems, absolute gray-value measurements will not be interpreted as direct equivalents of Hounsfield units. [2]

IMAGE PREPROCESSING

The proposed workflow will consist of:

DICOM acquisition → anonymization → image-quality assessment → orientation standardization → ROI localization → segmentation → standardized preprocessing → radiomic feature extraction → feature quality control → model development.

All preprocessing parameters will be prespecified.

Where resampling is required, a standardized isotropic voxel dimension will be selected and applied consistently.

Any preprocessing method capable of using information from the outcome labels will be performed only within the training dataset.

This restriction is essential to prevent information leakage.

MANDIBULAR REGION OF INTEREST

A standardized three-dimensional ROI will be selected within the mandibular body.

The preferred region will encompass both cortical and trabecular bone while avoiding:

tooth structures;

obvious restoration artefacts;

mandibular canals;

pathological lesions;

extraction sockets where possible.

A standardized anatomical region, such as the premolar/interforaminal region, will be considered because it offers relatively reproducible visualization of cortical and trabecular bone.

The final ROI definition will be established before analysis and documented using representative CBCT images.

SEGMENTATION

Segmentation will initially be performed using a standardized semi-automated or manual method.

Two trained observers will independently segment a predefined subset of examinations.

The subset will represent approximately 20–30% of the study population.

The following will be assessed:

intraobserver agreement;

interobserver agreement;

segmentation reproducibility;

feature stability.

Intraclass correlation coefficients will be calculated for continuous measurements and radiomic features where appropriate.

Features demonstrating inadequate reproducibility will be excluded.

CONVENTIONAL MANDIBULAR VARIABLES

Mandibular cortical thickness

Cortical thickness will be measured at prespecified anatomical locations using a standardized measurement protocol.

Measurements will be recorded in millimeters.

Where bilateral measurements are available, both sides will be retained initially.

The mean, minimum and side-to-side difference may be considered candidate variables.

Cortical morphology

Cortical morphology will be classified according to a predefined radiomorphometric classification.

The classification protocol will be standardized before analysis.

Bilateral asymmetry

For continuous bilateral measurements:

[\text{Absolute asymmetry}=|R-L|]

and

[\text{Relative asymmetry}=\frac{|R-L|}{(R+L)/2}]

may be calculated.

These variables will be evaluated only if bilateral measurements demonstrate adequate reproducibility.

CBCT RADIOMIC ANALYSIS

Radiomic analysis will follow standardized feature definitions based on the IBSI framework. [3]

The radiomic feature groups will include:

First-order statistics

Potential features include:

mean;

median;

variance;

standard deviation;

skewness;

kurtosis;

entropy;

minimum;

maximum;

percentile values.

Texture features

Potential texture families include:

gray-level co-occurrence matrix;

gray-level run-length matrix;

gray-level size-zone matrix;

gray-level dependence matrix;

neighboring gray-tone difference matrix.

Shape descriptors

Where appropriate, the ROI will be characterized by:

volume;

surface area;

compactness;

sphericity;

elongation.

Shape variables will be interpreted cautiously because the study ROI is anatomically standardized rather than representing an entire lesion.

RADIOMIC QUALITY CONTROL

The radiomic pipeline will include:

image-quality assessment;

segmentation reproducibility;

standardized voxel preprocessing;

intensity discretization according to a prespecified protocol;

feature extraction using standardized definitions;

reproducibility filtering;

variance filtering;

correlation reduction;

model-based feature selection.

Features with ICC <0.80 will be considered insufficiently reproducible for the primary analysis.

Radiomic feature extraction will be performed without knowledge of DXA outcome classification during the initial extraction process.

CLINICAL VARIABLES

The following variables will be considered when available:

age;

sex;

body mass index;

menopausal status where applicable;

history of fragility fracture;

family history of osteoporosis;

smoking;

alcohol consumption;

glucocorticoid exposure;

antiresorptive medication;

diabetes mellitus;

thyroid disease;

parathyroid disease;

chronic kidney disease;

other clinically relevant skeletal risk factors.

Clinical variables will be selected based on established biological relevance rather than univariable statistical significance alone.

MODEL ARCHITECTURE

Four models will be constructed.

Model A: Clinical model

Clinical predictors only.

Model B: Conventional mandibular model

Clinical predictors + conventional mandibular measurements.

Model C: Radiomics model

CBCT radiomic predictors.

Model D: Integrated MBHS model

Clinical predictors + conventional mandibular measurements + radiomic predictors.

This architecture will permit assessment of the incremental value of radiomics.

FEATURE SELECTION

Feature selection will occur within the development dataset only.

The proposed sequence is:

Quality filtering → reproducibility filtering → missingness assessment → variance filtering → correlation filtering → regularized/embedded selection → model fitting.

No feature selection will be performed using the independent test dataset.

During nested cross-validation, the complete feature-selection process will be repeated independently within each training fold.

This is intended to reduce optimism and prevent information leakage.

MACHINE-LEARNING METHODS

The following algorithms will be evaluated:

penalized logistic regression;

random forest;

gradient-boosted trees;

support vector machine.

The primary candidate model will be selected based on:

discrimination;

calibration;

stability;

interpretability;

computational reproducibility.

A complex model will not be preferred solely because it produces a higher training AUROC.

MANDIBULAR BONE HEALTH SCORE

The final model will produce an individualized predicted probability:

[P(\text{low BMD}|X)]

where (X) represents the patient's selected clinical, conventional imaging and radiomic predictors.

The proposed MBHS will be expressed as:

[\boxed{\text{MBHS}=100\times P(\text{low BMD}|X)}]

Thus:

0 represents very low predicted probability;

50 represents an estimated probability of 0.50;

100 represents very high predicted probability.

The score is intentionally continuous.

This avoids premature categorization and preserves information for future validation.

Potential categories may subsequently be defined as:

Low-risk MBHS

Intermediate-risk MBHS

High-risk MBHS

However, these categories will only be clinically proposed after validation.

No threshold will be described as a validated clinical cutoff in the initial development study.

EXPLAINABLE ARTIFICIAL INTELLIGENCE

SHAP analysis will be applied to the final model.

SHAP assigns feature-attribution values describing how individual variables contribute to a specific prediction. [4]

Two levels of interpretation will be provided.

Global explanation

The mean absolute SHAP value of each feature will be calculated.

Features will be ranked according to their overall contribution to model predictions.

Individual explanation

For each representative patient, the model output will be accompanied by:

predicted MBHS;

baseline prediction;

features increasing predicted risk;

features decreasing predicted risk.

SHAP values will be interpreted as explanations of model behavior and not as evidence of biological causality.

DATASET PARTITIONING

The preferred development framework will consist of:

Development dataset: 70%

Independent test dataset: 30%

The development dataset will undergo nested cross-validation.

The independent test dataset will remain completely isolated until the final locked model is evaluated.

If the available sample size is insufficient to support a reliable independent test set, repeated nested cross-validation will be used instead.

The chosen strategy will be justified according to sample size.

MISSING DATA

The amount and pattern of missing data will be documented.

Variables with substantial missingness will be evaluated before inclusion.

When appropriate, multiple imputation using chained equations will be considered.

Imputation will occur within model-development folds to avoid information leakage.

Complete-case analysis will be used only when justified by the pattern and amount of missingness.

STATISTICAL ANALYSIS PLAN

Continuous variables will be summarized as:

mean ± standard deviation when approximately normally distributed; or

median and interquartile range when non-normally distributed.

Categorical variables will be reported as frequencies and percentages.

Baseline characteristics will be described according to DXA-defined bone status.

Statistical significance will not be used as the sole criterion for predictor selection.

Primary performance measure

The primary performance measure will be:

Area under the receiver operating characteristic curve (AUROC).

The AUROC will be reported with 95% confidence intervals.

Secondary performance measures

The following will be reported:

sensitivity;

specificity;

positive predictive value;

negative predictive value;

accuracy;

F1 score;

area under the precision-recall curve.

Confidence intervals will accompany major performance estimates.

CALIBRATION

Calibration will be evaluated using:

calibration intercept;

calibration slope;

calibration plot;

Brier score.

A model demonstrating high discrimination but poor calibration will not be considered clinically adequate.

MODEL COMPARISON

The following comparisons will be prespecified:

Primary comparison

Model B versus Model D.

This will determine whether adding radiomic information to conventional mandibular assessment provides incremental predictive value.

Secondary comparisons

Model A versus Model B;

Model B versus Model C;

Model B versus Model D;

Model C versus Model D.

AUROC differences will be evaluated using appropriate paired methods because predictions are generated for the same participants.

INTERNAL VALIDATION

Nested cross-validation will be used for:

feature selection;

hyperparameter optimization;

model comparison.

Bootstrap resampling will be considered for optimism correction and confidence intervals.

All preprocessing, imputation, feature selection and model fitting steps will occur within the resampling framework.

DECISION-CURVE ANALYSIS

Decision-curve analysis will be used to estimate the net clinical benefit of MBHS-guided referral.

The model will be compared with:

treat-all strategy;

treat-none strategy;

conventional mandibular assessment.

The analysis will focus on clinically plausible probability thresholds at which referral for formal skeletal assessment would be considered.

The objective is not to demonstrate that MBHS diagnoses osteoporosis, but to determine whether using the score could theoretically improve the selection of patients for further evaluation.

SAMPLE-SIZE CONSIDERATIONS

Sample size will be determined before final model development.

The calculation will account for:

expected prevalence of low BMD;

anticipated model performance;

number of effective candidate predictors;

expected degree of regularization;

desired precision of performance estimates;

anticipated missingness;

need for internal validation.

Because radiomic analysis can generate a large number of candidate variables, the study will not rely on a simplistic fixed events-per-variable rule.

Where the available sample size is limited, dimensionality reduction and penalized models will be prioritized.

The final sample-size justification will be documented transparently.

BIAS CONTROL

Potential biases will be addressed at each stage.

Selection bias

Consecutive eligible examinations will be preferred rather than convenience sampling.

Verification bias

All participants included in model development should have the same reference-standard definition.

Measurement bias

Standardized CBCT measurements and reproducibility assessment will be used.

Spectrum bias

Participant characteristics will be reported to determine whether the cohort represents the intended screening population.

Data leakage

All preprocessing and feature selection will be performed inside the training data.

Overfitting

Nested validation, feature reduction and model regularization will be employed.

Algorithmic bias

Model performance will be examined across relevant demographic and technical subgroups when sample size permits.

ETHICAL CONSIDERATIONS

The study protocol will be submitted to the Institutional Ethics Committee of [institution].

Because the investigation uses previously acquired clinically indicated CBCT and DXA examinations, no additional radiation exposure will be administered.

Patient identifiers will be removed before analysis.

Data will be stored using secure institutional procedures.

A waiver of informed consent will be requested for retrospective analysis when permitted by institutional regulations.

The study will not alter clinical treatment decisions during the development phase.

PROPOSED CLINICAL WORKFLOW

The intended future clinical workflow is:

Routine dental CBCT

↓

Automated mandibular ROI identification

↓

Radiomic + conventional feature extraction

↓

MBHS calculation

↓

Low / intermediate / high predicted risk

↓

Clinical review

↓

Referral for osteoporosis assessment when appropriate

↓

DXA confirmation according to clinical guidelines

The proposed workflow deliberately places clinical review between MBHS output and formal osteoporosis assessment.

EXPECTED CONTRIBUTION OF THE STUDY

The proposed MBHS is intended to address three limitations of conventional mandibular osteoporosis screening.

First, conventional radiomorphometric assessment relies on a small number of measurements and can be observer-dependent.

Second, simple CBCT gray-value approaches are limited by interscanner and acquisition variability. [2]

Third, complex radiomic models may be difficult to interpret without an explainability framework.

The proposed approach addresses these limitations by combining structural imaging information with radiomic analysis and explainable prediction.

DISCUSSION

The central premise of this study is that routinely acquired mandibular CBCT examinations may contain clinically relevant information about systemic skeletal status that is currently underused.

Previous work has demonstrated an association between mandibular radiomorphometric characteristics and reduced BMD. However, the evidence does not support treating a single mandibular measurement as a diagnostic substitute for DXA. The recent systematic review and meta-analysis by Heuchert et al. identified mandibular cortical width and other indices as potentially useful screening markers but also highlighted methodological limitations and variability between studies. [1] (Springer)

The proposed MBHS therefore differs conceptually from a conventional mandibular index. Rather than asking whether one measurement is below a predefined threshold, the model integrates multiple complementary characteristics.

The first component is conventional mandibular morphology. Cortical thickness and cortical integrity remain clinically interpretable and provide an understandable baseline against which more complex models can be compared.

The second component is CBCT radiomics. Trabecular bone is spatially heterogeneous, and this organization may not be adequately represented by a single thickness measurement. Texture descriptors may capture differences in spatial arrangement that are difficult to identify visually.

However, radiomic analysis of CBCT requires particular methodological caution. CBCT gray values vary according to acquisition and reconstruction characteristics and cannot simply be treated as equivalent to standardized CT attenuation values. [2] The proposed model therefore emphasizes structural and texture characteristics rather than presenting CBCT gray values as direct measurements of mineral density.

The third component is clinical information. Skeletal health is influenced by age, sex, body composition, hormonal status, medication exposure and systemic disease. Imaging should therefore be considered in conjunction with relevant clinical information rather than as an isolated biomarker.

Explainability is another central component. A model that simply produces a risk percentage may be difficult for clinicians to trust or evaluate. SHAP provides a framework for examining how individual predictors contribute to model predictions. [4] This could allow the future system to provide a clinically understandable explanation alongside the MBHS.

For example, a future output could state that an individual's predicted risk was influenced primarily by reduced cortical thickness, specific trabecular texture characteristics and age. Such an explanation would not imply causality, but it could make the screening result more transparent.

The clinical purpose of MBHS is deliberately limited. It is not proposed as an alternative to DXA. Instead, it would function as an opportunistic case-finding mechanism. Patients who are flagged could subsequently undergo appropriate clinical assessment and formal skeletal evaluation.

This distinction is particularly important because the consequences of false-positive screening and false-negative screening are different. A screening system should therefore be evaluated not only according to AUROC but also according to calibration and potential clinical utility.

Decision-curve analysis will consequently form an important part of the proposed evaluation. A model with excellent statistical discrimination but little practical benefit should not automatically be considered clinically useful.

STRENGTHS

The proposed study has several strengths.

First, it uses routinely acquired CBCT examinations and therefore does not require additional radiation exposure for the proposed screening mechanism.

Second, it integrates conventional and quantitative imaging information rather than replacing clinically interpretable measurements with a purely black-box model.

Third, radiomic reproducibility is incorporated before final feature selection.

Fourth, nested validation is used to reduce overfitting and information leakage.

Fifth, calibration and decision-curve analysis are included alongside discrimination measures.

Sixth, SHAP analysis provides a mechanism for patient-level and global model interpretation.

Finally, the framework is designed according to contemporary prediction-model and AI reporting standards. TRIPOD+AI specifically addresses prediction models developed using statistical or machine-learning methods, while STARD-AI provides additional guidance for AI-based diagnostic-accuracy research. [5,6] (BMJ)

LIMITATIONS

Several limitations should be recognized.

The retrospective nature of the proposed study may introduce selection bias because individuals with both dental CBCT and DXA examinations may not represent the general population undergoing dental imaging.

CBCT acquisition heterogeneity may affect radiomic features. Scanner manufacturer, reconstruction algorithm, voxel size and field of view can introduce systematic differences.

Mandibular pathology may also influence radiomic characteristics independently of systemic BMD. Periodontal disease, severe tooth loss, previous surgery, inflammatory changes and local bone lesions may therefore act as confounding factors.

The large number of candidate radiomic variables creates a substantial risk of overfitting, particularly if the sample size is modest. Rigorous feature reduction and nested validation are therefore essential.

Finally, an internally validated model cannot be assumed to generalize to other populations. External validation across institutions and CBCT platforms will be required.

FUTURE RESEARCH

Future research should proceed through sequential stages.

Phase I — Development

Derive the MBHS from a well-characterized retrospective cohort.

Phase II — Internal validation

Evaluate model optimism, calibration and stability using nested resampling.

Phase III — External validation

Test the locked model in independent populations and across different CBCT systems.

Phase IV — Prospective clinical validation

Determine whether MBHS can identify patients who would otherwise remain unrecognized as candidates for osteoporosis assessment.

Phase V — Clinical utility study

Determine whether MBHS-guided referral improves osteoporosis case finding without causing unacceptable unnecessary referrals.

Phase VI — Longitudinal evaluation

Investigate whether MBHS predicts clinically relevant outcomes such as future fragility fracture.

CONCLUSION

The proposed Mandibular Bone Health Score represents a novel framework for transforming routinely acquired dental CBCT examinations into an opportunistic source of skeletal-health information.

By integrating conventional mandibular morphology, CBCT radiomic characteristics and selected clinical predictors within an explainable artificial intelligence framework, the proposed model aims to identify individuals who may have an increased probability of low BMD.

The MBHS is intended to complement—not replace—established osteoporosis assessment. DXA remains the reference standard for formal BMD evaluation.

If successfully validated, the approach could provide a low-burden pathway for identifying individuals who might otherwise remain unscreened, particularly when CBCT imaging has already been obtained for an unrelated dental indication.

Table 1: Candidate predictor domains

Domain	Candidate variables
Demographic	Age, sex
Anthropometric	BMI
Clinical	Fragility fracture, family history, smoking, alcohol
Medication	Glucocorticoids, antiresorptive therapy
Systemic	Diabetes, thyroid/parathyroid disease, renal disease
Conventional CBCT	Cortical thickness, cortical morphology, bilateral asymmetry
Radiomics	First-order, GLCM, GLRLM, GLSZM, GLDM, NGTDM
Technical	Scanner, voxel size, field of view, reconstruction protocol

However, clinical implementation should occur only after rigorous internal and external validation, assessment of calibration and clinical utility, evaluation across CBCT platforms and prospective assessment of its effect on osteoporosis case finding.

PROPOSED TABLES

Table 1, Table 2 and Table 3 mention below

PROPOSED FIGURES

Figure 1. Conceptual framework of opportunistic mandibular osteoporosis screening.

Routine dental CBCT is analyzed without additional radiation exposure, followed by mandibular segmentation, quantitative feature extraction, MBHS calculation and risk-based clinical referral.

Figure 2. Standardized mandibular ROI.

Representative CBCT views demonstrating the anatomical region selected for conventional measurement and radiomic analysis.

Figure 3. Radiomic feature-extraction pipeline.

Schematic representation of preprocessing, segmentation, discretization, radiomic extraction, reproducibility filtering and feature selection.

Figure 4. MBHS artificial intelligence architecture.

Integration of clinical variables, conventional mandibular measurements and CBCT radiomics into the final predictive model.

Figure 5. Explainable AI output.

Illustration of a hypothetical patient-level MBHS prediction with SHAP-based feature contributions.

Figure 6. Proposed clinical pathway.

CBCT ⟶ MBHS ⟶ clinical review ⟶ appropriate osteoporosis assessment ⟶ DXA confirmation.

Figure 7. Model-validation framework.

Development cohort ⟶ nested cross-validation ⟶ locked model ⟶ independent external validation.

REPORTING FRAMEWORK

The completed validation study will be reported according to:

TRIPOD+AI: primary framework for clinical prediction-model development and evaluation. [5] (BMJ)

STARD-AI: applicable diagnostic-accuracy reporting elements, particularly dataset description, AI index-test methodology, evaluation and bias/fairness considerations. [6] (Nature)

PROBAST: assessment of risk of bias and applicability of the prediction model. [7] (DOI)

IBSI: standardization of radiomic feature definitions and extraction. [3] (PMC)

DECLARATIONS

Ethics approval

To be obtained from the Institutional Ethics Committee of [institution] before commencement of retrospective data analysis.

Consent for participation

A waiver of informed consent will be requested where permitted for retrospective analysis of anonymized clinical data.

Consent for publication

Not applicable to anonymized retrospective data unless individual patient images are identifiable.

Funding

[Insert funding source / No external funding.]

Conflict of interest

The authors declare **[no conflict of interest / insert relevant disclosures].**

Data availability

Data availability will be determined by institutional policy and applicable ethical and privacy regulations. Individual-level CBCT and DXA data will not be publicly released where such release could compromise participant confidentiality.

Table 2: Proposed model architecture

Model	Predictors	Purpose
Model A	Clinical	Clinical baseline
Model B	Clinical + conventional CBCT	Conventional imaging benchmark
Model C	Radiomics	Quantitative imaging model
Model D	Clinical + conventional CBCT + radiomics	Final integrated MBHS

Table 3: Planned model-performance measures

Domain	Measure
Discrimination	AUROC
Classification	Sensitivity, specificity
Predictive value	PPV, NPV
Overall performance	Brier score
Calibration	Intercept, slope, calibration plot
Classification balance	F1 score
Imbalanced outcome assessment	AUPRC
Clinical utility	Decision-curve analysis
Explainability	SHAP values

Author contributions

[Insert individual contributions according to the CRediT taxonomy.]

PROTOCOL STATUS

This manuscript describes a proposed conceptual and methodological framework. No participant-level results, diagnostic-performance estimates or model-derived feature rankings are reported at this stage.

Numerical performance estimates should be introduced only after completion of the prespecified analysis of the study dataset.

REFERENCES

Heuchert, J., Kozieł, S., & Spinek, A. E. (2024). Radiomorphometric indices of the mandible as indicators of decreased bone mineral density and osteoporosis: Meta-analysis and systematic review. *Osteoporosis International, 35*(3), 401–412. https://doi.org/10.1007/s00198-023-06949-7

Pauwels, R., Jacobs, R., Singer, S. R., & Mupparapu, M. (2015). CBCT-based bone quality assessment: Are Hounsfield units applicable? *Dentomaxillofacial Radiology, 44*(1), 20140238. https://doi.org/10.1259/dmfr.20140238

Zwanenburg, A., Vallières, M., Abdalah, M. A., Aerts, H. J. W. L., Andrearczyk, V., Apte, A., et al. (2020). The Image Biomarker Standardization Initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology, 295*(2), 328–338. https://doi.org/10.1148/radiol.2020191145

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 30* (pp. 4765–4774). Curran Associates.

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., et al. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, e078378. https://doi.org/10.1136/bmj-2023-078378

Sounderajah, V., Guni, A., Liu, X., Collins, G. S., Karthikesalingam, A., Markar, S. R., et al. (2025). The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. *Nature Medicine, 31*, 3283–3289. https://doi.org/10.1038/s41591-025-03953-8

Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., et al. (2019). PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine, 170*(1), 51–58. https://doi.org/10.7326/M18-1376

Moons, K. G. M., Wolff, R. F., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., et al. (2019). PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. *Annals of Internal Medicine, 170*(1), W1–W33. https://doi.org/10.7326/M18-1377

How to Cite this Article: Gupta N., Mishra J., Saurabh S., Ekka R.K.. (2026). Development of a Personalized Mandibular Bone Health Score using Explainable Artificial Intelligence and Cone-beam Computed Tomography Radiomics for Opportunistic Osteoporosis Screening. *Academia Journal of Medicine*, ***9***(2), 130-140. https://doi.org/10.48165/ajm.2026.9.02.23