**Dentomaxillofacial Radiology (2023) 52**, 20220225

birpublications.org/dmfr

SYSTEMATIC REVIEW

# Radiomics in bone pathology of the jaws

[1]Glaucia Nize Martins Santos, [1]Helbert Eustáquio Cardoso da Silva, [2]Filipe Eduard Leite Ossege, [1]Paulo Tadeu de Souza Figueiredo, [1]Nilce de Santos Melo, [1]Cristine Miron Stefani and [1]André Ferreira Leite

*[1]Dentistry Department, Faculty of Health Science, University of Brasília, Brasilia, Brazil; [2]Mechanical Engineering Department, Faculty of Technology, University of Brasília, Brasília, Brazil*

**Objective:** To define which are and how the radiomics features of jawbone pathologies are extracted for diagnosis, predicting prognosis and therapeutic response.
**Methods:** A comprehensive literature search was conducted using eight databases and gray literature. Two independent observers rated these articles according to exclusion and inclusion criteria. 23 papers were included to assess the radiomics features related to jawbone pathologies. Included studies were evaluated by using JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies.
**Results:** Agnostic features were mined from periapical, dental panoramic radiographs, cone beam CT, CT and MRI images of six different jawbone alterations. The most frequent features mined were texture-, shape- and intensity-based features. Only 13 studies described the machine learning step, and the best results were obtained with Support Vector Machine and random forest classifier. For osteoporosis diagnosis and classification, filtering, shape-based and Tamura texture features showed the best performance. For temporomandibular joint pathology, gray-level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), first-order statistics analysis and shape-based analysis showed the best results. Considering odontogenic and non-odontogenic cysts and tumors, contourlet and SPHARM features, first-order statistical features, GLRLM, GLCM had better indexes. For odontogenic cysts and granulomas, first-order statistical analysis showed better classification results.
**Conclusions:** GLCM was the most frequent feature, followed by first-order statistics, and GLRLM features. No study reported predicting response, prognosis or therapeutic response, but instead diseases diagnosis or classification. Although the lack of standardization in the radiomics workflow of the included studies, texture analysis showed potential to contribute to radiologists' reports, decreasing the subjectivity and leading to personalized healthcare.
*Dentomaxillofacial Radiology* (2023) **52**, 20220225. doi: 10.1259/dmfr.20220225

**Cite this article as:** Santos GNM, da Silva HEC, Ossege FEL, Figueiredo PTS, Melo NS, Stefani CM, et al. Radiomics in bone pathology of the jaws. *Dentomaxillofac Radiol* (2023) 10.1259/dmfr.20220225.

**Keywords:** Radiomics; Jaws; Texture analysis; Bone; Systematic Review

## Introduction

Radiomics is the process of extracting quantitative information from radiological images, and is designed to develop decision-support tools, taking a central role in the context of personalized precision medicine. Mining data from digital images and combine them with other patient characteristics promises to increase precision in diagnosis, assessment of prognosis, and prediction of therapy response[1–5].

Radiomic analysis begins with the choice of a high-quality imaging protocol. Such imaging methods include

Correspondence to: Dr Glaucia Nize Martins Santos, E-mail: nize.gal@gmail.com

Received 29 June 2022; revised 02 September 2022; accepted 02 October 2022; published online 16 November 2022

X-ray, CT, MRI, Nuclear Medicine, Positron Emission Tomography and Ultrasound[3]. After a prediction target is well stablished, regions that may contain value information, the region of interest (ROI) or volume of interest (VOI), are identified from these images. Then, they are segmented manually or computer-aided, and quantitative features based on texture, shape and gray level statistics within images are extracted to generate a report, which can be combined with demographic, clinical or genomic information from a database. These data are then mined to develop diagnostic, predictive or prognostic models to outcomes of interest[1,3,5].

The main part of radiomics is the process of analyzing texture, also known as feature extraction[6], and it is classified in two types: semantic features, which are commonly used in radiologic evaluation, and agnostic or non-semantic features, which comprise quantitative descriptors that are mathematically extracted from radiologic images, grounded on shape, intensity, texture, and filtering information[1,3,4,7]. This stage produces a large number of radiomics features and only the "highly" informative ones are selected based on the user defined criteria, developing the so-called radiomics signature[2,3].

The concept of radiomics has most broadly been applied in the field of oncology[7], since digitally encrypted medical images hold information related to tumor pathophysiology, which can be transformed into mineable quantitative high-dimensional data[5]. Its potentials include the differential diagnosis between neoplasms, correlation with molecular biology and genomics, the prediction of survival and the evaluation of the response to treatment[8]. Many other medical specialties are benefiting from radiomics studies, such as neuroimaging, cardiac imaging, and gastroenterology[2]. Although the literature is still scarce in dentistry field, the possibilities and perspectives of the use of radiomics-based investigations combined with the oral radiologists' analyses are growing through the years. A review of the state-of-the-art of using radiomics and machine learning (ML) for imaging in oral healthcare was recently described, which showed image segmentation and optimization applications, and pathology detection, classification and diagnosis implementations[4].

Regarding jawbone alterations, it is challenging to differentiate imaging lesions within the maxilla and mandible due to the appearance similarity among a wide variety of pathologies, mainly lesions from odontogenic and osseous sources. The response of the cancellous and cortical bone to pathologic stimulus can be expressed by osteolytic or osteoblastic behavior. Thus, most lesions within the jaws are possibly classified as lytic, sclerotic, or a mixture of both[9–11]. Due to this large scenario of the dentomaxillofacial imaging, the goal of this study is to define which are the radiomics features of bone pathologic alterations of the jaws and how they are extracted for diagnosis, predicting prognosis and therapeutic response. In this systematic review, radiomic workflow with emphasis in feature analysis was investigated related to bone changes of the jaws.

## Methods

*Study design*

A systematic review that analyzed radiomic features was performed to answer the question: "In dentomaxillofacial imaging, which are the radiomics features of bone pathologic changes of the jaws and how are they extracted for diagnosis, predicting prognosis and therapeutic response?"

*Eligibility criteria*

The PFO (Participants, Factors and Outcomes) acronym was used to define inclusion and exclusion criteria. As inclusion criteria, dentomaxillofacial imaging of bone alterations of the jaws (P) and radiomic analysis (F) were evaluated. Lastly, the outcomes measured (O) were feature extraction and their usefulness in detecting or classifying bone alterations of the jaws (sensitivity, specificity, ROC, AUC)..

*Inclusion criteria:* All published studies using radiomics in bone pathologic alterations of the jaws, that reported partially or completely the radiomics workflow, including feature extraction parameters.

*Exclusion criteria:* The following exclusion criteria were applied:

(1) Reviews, editorials, letters, personal opinions, book chapters and conference abstracts;
(2) Ultrasound imaging and elastography applied to the maxillofacial region;
(3) Scanned images;
(4) Non-human studies;
(5) Maxillofacial trauma, peripheric lesions with bone destruction, anatomic variations, metastatic lesions and abnormalities;
(6) Testing cases not based in imaging exams;
(7) Studies that merely differentiated healthy from pathological areas.

*Information sources and search strategy*

Individual search strategies for each of the following bibliographic databases were developed: Embase, IEEE Xplore, Lilacs, Livivo, PubMed, Science Direct, Scopus and Web of Science. The following terms: ("Intraoral radiography" OR "Extraoral radiography" OR "Dental panoramic radiographs" OR "Cone Beam Computed Tomography" OR "Magnetic Resonance Image" OR "PET-CT Scan" OR "Multislice Computed Tomography") AND ("osseous pathology" OR "bone pathology") AND (Jaws OR "Mandibular bones" OR "Maxillary Bones" OR "Alveolar bone") AND (Radiomics OR "Texture Analysis" OR "feature extraction") and their synonyms were used to develop the search strategies. A gray literature search was taken using

Google Scholar, JSTOR and ProQuest (Supplementary Material 1). The end search date was January 12, 2022 across all databases. Manual searches of reference lists were performed on relevant articles, theses and dissertations.

All references were managed by the software Mendeley®, Elsevier and duplicate records were removed.

*Study selection*

The study selection was completed in two phases. In Phase I, two reviewers (GNMS and HECS) independently reviewed the titles and abstracts of all identified electronic database records. Those not meeting the inclusion criteria were discarded. In Phase II, the same reviewers applied the inclusion criteria to the full text of the articles. Both phases were completed using the Rayyan QCRI online application (https://rayyan.qcri.org). The references lists of selected studies were critically assessed by both examiners (GNMS and HECS). Any disagreement in first or second phase was solved by discussion until consensus between the two authors was attained. When a consensus was not reached, a third reviewer (FELO) was involved to make a final decision.

*Data collection process and data items*

One author (GNMS) collected the required data from the selected articles and a second author (HECS) cross-checked the collected information. Any disagreement was resolved by consensus or the third reviewer's decision (FELO). For each of the included studies, the following items were recorded: authors, year of publication, country, sample size, objective of the study, imaging modality and protocol, reference exam, image processing techniques, segmentation method, software used, extracted features, machine learning classification method, statistical results and main conclusions. In case of incomplete required data, attempts were made to contact the authors to retrieve any pertinent missing information.

The included studies were also evaluated using the Radiomics Quality Score (RQS), proposed by Lambin et al.[5], which consists of a homogeneous evaluation criteria and reporting guidelines for assessment of radiomic studies. It is composed by five main sections, namely Data selection, Medical imaging, Feature extraction, Exploratory analysis, and Modeling, which include 16 criteria that carry different weights, such that a maximum of 36 points can be achieved. Full and clear reporting of information is required on all these aspects to minimize bias and enhance the usefulness of prediction models.

*Methodological quality assessment of included studies*

The Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Analytical Cross-Sectional Studies[12] was applied as the methodological quality assessment. Regarding JBI quality appraisal, two reviewers (GNMS and HECS) scored the items as Yes, No, Unclear and Not Applicable for the included articles. Any disagreement was resolved by consensus or a third author's decision (PTSF).

*Outcomes of interest*

Included manuscripts were synthesized in a qualitative and quantitative description focused on the following radiomics data: main objective, ROI/ VOI segmentation, software used, feature extraction, machine learning modeling/ classifier and their result rates.

*Synthesis of results*

A meta-analysis was planned since the data from the included studies were considered relatively homogeneous.

## Results

*Study selection*

In Phase I, 508 studies were retrieved from the 8 electronic databases. Afterwards, duplicate articles were removed, resulting in 445 remaining studies. After a comprehensive evaluation of titles and abstracts, 389 articles records were excluded. Therefore, 56 manuscripts were elected to conduct a full-text reading. Later, 35 studies were excluded, resulting in 21 studies at the end of Phase I from databases. From the reference lists of these 21 included studies, seven articles were selected to a full-text reading. Regarding the gray literature, 40 studies were retrieved from Google Scholar and one from ProQuest. No additional study was found across JSTOR.

After cautious evaluation of the 48 abstracts obtained from gray literature and the addition of the reference list results to them, 38 studies were discarded and 10 were assessed for full-text reading. Then, two studies were included. All the excluded articles and reasons for exclusion are described in Supplementary Material 2. Finally, 23 studies were selected for inclusion per study parameters (Figure 1).

*Studies' characteristics*

The 23 included studies were published in seven different types of journals: biomedical computer science[13–16], biomedical physics and engineering[17–19], radiology[20–27], general engineering/ computer science/ physics[28–32], biomedical research[33], oral and maxillofacial surgery[34] and periodontology[35]. Such studies were conducted in 12 different countries: Belgium[33], Brazil[29,35], China[23], Indonesia[15,31], Iran[13,17], Japan[21,22,26,27], Jordan[14,30], Korea[20,24], Poland[32], Turkey[16], the United Kingdom[19], and the United States of America[18,25,28,34]. All studies were published in english. Sample size ranged from 19[29] to 663[19] imaging exams. According to the exam modality, 10 studies used Cone Beam CT (CBCT)[13,16–18,23,28–30,34,35], five adopted Helicoidal CT[21,22,25,26], Dental Panoramic

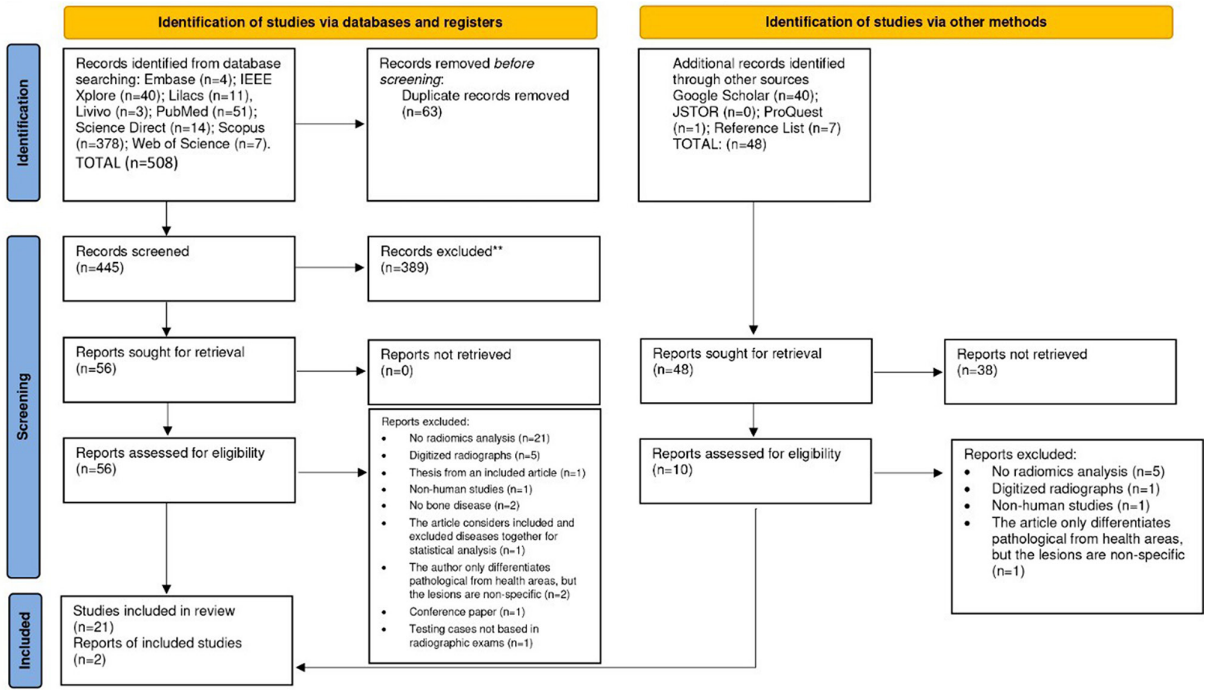


**Figure 1** PRISMA 2020 flow diagram for new systemic reviews which included searches of databases, registers and other sources. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Radiograph was applied in five studies[14,19,20,24,31], Periapical Images in one study[32] and two elected MRI[27,33]. One study used both Dental Panoramic Radiograph and Periapical Images[15]. Seven authors studied osteoporosis[14,15,19,20,24,25,30], four assessed temporomandibular joint (TMJ) pathology[17,28,33,34], one reported diabetes[22], three studies analyzed radicular cysts and granulomas[18,29,32], five authors studied odontogenic and non-odontogenic cysts and tumors[13,16,23,26,31] and three reported osseous inflammation/ infection[21,27,35]. A summary of the studies' descriptive characteristics can be found in Table 1.

*Methodological quality assessment of included studies*

All 23 studies were evaluated by using JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies[12]. The overall methodological assessment ranged from low quality: one to three "Yes" answers; moderate quality: four to six "Yes" answers; or high quality: seven or eight "Yes" answers. Eight[13–15,18,24,30–32] studies presented moderate methodological quality while 13[19–21,23,25–29,33–35] studies showed high methodological quality. Two studies showed low methodological quality[16,17].

All studies measured the outcomes in a valid and reliable way and presented an appropriate statistical analysis. However, many selected studies presented methodological problems related to inclusion and exclusion selection criteria[13–19,24,30–32], subjects description[13,14,16–18,30,31,35] and standardized criteria used for measurement of the condition[13,14,16,17,24,31]. In three studies[16,17,31], the exposure was not measured in a valid and reliable way. More information about methodological assessment may be found in Supplementary Material 3 .

*Radiomics results*

To answer the main question of this systematic review, the following characteristics were considered: main objective, image modality, segmentation, software used, features extracted, machine learning classifier and their result rates. The Radiomics Quality Score (RQS)[5] was assessed and the general result is showed in Figure 2. More detailed radiomics workflow of the included studies is presented in Table 2.

The main objectives of the included works were: (a) recognize a specific disease or a pathologic patter n[15,17,19–25,27,28,30,34,35] or (b) classify two or more similar diseases or pathologic patterns[13,14,16,21,24,26,29,31–33]. Most included studies used CBCT as imaging modality (43.4%), followed by the Dental Panoramic Radiography (26%). Also, the most tested diseases were osteoporosis (30.4%) and odontogenic and non-odontogenic cysts and tumors (21.7%). Manual and semi-automatic segmentations were reported in 86.3% of the studies, and automatic segmentation was also described.

Gray Level Co-occurrence Matrix (GLCM) was the most frequent statistical feature, described in 15 studies, followed by first-order statistics, described in 12 studies (histogram features in seven of them). Gray Level Run Length Matrix (GLRLM) were reported in seven studies,whereas 14 studies mined shape-based features.

**Table 1** Description of the studies' characteristics

| *Author,year, local* | *Reference exam* | *Objective* | *Imaging modality* | *n* | *Main conclusions* |
|---|---|---|---|---|---|
| [13] Abdolali et al., 2016 Iran | Histology | Radicular cyst, dentigerous cyst and keratocyst classification | CBCT | 125 | Using orthogonalized SPHARM (a combination of contourlet and SPHARM features) leads to better results for discriminating between three groups of diseases. SDA performance is superior to SVM in terms of classification accuracy |
| [14]Alzubaidi & Otoom, 2020 Jordan | DXA | Osteoporosis classification | DPR | 575 | The proposed SOM/LVQ method outperformed the SVM method across the 13 feature extractors in terms of the accuracy, sensitivity and specificity performance measures. |
| [28] Bianchi et al., 2020 (1), USA | Clinical parameters following the DC/TMD for TMJ | TMJ Osteoarthritis diagnosis | CBCT: 3D Accuitomo (J. Morita MFG. CORP Tokyo, Japan), FOV 40 × 40 mm; 90kVp, 5mAs, scanning time of 30.8 s and 0.08 $mm^3$ voxel size | 92 | Entropy, Energy, HarCor features were most accurate to differentiate between control and TMJ Osteoarthritis patients. A decreased value of Energy and increased values for HarCor and Entropy were associated with bone sclerosis/ loss. |
| [34] Bianchi et al., 2020 (2), USA | Clinical parameters following the DC/TMD for TMJ | TMJ Osteoarthritis diagnosis | CBCT: 3D Accuitomo scanner (J. Morita Mfg. Corp., Tokyo, Japan). FOV of 40 × 40 mm, 90 kVp, 5 mAs, scanning time of 30.8 s, and 0.08 $mm^3$ voxel size | 84 | 13 bone imaging biomarkers presented acceptable diagnostic performance for the diagnosis of TMJ Osteoarthritis, indicating that the texture and geometry of the subchondral bone microarchitecture may be useful for quantitative grading of the disease. |
| [29] De Rosa et al., 2020 Brazil | Histology | Periapical granuloma and radicular cyst classification | CBCT: Promax 3D scanner (Planmeca Oy, Helsinki, Finland), 96kVp, 12mA, FOV of 6 cm, voxel size of 200 µm and acquisition time of 12 s | 19 | Texture analysis showed potential for differentiating between radicular cysts and periapical granulomas as there was an association between the five texture parameters in the characterization of the lesions. |
| [35] Gonçalves et al., 2020 Brazil | Oral Radiologist diagnosis | Grade C periodontitis diagnosis | CBCT: i-CAT 3D Imaging System (Imaging Sciences International, Hatfield, PA), 37.07 mA, 120 kVp, voxel size of 0.20 mm, exposure time of 26.9 s and FOV of 8.0 × 8.0 mm | 34 | There were significant statistical differences regarding the textural parameters between groups. This may bring valuable meaning to the identification of regions already affected by inflammatory processes that could be unnoticed. |
| [17] Haghnegahdar et al., 2016 Iran | - | TMJ disorders diagnosis | CBCT | 264 | k-NN classifier achieved a very good accuracy and showed desirable sensitivity and specificity results. The proposed method could help automatically diagnose TMJ disorders at their initial stages. |

(Continued)

**Table 1** (Continued)

| *Author,year, local* | *Reference exam* | *Objective* | *Imaging modality* | *n* | *Main conclusions* |
|---|---|---|---|---|---|
| [20]Hwang et al., 2017 Korea | DXA | Osteoporosis diagnosis | DPR: Cranex 3 + Ceph panoramic apparatus (Soredex Co, Helsinki, Finland), 67–71 kV at 10 mA (exposure time, 19.5 s). Images were read using a FCR XG5000 cassette reader (Fuji film Co, Tokyo, Japan) at 170 dpi | 454 | Endosteal margin area was an effective ROI that showed statistically significant differences in FD, GLCM and strut variables between osteoporotic and non-osteoporotic patients, whereas the medullary portions of DPR showed few distinguishing features. It was also found that the strut variables showed the highest sensitivity, specificity and accuracy using the decision tree and SVM. |
| [21] Ito et al., 2021 (1), Japan | MRI alterations + no CT and no clinical alterations | MRONJ diagnosis | CT: 64-multidetector row CT scanner (Aquilion 64; Toshiba Medical Systems, Tokyo, Japan) using the craniomaxillofacial protocol; tube voltage, 120 kV; tube current, 100 mA; FOV 240 × 240 mm; and helical pitch, 41. The protocol consisted of axial (0.50 mm) and multiplanar (3.00 mm) images | 25 | CT was able to quantitatively assess texture features of normal mandibular bone marrow and that with MRONJ. Texture analysis may be useful as a new method for detecting stage 0 MRONJ using CT. |
| [22] Ito et al., 2021 (2), Japan | Blood examination | Diabetes mellitus Type two diagnosis | CT: 64-multidetector row CT system (Aquilion 64; Toshiba Medical Systems, Tokyo, Japan). Protocol for craniomaxillofacial: tube voltage, 120 kV; tube current, 100 mA; FOV 240 × 240 mm; and helical pitch, 41. The imaging included axial (0.50 mm) and multiplanar (3.00 mm) images. | 32 | The value of entropy and difference entropy were significantly higher in diabetes mellitus patients than in non- diabetes mellitus patients. From the histogram, GLCM, and GLRLM results, mandibular condylar bone marrow of diabetes mellitus patients was overall non-uniform, but uniform locally, and had high brightness values. These results suggested the presence of osteosclerosis in the mandibular condyle of these patients. There was a correlation between all texture parameters and HbA1c. From this result, it is considered that the changes in the mandibular condyle bone marrow become more prominent in severe diabetes mellitus patients. CT texture analysis may have the potential to detect diabetes from the mandibular condyle bone marrow. |

(Continued)

**Table 1** (Continued)

| *Author,year, local* | *Reference exam* | *Objective* | *Imaging modality* | *n* | *Main conclusions* |
|---|---|---|---|---|---|
| [23]Jiang et al., 2021 China | Surgical exploration | Jaw simple bone cyst diagnosis | CBCT: DCTPRO, VATECH (Yongin-Si, Republic of Korea); FOV of 16 × 7 cm and a voxel size of 0.16 mm; 90.0 kV and 9 mA and 24 s scanning time | 38 | GLCM contrast, NGTDM contrast, and GLCM variance may be the characteristic imaging features of simple bone cysts of the jaw. The two non-texture features (volume and size) were significantly different between the groups. |
| [24]Kavitha et al., 2014 Korea | DXA | Osteoporosis diagnostic and classification | DPR: OP-100D, Imaging Instrumentarium, Tusuula, Finland | 141 | The combination of textural features (FD and GLCM) and MCW using three classifiers showed better discriminatory values to identify osteoporosis or low BMD than using only MCW. The overall accuracy of the classification for all feature sets using the SVM classifier was slightly higher than those using the naïve Bayes or k-NN classifiers. |
| [25]Kawashima et al., 2019 USA | DXA | Osteoporosis diagnosis | CT Non-contrast head scans axially acquired on 64-detector row CT scanners (Lightspeed VCT; GE Healthcare, Milwaukee, WI), 120 kV, 225 mA, 1 s/rotation | 58 | Results demonstrated the ability of a texture analysis to distinguish between osteoporosis and normal bone mineral density, despite potential osteoarthritic changes. |
| [30]Marar et al., 2020 Jordan | DXA | Osteoporosis diagnosis | CBCT: CS8100 3D machine, 75kvp and 4mAs (Carestream Dental LLC, Atlanta, GA) | 120 | Applying the ANN to distinguish between healthy and osteoporotic persons resulted to almost 98% successful classification rate of the testing set. With the help of the proposed scheme, dentists could be able to predict osteoporosis accurately and efficiently. |
| [27]Muraoka et al., 2022 Japan | Clinical symptoms and image findings | Acute osteomyelitis diagnosis | MRI: 1.5 T superconducting unit (InteraAchieva® 1.5T Nova; Philips Medical Systems, Best, Netherlands) and a 5-channel phased array coil. The STIR images were obtained using a spin echo sequence with the following parameters: TR/TE/ TI = 2500/60/180 ms; other conditions were: 6.0 mm section thickness, 320 × 256 matrix, 230 × 195.5 mm FOV | 38 | The histogram and texture features of the bone marrow in the mandible were significantly different in those with acute osteomyelitis than in those without it. The 90th percentile was higher in patients with acute osteomyelitis, suggesting that they exhibit regions with higher brightness values in the bone marrow. The sum averages were higher in acute osteomyelitis patients, suggesting that they have a region with a heterogeneity of tissue density in the bone marrow. Heterogeneity is higher in the bone marrow of acute osteomyelitis group than in that of the control group. |

(Continued)

**Table 1** (Continued)

| *Author,year, local* | *Reference exam* | *Objective* | *Imaging modality* | *n* | *Main conclusions* |
|---|---|---|---|---|---|
| [31]Nurtanio et al., 2013 Indonesia | - | Cyst and Tumor Lesions classification | DPR: Cranex 2.5 + Soredex dental panoramic X-ray Machine model PT-12SA, Tusuula, Finland | 133 | The highest performance achieved as the result of the combination of FO and GLRLM. GLCM did not achieve good accuracy. |
| [26]Oda et al., 2019 USA | Histology | Dentigerous cyst, keratocyst and ameloblastoma classification | CT: Non-contrast 64- detector row CT scanner (Lightspeed VCT; GE Healthcare, Milwaukee, WI) with 120 kV, 225 mA and 1 s/rotation, and 1.25 mm thick images were reconstructed using soft tissue and bone algorithms. Axial 1.25 mm images in soft tissue algorithm reconstruction were used | 98 | GLCM, GLRL, GLGM, Laws features and Chi-square features showed significant differences when comparing dentigerous cysts and odontogenic keratocysts. Dentigerous cysts contain fluid with or without inflammatory cells, cholesterol or hyaline bodies, and may not have a great impact on CT texture features. Odontogenic keratocysts are comprised of keratin debris and fluid which may correspond to a highly complicated pattern of design on CT. Differences between histogram features reflected the high density component in odontogenic keratocysts. |
| [18]Okada et al., 2015 USA | Endodontic diagnosis and histology | Periapical cysts and granulomas classification | CBCT: NewTom 3G scanner, QR Srl, Verona, Italy. 360 images at 1◦ intervals in 36 s, with reconstructed image resolution of 512 × 512 pixels and 12 bits per pixel (4096 grayscale). The pixel size was 0.25 × 0.25 mm. The axial slice thickness was 0.2 mm | 45 | Experimental results of the authors showed that CBCT diagnosis can be as accurate as histopathology for differentiating the periapical lesions. |
| [33]Orhan et al., 2021 Belgium | Clinical parameters following the DC/TMD for TMJ | TMJ pathologies classification | MRI: 1.5 T imaging unit (Signa Horizon, GE Electric, Milwaukee; Gyroscan Intera, Philips Medical Systems, Washington; Magnetom SP 4000, Siemens, Erlangen, Germany) with the help of dual-surface coils (3-inch and 6 × 8 cm surface coils). Acquisitions of axial, sagittal, and coronal planes using fast spin-echo sequences. The images were taken in the closed, partially opened, and maximally opened mouth positions to detect disc displacements. | 214 | For disc displacement, the RF classifier was the best method in the validation set on diagnostic performance by four indicators. k-NN and RF were found to be the best methods for identifying the mandibular condyle changes, whereas the RF classifier was the best machine learning approach for quantifying TMJ disc placements on MRI. |

(Continued)

**Table 1** (Continued)

| *Author,year, local* | *Reference exam* | *Objective* | *Imaging modality* | *n* | *Main conclusions* |
|---|---|---|---|---|---|
| [32]Pociask et al., 2021 Poland | Histology | Periapical cysts and granulomas classification | Periapical radiographs: Gendex Kavo 765 DC Intraoral X-ray System, Biberach, Deutschland; 65kV and 7mA; exposure time of 0.1 s and recorded on phosphor plates with a secondary readout of five detectors (CS 7600, Carestream Dentak LLC, Atlanta, GA) connected to a Kamsoft computer system. Resolution from 490 × 649 to 1528 × 2024 pixels | 62 | The differentiation between periapical cysts and granulomas was possible, but a definite distinction was not feasible. The most important information about the differentiation of lesions was found in the border of the lesion. Granulomas create a fibrous capsule, while radicular cysts are lined with epithelium. This feature influences textural analysis of the given lesions. High cross-correlation within the group of texture features obtained from the run length matrix, indicating that the structure was isotropic. |
| [19]Roberts et al., 2013 United Kingdom | DXA | Osteoporosis diagnosis | DPR: Planmeca PM2002CC (Planmeca Oy, Helsinki, Finland), Cranex 3DC (Soredex, Tuusula, Finland) 70 kV (constant potential) at 8 mA for 15 s. In Leuven, ADC Solo (Afga, Mortsel, Belgium) was used as the photostimulable phosphor plate system, but other centers used analog images. | 663 | Texture classifiers based on co-occurrence statistics perform much better than those based on FD that have been investigated previously. The combined classifier using cortical texture and width results showed a significantly stronger association with osteoporosis at the femoral neck than width-only methods, but at other skeletal sites there is little if any improvement. |
| [15]Sela & Widyaningrum, 2015 Indonesia | DXA | Osteoporosis diagnosis | DPR: Panoura deluxe dental X-ray unit with 70–80 kVp, 12mA, and 12 s exposure (Yoshida Dental Mfg. Co., Ltd., Japan). Periapical radiograph: DBSWin 4.5, Durr Dental (Bietigheim-Bissingen, Deutschland) | 69 | The proposed model could perform for osteoporosis detection using the selected porous trabecular bone features. |
| [16]Yilmaz et al. 2017 Turkey | Clinical, radiographic, and histopathological features | Periapical cyst and keratocyst classification | CBCT: KODAK K9500 Trophy device (Carestream Health, Rochester, NY) | 50 | The use of features extracted from datasets for classification of dental lesions as periapical cysts or keratocystic odontogenic tumor was determined to be appropriate. The best performance results were achieved using the SVM classifier. |

ANN, Artificial Neural Networks; BMD, Bone mineral density; CBCT, Cone Beam Computed Tomography; DC/TMD, Diagnostic Criteria for Temporomandibular Disorders; DPR, Dental Panoramic Radiograph; DXA, Dual-energy X-ray Absorptiometry; FD, Fractal Dimension; FO, First-order statistics texture; FOV, Field of View; GLCM, Gray-Level Co-occurrence Matrix; GLRL, Gray Level Run-Length; GLRLM, Gray Level Run Length Matrix; HarCor, Haralick Correlation; MCD, Mandibular Cortical Degree; MCW, Mandibular Cortical Width; MRI, Magnetic Resonance Imaging; MRONJ, Medication-related Osteonecrosis of the Jaw; NGTDM, Neighboring Gray Tone Difference Matrix; RF, Random Forest; ROI, Region of Interest; SDA, Sparse Discriminant Analysis; SOM-LVQ, Self-Organizing Map - Learning vector quantization; SPHARM, Spherical Harmonics; SVM, Support Vector Machine; TMJ, Temporomandibular Joints; k-NN, k-Nearest Neighbors.

Only 13 (56.5%) authors described the machine learning classifier step. Most studies developed the machine learning classifier based on supervised learning Support Vector Machine (SVM) (61.5%) and random forest classifier (38.5%). For osteoporosis diagnosis and classification, filtering, shape-based and Tamura

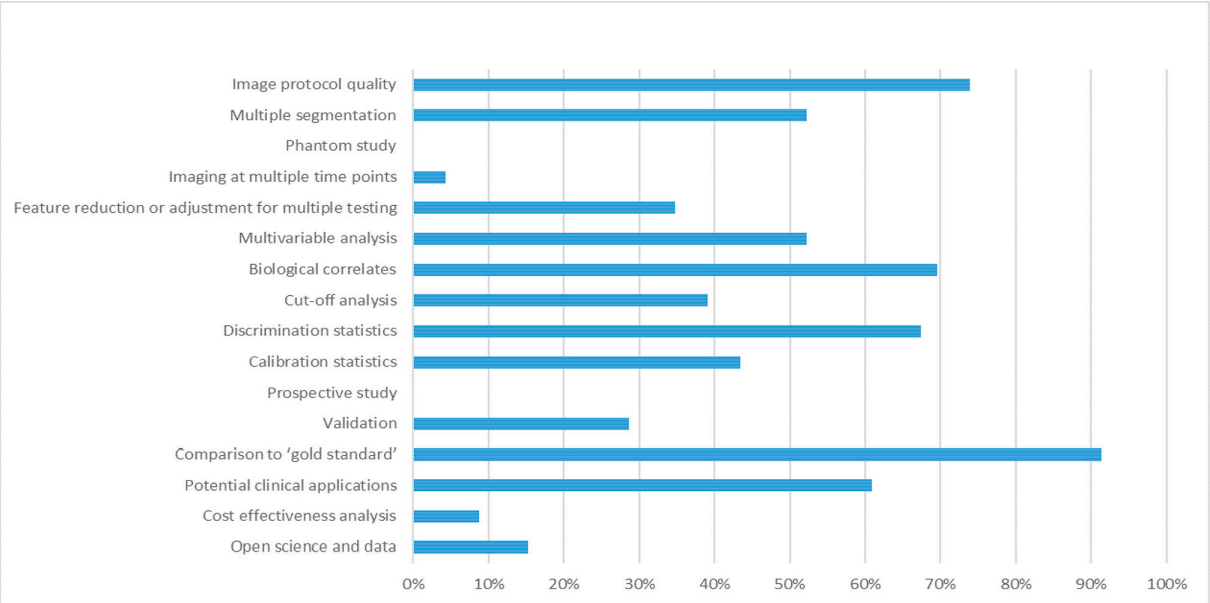


**Figure 2** Radiomics Quality Score (RQS).

texture features showed the best performance. For TMJ pathology, GLCM, GLRLM, Gray Level Size Zone Matrix (GLSZM) first-order statistics and shape-based analyses showed the best results. Considering odontogenic and non-odontogenic cysts and tumors, contourlet and Spherical Harmonics (SPHARM), first-order statistical, GLRLM, GLCM features had better indexes. For odontogenic cysts and granulomas, first-order statistical analysis showed better classification results.

*Osteoporosis:* All the seven[14,15,19,20,24,25,30] included studies that explored osteoporosis used the Dual energy X-ray Absorptiometry (DXA) as the reference bone density exam. Four authors[14,19,20,24] chose Dental Panoramic Radiographs, one Dental Panoramic Radiographs with Periapical Radiographs[15] , one non-contrast head CT[25] and one mandibular CBCT[30] to extract features. Manual or semi-automatic segmentation was applied in five studies, while automatic segmentation was performed in two studies[15,19]. The extracted features included histograms[14,25], filtering[24], GLCM[19,20,24,25], Gray-Level RunLength (GLRL)[25], Gray-Level Gradient Matrix (GLGM) [25], Law features[25] and Tamura features for 3D images[30]. Other morphological/ structure or shape features mined were the Mandibular Cortical Width (MCW)[19,24], Strut analysis[20], Fractal Dimension (FD)[19,20,24], thickness and roughness[24], shape-based porous trabecular [14,15] and trabecular features[30].
The machine learning algorithms used were Learning Vector Quantization (LQV)[14], Decision Tree[15,20], SVM[20,24], Naïve Bayes[24],k-Nearest Neighbors (k-NN)[24], Back Propagation Artificial Neural Networks (BPANN)[30], and Random Forest Classifier[19]. One author[25] did not developed a model classifier.

The most expressive results were observed in osteoporosis level and osteoporosis diagnosis. Regarding osteoporosis level, accuracy 92.6% was obtained using Gabor-based filter algorithm in SOM/ LVQ (Self-Organizing Map/ Learning Vector Quantization) method[14].For osteoporosis diagnosis, it was observed: accuracy 96.25% using Strut variables in the endosteal margin area in Decision Tree and 96.9% in SVM methods[20]; accuracy of 96.8% using FD plus MCW in SVM classifier method[24]; accuracy of 97.917% using Tamura texture features in BPANN method[30] and accuracy of 73.33% using porosity feature in Decision Tree method[15]. Area Under the Curve (AUC) value for identifying osteoporosis at femoral neck was 0.872 using combined MCW and GLCM in random forest classifier method[19]. Moreover, an author[25] evaluated the texture features without machine learning classifier method, which demonstrated that 31 features had significant differences between normal Bone Mineral Density (BMD) and osteoporosis in the left condyle and that 22 features had a statistically significant difference in both sides of the condyle, being six histogram, three GLCM, nine GLRL, two GLGM and two Law features

*TMJ pathology:* Four studies intended to evaluate the diagnostic performance of radiomic features for TMJ, being two for osteoarthritis[28,34], and two for TMJ disorders[17,33].Three[28,33,34] of them used the clinical parameters following the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) as the reference exam. CBCT was the imaging exam used in three studies[17,28,34] and MRI in one[33]. Manual segmentation was done in three studies[28,33,34], while one study did not mention how the segmentation was applied[17]. The radiomic features include first-order/histogram of oriented gradients[17,33], shape[33], GLCM[28,33,34], GLRLM[28,33,34] and GLSZM[33]. Also, semantic features including biomolecular (proteins

**Table 2** Radiomics worflow

| Author, year | Software | Imaging processing | Segmentation | Features extracted | Classifier | Validation | Results | RQS |
|---|---|---|---|---|---|---|---|---|
| *Osteoporosis* | | | | | | | | |
| [14]Alzubaidi et al., 2020 | - | Histogram Equalization | Manual | Agnostic features: shaped-based, first-order and filtering | Supervised learning: SOM/LQV | Internal | Gabor-based algorithm achieved an accuracy of 92.6%, a sensitivity of 97.1%, and a specificity of 86.4% in SOM/LVQ classification method | 13 |
| [20]Hwang et al., 2017 | In house developed software (MATLAB and ImageJ) | Black and white normalization; Upsampling; gaussian blur; Density correction; Binarization; Skeletonization | Manual | Agnostic features: shaped-based and GLCM | Supervised learning: Decision tree and SVM | Internal | Strut variables in the endosteal margin area were sensitivity, specificity, and accuracy of the 97.1%, 95.7 and 96.25 using the decision tree and 97.2%, 97.1 and 96.9% using SVM | 12 |
| [24]Kavitha et al., 2014 | In house developed software | Histogram equalization; Clustering thresholding; High-pass filtering | Manual | Agnostic features: shaped-based and GLCM | Supervised learning: Naïve Bayes, k-NN and SVM | Internal | The combinations of FD plus MCW (95.3%, 92.1%, 96.8%) and GLCM plus MCW (93.7%, 89.5%, 94.2%) for femoral neck BMD showed the highest diagnostic accuracy using the naïve Bayes, k-NN and SVM classifiers, respectively. | 9 |
| [25]Kawashima et al., 2019 | In house developed software | Kernels reconstructions | Manual | Agnostic features: first- order, GLCM, GLRLM, GLGM, Law features | - | - | 22 features including six histogram, 3 GLCM, 9 GLRLM, 2 GLGM and 2 Law's features demonstrated a statistically significant difference in both sides of the condyle. | 8 |
| [30]Marar et al., 2020 | MATLAB | Gray Scale Image Conversion; Illumination filtering based on Retinex method; Canny edge detection; Otsu segmentation | Semi-automatic | Agnostic features: shaped-based and Tamura texture features for 3D images | Supervised learning: BPANN | Internal | The performance of the suggested feed forward BPANN classifier was measured by precision, recall, and accuracy which were 0.96, 1, and 97.917%, respectively. | 7 |
| [19]Roberts et al., 2013 | - | Square-rooted German-McClure kernel function normalization | Automatic | Agnostic features: Shaped-based and GLCM | Supervised learning: Random Forest Classifier | External | AUC values for identifying osteoporosis at femoral neck were 0.830, 0.824, and 0.872 using, respectively, cortical width alone, cortical texture (GLCM) alone and combined width with texture. At 80% sensitivity, these classifiers produced specificity values of 74.4%, 73.6%, and 80.0%, respectively. | 22 |
| [15]Sela et al., 2015 | Custom computer program | Tophat and bothat filtering and histogram equalization | Automatic | Agnostic features: Shaped-based analysis (porosity) | Supervised learning: Decision tree | Internal | From 54 training data, 49 data can be detected correctly. The results of performance detection obtained accuracy, sensitivity, specificity were 73.33%, 72.23%, and 72.23%, respectively. | 14 |
| TMJ Pathology | | | | | | | | |

(Continued)

**Table 2** (Continued)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Osteoporosis* | | | | | | | | |
| [28]Bianchi et al., 2020 (1) | 'BoneTexture' module in 3D Slicer; Ibex; ITK-Snap | Optimization using i-Dixel (J. Morita MFG. CORP Tokyo, Japan) filter: G_103 + H_009 | Manual | Semantic features: biomolecular (saliva and blood), and clinical variables Agnostic features: GLCM, GLRLM and shaped-based (bone morphometry) | Supervised learning: Logistic Regression, Random Forest, LightGBM and XGBoost | Internal | XG-Boost + LightGBM with the interaction features (radiomics, clinical and protein information) achieved the best average accuracy of 0.823, AUC 0,870, and F1-score 0,823 to determine disease status. | 22 |
| [34]Bianchi et al., 2020 (2) | 'BoneTexture' module in 3D Slicer; ITK-Snap | Mirroring to the right side-and creation of a spatial orientation matrix | Manual | Agnostic features: Shaped-based (bone morphometry); GLCM, GLRLM | - | - | Osteoarthritis with AUC values were between 0.620 and 0.710. Good diagnostic performance, especially for energy and entropy, with AUC of ≥ 0.70 | 12 |
| [17]Haghnegahdar et al., 2016 | - | Gray level and LBP | not mentioned | Agnostic feature: first-order | Supervised learning: k-NN, SVM, Naïve Bayes and Random Forest Classifier | Internal | k-NN was the best classifier, by 92.42% accuracy, 94.70% sensitivity and 90.15% specificity. | 12 |
| [33]Orhan et al., 2021 | Huiying Radiomic platform and Radcloud platform | - | Manual | Agnostic features: shaped-based, first-order, GLCM, GLRLM and GLSZM analysis | Unsupervised and supervised learning: k-means clustering learning; Logistic regression, random forest, decision tree, k-NN, XGBoost and SVM | Internal | AUC of k-NN and random forest were high, with a range of 0.89 and 0.77 for the training set and validation set, respectively. | 17 |
| Odontogenic and non- odontogenic cysts and tumors | | | | | | | | |
| [13]Abdolali et al., 2016 | MATLAB | Active contour method based on symmetry analysis | Automatic | Agnostic features: shaped-based and contourlet alone, contourlet + conventional SPHARM, and contourlet + orthogonalized SPHARM analysis | Supervised learning: SVM, SDA | Internal | A combination of contourlet and SPHARM features leads to better results. The classification accuracy of 94.29 and 96.48% is achieved using SVM and SDA classifier | 16 |
| [23]Jiang et al., 2021 | ITK-Snap | Active contour methods | Manual | Semantic features: volume, size, solidity, eccentricity Agnostic features: first-order, GLCM, GLRLM, GLSZM, and NGTDM | - | - | The absolute value of correlation coefficient was 0.487–0.775. GLCM contrast, NGTDM contrast, and GLCM variance were the features with the highest correlation coefficients. | 10 |
| [31]Nurtanio et al. 2013 | MATLAB | Black and white normalization; gaussian filter | Manual | Agnostic features: first-order, GLCM, and GLRLM | Supervised learning: SVM | Internal | It was obtained up to 84.62% accuracy using first-order statistical features, 87.18% using the combination of first-order statistical features and GLRLM, and 87.18% using the combination of first-order statistical features, GLCM and GLRLM. | 12 |

(Continued)

**Table 2** (Continued)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Osteoporosis* | | | | | | | | |
| [26]Oda et al., 2019 | OsiriX (Pixmeo SARL, Bernex, Switzerland) and in-house-developed MATLAB | - | Manual | Semantic features: region, tooth within the lesion, volume, septum Agnostic features: first-order, GLCM, GLRLM, GLGM, Laws features and Chi-square features | - | - | Kruskal-Wallis analysis: nine histogram features, 1 GLCM feature, 3 GLRLM features, 4 GLGM features, 2 Laws features, and 2 Chi-square features showed significant differences among the different types of lesions. Mann-Whitney's U test analysis: nine histogram features, 1 GLCM feature, 3 GLRLM features, 4 GLGM features, 2 Laws features, and 2 Chi-square features showed significant differences between DC and OKC; four histogram features and 1 Chi-square feature showed significant differences between OKC and AM. two histogram features showed significant differences between DC and AM. | 10 |
| [16]Yilmaz et al., 2017 | In-house-developed MATLAB and Rapid Miner data mining software | - | Manual | Agnostic features: first-order and GLCM | Supervised learning: k-NN, Naïve Bayes, decision trees, random forest, neural network, and SVM | Internal | SVM achieved the best performance with the vector consisting of the single selected feature obtained from 3D GLCM, with a sub feature vector consisting of two values obtained from 3D GLCM, and with a sub feature vector consisting of single value obtained from 3D GLCM, respectively. The SVM classifier achieved the best performance of 100% accuracy after feature reduction. | 15 |
| Odontogenic Cysts and Granulomas | | | | | | | | |
| [29]De Rosa et al. (2020) | Mazda | Images were converted into bitmap format with loss resolution. | Manual | Semantic features: age, gender, lesion diameter and volume, tooth Agnostic feature:GLCM | - | - | The five parameters (angular second moment, sum of squares, sum of average and contrast, and correlations in all directions) were selected for ROC curve as they achieved the AUC of 81.2% in the other half validation set. | 7 |
| [18]Okada et al. 2015 | MATLAB | Binarization | Semi-automatic | Agnostic feature: first-order | Supervised learning: LDA-AdaBoost method | External | Accuracy of LDA-AdaBoost method using median and minimum intensity was 94.1% for endodontic diagnosis and 78.9% for histopathology as gold standard. | 11 |

(Continued)

**Table 2** (Continued)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Osteoporosis* | | | | | | | | |
| [32]Pociask et al. 2021 | MaZda | Images were converted into JPG format, with loss resolution and without scale preservation. | Manual | Agnostic features: First-order (histogram-based method), second -order analysis (14 Haralick features), GRLM, GTDM and LBP | Unsupervised and supervised learning: Logistic regression; 2D and 3D t-SNE models | Internal | five parameters (angular second moment, sum of squares, sum of average and contrast) were selected given their potential to differentiate periapical cysts from granulomas as they achieved the AUC of 81.2%. The best predictor variable was YS6GlcmZ4Entropy texture in logistic regression method. | 15 |
| Bone inflammation/ infection | | | | | | | | |
| [35]Gonçalves et al., 2020 | MaZda | Images were converted into bitmap format with loss resolution | Manual | Agnostic feature: GLCM | - | - | Comparisons between groups showed statistically significant differences for all parameters. In the inter group analysis, there were statistically significant differences in inverse difference moment, angular second moment, entropy, sum of average, sum of variance, sum of entropy, difference of variance and difference of entropy parameters. | 8 |
| [21]Ito et al., 2021 (1) | Open-access LIFEx software | - | Manual | Semantic features: age, sex, MRONJ stage, bisphosphonates, medical history Agnostic features: GLRLM and GLZLM | - | - | Among 37 texture features, the bone marrow of the mandible with stage 0 MRONJ and the contralateral normal mandibular bone marrow revealed significant differences in 6 GLRLM features and 4 GLZLM features | 7 |
| [27]Muraoka et al., 2022 | MaZda | Intensity normalization using MaZda default | Manual | Agnostic features: first-order, GLCM, GLRLM | - | - | These radiomics features in the acute osteomyelitis group were significantly higher than in the non-osteomyelitis group ($p < 0.001$) | 9 |
| Diabetes Mellitus | | | | | | | | |
| [22]Ito et al., 2021 (2) | MaZda | Intensity normalization using MaZda default | Manual | Agnostic features: first-order, GLRLM, GLCM and wavelets | - | - | One histogram feature, 15 GLCM features, and 4 GLRLM features showed significant differences between the diabetes and control patients cohort | 10 |

AM, Ameloblastoma; AUC, Area Under the Curve; BMD, Bone Mineral Density; BPANN, Back Propagation Artificial Neural Networks; DC, Dentigerous Cyst; FD, Fractal Dimension; GLCM, Gray-Level Co-occurance Matrix; GLGM, Gray-Level Gradient Matrix; GLRLM, Gray-Level Run-Length Matrix; GLSZM, Gray-Level Size Zone; GLZLM, Gray-Level Zone Length Matrix; GTDM, Game Theory Based Decision Making; LBP, Local Binary Patterns; LDA, Linear Discriminant Analysis; LVQ, Learning Vector Quantization; LightGBM, Light Gradient Boosting Machine; MCW, Mandibular Cortical Width; MRONJ, Medication-Related Osteonecrosis of the Jaw; NGTDM, Neighborhood Gray Tone Difference Matrix; OKC, Odontogenic Keratocyst; RQS, Radiomics Quality Score; SDA, Sparse Discriminant Analysis; SOM, Self-Organizing Map; SPHARM, Spherical Harmonics; SVM, Support Vector Machine; XGBoost, Extreme Gradient Boosting; k-NN, k-Nearest Neighbors; t-SNE, t-Distributed Stochastic Neighbor Embedding.

of serum and saliva)[28], clinical parameters (age, gender, signs and symptoms)[28,34], and imaging characteristics (bone morphometry)[28,34] were correlated to radiomic analysis in two studies. The machine learning algorithms used were Logistic Regression[28,33], random forest classifier[17,28,33], Light Gradient Boosting Machine (LightGBM)[28], Extreme Gradient Boosting (XGBoost)[28,33], k-NN[17,33], SVM[17,33], Naïve Bayes[17],

Decision Tree[33], k-means clustering[33]. One author[34] did not developed any model classifier.
The best classifiers' results were: accuracy of 82.3% using clinical, protein information and radiomics features (GLCM, GLRLM and bone morphometry) in XG-Boost+LightGBM method,[28] accuracy of 92.42% using first-order statistics analysis in k-NN classifier method[17] and AUC of 89% using shaped-based, first-order, GLCM, GLRLM and GLSZM analysis in k-NN method[33]. Without the classifier method, an author described a AUC ≥ 0.70 for energy and entropy features[34].

*Odontogenic and non-odontogenic cysts and tumors:* Five studies aimed to diagnose cysts[23],classify cysts[13,16] or differentiate cysts from tumor lesions[26,31]. There were two standard comparators, histological exams[13,16,26] and surgical exploration[23]. One author[31] did not mention a reference exam. CBCT[13,16,23], Dental Panoramic Radiography[31], and non-contrast CT[26] were the imaging exams explored. Manual segmentation was done in four studies[16,23,26,31], while automatic segmentation was performed in one study[13]. The mined radiomics features were edge- and region-based analysis (shape)[13], first order[16,23,26,31], GLCM[16,23,26,31], GLRLM[23,31], GLSZM[23], Neighboring Gray Tone Difference Matrix (NGTDM)[23], GLRL[26], GLGM[26], Law features[26], $\chi^2$ features[26] and coutourlet[13]. Semantic features, such as volume, size, solidity and eccentricity[23] and region, tooth within the lesion, volume, septum[26], were also described. Only three authors[13,16,31] developed the machine learning model, with the algorithms SVM[13,16,31], k-NN[16], Naïve Bayes[16], decision trees[16], random forest classifier[16] and neural network[16].
The most expressive results were: accuracy of 96.48% in classifying radicular cyst, dentigerous cyst and keratocysts, using a combination of contourlet and SPHARM features in Sparse Discriminant Analysis (SDA) classifier[13]; accuracy of 87.18% in differentiating cysts from tumors, using the combination of first-order statistical features and GLRLM; and accuracy of 87.18%, using the combination of first-order statistical features, GLCM and GLRLM, with SVM classifier[31]. The SVM classifier achieved the best classification performance in differentiating periapical cysts from keratocysts, with 100% accuracy after feature reduction based on first-order and GLCM features[16]. Analyzing features without the classifier model, one author[23] found that two semantic features (volume and size) were significantly different between the keratocyst and simple bone cyst groups. Among the texture features, GLCM contrast, NGTDM contrast, and GLCM variance had the highest correlation coefficients[23]. Another author[26] reported that significant differences were obtained between dentigerous cyst and keratocyst by using Mann–Whitney's *U* test analysis, nine histogram features, one GLCM, three GLRL, four GLGM, two Law features, and two $\chi^2$ features. ; Furthermore, four histogram and one $\chi^2$ features showed significant differences between keratocyst and ameloblastoma, and two histogram features showed significant differences between dentigerous cyst and ameloblastoma[26].

*Odontogenic cysts and granulomas:* Three authors[18,29,32] proposed the differentiation between radicular cysts and granulomas. Two of them used CBCT images[18,29] and one Periapical Radiographs[32], which were compared to histology and clinical endodontics findings. Manual segmentation was applied in two studies[29,32], while semi-automatic segmentation was performed in one study[18]. The radiomics features were GLCM[29], histogram-based method[18,32], second-order features (14 Haralick features)[32], Run-length Matrices (RLM)[32], Gray-tone Difference Matrices (GTDM)[32], Local Binary Patterns (LBP)[32]. One author used as semantic features age, gender, diameter, volume and tooth involved[29].
Two authors developed the machine learning models, in which the learning methods were Linear Discriminant Analysis (LDA)-AdaBoost method[18], logistic regression[32], and 2D and 3D t-Distributed Stochastic Neighbor Embedding (t-SNE) models[32].

The best results were: accuracy of 94.1% in classifying radicular periapical cysts and granulomas, using first-order statistical analysis in LDA-AdaBoost classifier[18]; and the best predictor variable was YS6GlcmZ4Entropy texture in logistic regression method[32]. Without machine learning classifier model, one author described five parameters (angular second moment, sum of squares, sum of average, correlation and contrast) which were selected due to their potential to differentiate periapical cysts from granulomas based on their sensitivity and specificity, as they achieved the AUC of 81.2% in the validation set[29].

*Bone inflammation/ infection:* One author proposed to detect patients with grade C periodontitis using CBCT[35], one to detect stage 0 Medication-related Osteonecrosis of the Jaw (MRONJ) using CT[21] and another to early diagnose suppurative osteomyelitis using MRI[27]. All of them applied manual segmentation and no machine learning model was developed. The radiomics features were GLCM[27,35], GLRLM[21], GLZLM[21], and histogram feature[27].
Comparisons between periodontal disease groups showed statistically significant differences for all parameters. In the intergroup of periodontitis level analysis, there were statistically significant differences in the following parameters: inverse difference moment, angular second moment, entropy, sum of average, sum of variance, sum of entropy, difference of variance and difference of entropy[35].

Among 37 texture features, the bone marrow of the mandible with stage 0 MRONJ and the contralateral

normal mandibular bone marrow revealed significant differences in six GLRLM and four GLZLM features. Moreover, these texture features exhibited a moderate diagnostic performance[21].

Finally, the histogram and the eight GLCM features in the acute osteomyelitis group were significantly higher than in the non-osteomyelitis group[27].

*Diabetes mellitus:* One author[22] proposed the assessment of the mandibular condylar bone marrow in diabetes mellitus using CT. The reference exam was defined as HbA1*c* ≥ 6.5% on blood test. Manual segmentation was performed and the radiomics features selected were histogram feature, GLRLM, GLCM, and wavelets. No machine learning model was developed. One histogram, 15 GLCM, and 4 GLRLM features showed significant differences between the diabetes and control patients.

*Quantitative synthesis and evidence certainty*

The selected studies were all descriptive and used similar methods, which reduced the possibility of misinterpretation. Results were considered homogeneous enough, but did not have compatible quantitative data to run a meta-analysis.

Due to the qualitative nature of included studies, an analysis of the certainty of evidence was not performed.

## Discussion

This is the first systematic review that discussed which are the texture features of bone pathologic changes of the jaws and how they are extracted for diagnosis, predicting prognosis and therapeutic response in dentomaxillofacial imaging. According to the obtained results, agnostic features based on shape, intensity, texture and filtering information were mined from Periapical, Dental Panoramic Radiographs, CBCT, Helicoidal CT and MRI images. Six different jawbone alterations were studied: osteoporosis, TMJ pathology, diabetes, radicular cysts and granulomas, odontogenic and non-odontogenic cysts and tumors, and osseous inflammation/infection. No study reported predicting response, prognosis or therapeutic response, but instead diagnosis or classification of the aforementioned diseases. The small number of included articles and their main objective are in agreement with a previous meta-analysis[8], which highlighted that 91% of radiomics studies concern oncological applications, and 81% of them are for diagnostic purposes. This systematic review revealed the potentials and also the limitations of the previous published studies, including the variability of imaging techniques and protocols.

The literature recognizes the importance of using standardized imaging protocols to eliminate unnecessary confounding variability when dealing with radiomics, including scanner manufacturer, model and calibrations. Different results may arise when diverse filters and thresholding are used[1,36]. From the different imaging modalities employed in the included studies, a wide range of protocols settings were applied. Some studies have not reported the acquisition protocol. Software variability may also lead to different results even when a feature is measured from the same ROI/VOI obtained from identicalscans[36]. From the included studies, six authors[15,16,20,24–26] used in-house developed software, without providing further details, and three authors[14,17,19] did not even cited which software was used. In addition, JPEG and BMP images formats were used instead of DICOM in two included studies[29,32]. To facilitate interoperability of radiomic features and make standardization available, differences in algorithms and software implementations should be elucidated by using open-source software or releasing source code publicly[5,36].

Another crucial step related to radiomics is the segmentation of the analyzed structures. The majority of the included studies employed manual or semi-automatic segmentation. A previous work reported that manual outlining by experts is considered the ground truth in tumor segmentation, although its result can be unreliable due to inter- and intraobserver segmentation variability[37]. On the other hand, some authors assured that automatic and semi-automatic segmentation methods are preferred due to their robustness and significantly higher levels of reproducibility, with minimum user input[38].

According to this review results, first statistical and the GLCM second statistical-based techniques were the most mined feature algorithms, which corroborates with previous literature[3,39]. In a statistical-based model, first-order statistics evaluate the gray-level frequency distribution from the pixel/ voxel intensity histogram in a given area of interest, while GLCM captures spatial relationships of pairs of pixels/ voxels with pre-defined gray-level intensities, in different directions (horizontal, vertical, or diagonal for a 2D analysis or 13 directions for a 3D analysis), with a predefined distance between them[40]. Regarding the machine learning model development, SVM and random forest were the most applied classifiers in this review. An author highlighted that SVM was one of the first highly successful models, although required a careful feature selection[41]. Also, a previous study compared 12 supervised classifiers in predicting overall survival in lung cancer patients, with CT images and found the random forest classifier as the best performance method[38]. Since radiomics is a new and expanding area of research, better classification algorithms will be developed, so that the optimal method is not yet clearly defined and will depend on the application[3].

As explained, radiomics studies involve multiple complex subprocesses, each one affected by a wide range of decisions, and a sort of software and mathematical approaches to segmentation, texture mining

and statistical analysis[5,36,42]. According to this review results, there is no consensus about the selection of proper texture features that could be relevant in diagnosis of pathological changes of the jaws. In particular, many textural indices show a lack of reproducibility and standardization, an obstacle also reported by other studies[1,5,7,36], including a phantom experiment[42]. These aspects are specially related to the deficiency of imaging protocol standardization in retrospective study designs, poor calibration statistics and restricted source code and data, as observed in the included articles of this review, by means of the Radiomics Quality Score (RQS) proposed by Lambin et al.[5]. The cut-off values were often arbitrarily chosen, a practice previously criticized [7], which may lead to biased validation. This also may increase the risk of false-positive results and consequently delays the translation to clinical practice[43]. Furthermore, radiomics studies based on retrospectively collected data have low level of evidence, because imaging protocols, including acquisition and reconstruction settings, are often not controlled or standardized, and they mainly serve as examples[7]. The solution would include imaging protocols standardization, development of generalizable models, and larger samples. Multicentric efforts are required, since data-sharing enables highly powered prospective studies and accelerates the development and validation of radiomic signatures derived from new and existing data[5].

The conventional visual analysis is based only on the lesion's image behavior, leaving gaps in the diagnosis. Therefore, radiomics could be recommended as an auxiliary tool in the elaboration of the radiologists' report, since the analysis is independent of subjective evaluations[29,35]. Thus, the use of texture analyses in imaging is a striking option, potentially allowing the development of a novel form of disease biomarkers, when integrated to clinical and genomic patient data from multiple sources. , Therefore, the texture analysis has the potential to enhance diagnosis accuracy and delivery personalized healthcare, preventing the need for more invasive steps, such as biopsy or surgery,[1,42,44–46].

To gain the trust of health professionals, recognized institutions, and patients, an imaging diagnostic system must be transparent, interpretable, and explainable[47]. Significant progress has been made, but further improvements are imperative to achieve routine utilization of radiomics due to the insufficient evidence. There are notable differences in terms of sample size, methodology, performance metrics, and clinical utility[7,48]. Important efforts have been made to address these issues, including the cooperation of different institutions, such as the American College of Radiology Imaging Network and the Canadian Institute of Health Research[48]. Working groups that include radiologists, physicists, applied mathematicians, and computer scientists aim to improve the field and educate people on radiomics use as a reliable part of a decision-support system in clinical assistance[1,48]. Another milestone in radiomics is the creation of a predictive tool model. It requires the involvement of different imaging centers from all over the world to provide data and a worldwide standardization of the radiomics process[49]. Although there are many challenges to overcome, the potential benefits to precision health care are enormous and may revolutionize radiology practice in the near future.

*Limitations*

In this review, neither the size of the ROI/ VOI nor the 3D *vs* 2D images influences were evaluated. In addition, the imaging pre- and post-processing techniques were not deeply explored, and the segmentation phase, which is a crucial radiomics step, was not critically assessed. As a great variety of statistical approaches and different overlapping reduction methods were used, a carefully work focused on this subject is strongly suggested. Although the included studies reported jawbone alterations, few of them were published in dentistry journals, which shows that there is still a lack of dentists' engagement in this area of study.

## Conclusion

Feature extraction and analysis are just one part of the growing field of radiomics. GLCM feature was the most frequent statistical feature, followed by first-order statistics, and GLRLM features. Shape-based features were also recurrently mined. No study reported predicting response, prognosis or therapeutic response, but instead disease diagnosis or classification. Due to the lack of standardization showed in this review, there are multiple software and statistical approaches, with the purpose of mining texture parameters, selecting their useful meaning, and building the machine learning models, so that comparison between studies and reproduction of the results is challenging. Aiming to uniform the reports, following the RQS requirements is strongly suggested. Although the limitations of this review, texture analysis showed potential to contribute to radiologists' reports, decreasing the subjectivity when using mathematical approaches, and providing unique, objective and reliable disease information.. Therefore, radiomics analysis is a potential quantitative, non-invasive and accessible tool for personalized healthcare.

*Protocol and registration*

This systematic review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Checklist[50]. The protocol was registered in PROSPERO database (University of York) (http://www.crd.york.ac.uk/PROSPERO)[51] under number CRD42022312507.

## REFERENCES

1. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016; **278**: 563–77. https://doi.org/10.1148/radiol.2015151169
2. Scapicchio C, Gabelloni M, Barucci A, Cioni D, Saba L, Neri E. A deep look into radiomics. *Radiol Med* 2021; **126**: 1296–1311. https://doi.org/10.1007/s11547-021-01389-x
3. Parekh V, Jacobs MA. Radiomics: a new application from established techniques. *Expert Rev Precis Med Drug Dev* 2016; **1**: 207–26. https://doi.org/10.1080/23808993.2016.1164013
4. Leite AF, Vasconcelos K de F, Willems H, Jacobs R. Radiomics and machine learning in oral healthcare. *Proteomics Clin Appl* 2020; **14**(3): e1900040. https://doi.org/10.1002/prca.201900040
5. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017; **14**: 749–62. https://doi.org/10.1038/nrclinonc.2017.141
6. Veena D, Jatti A, Joshi R, D. KS. Characterization of dental pathologies using digital panoramic X-ray images based on texture analysis. In: *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf [Internet]*. ; 2017. Available from: https://pubmed.ncbi.nlm.nih.gov/29059942/
7. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging-"how-to" guide and critical reflection. *Insights Imaging* 2020; **11**(1): 91. https://doi.org/10.1186/s13244-020-00887-2
8. Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, et al. Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. *Eur Radiol* 2020; **30**: 523–36. https://doi.org/10.1007/s00330-019-06360-z
9. Mosier KM. Lesions of the jaw. semin ultrasound, CT MRI. 2015; 444–50.
10. Regezi JA. Odontogenic cysts, odontogenic tumors, fibroosseous, and giant cell lesions of the jaws. *Mod Pathol* 2002; **15**: 331–41. https://doi.org/10.1038/modpathol.3880527
11. Baumhoer D. Bone-Related Lesions of the Jaws. Surg Pathol Clin. Internet]. 2017. Available from: https://pubmed.ncbi.nlm.nih.gov/28797509/
12. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Chapter 7: Systematic reviews of etiology and risk. In: *Joanna Briggs Institute Rewiewer's Manual for Evidence Synthesis*. The Joanna Briggs Institute; 2017, pp. 1–5.
13. Abdolali F, Zoroofi RA, Otake Y, Sato Y. Automated classification of maxillofacial cysts in cone beam CT images using contourlet transformation and spherical harmonics. *Comput Methods Programs Biomed* 2017; **139**: S0169-2607(16)30389-3: 197–207: . https://doi.org/10.1016/j.cmpb.2016.10.024
14. Alzubaidi MA, Otoom M. A comprehensive study on feature types for osteoporosis classification in dental panoramic radiographs. *Comput Methods Programs Biomed* 2020; **188**: S0169-2607(19)31369-0: 105301: . https://doi.org/10.1016/j.cmpb.2019.105301
15. Sela EI, Widyaningrum R. Osteoporosis detection using important shape-based features of the porous trabecular bone on the dental X-ray images. *Ijacsa* 2015; **6**(9). https://doi.org/10.14569/IJACSA.2015.060933
16. Yilmaz E, Kayikcioglu T, Kayipmaz S. Computer-aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography. *Comput Methods Programs Biomed* 2017; **146**: S0169-2607(16)30429-1: 91–100: . https://doi.org/10.1016/j.cmpb.2017.05.012
17. Haghnegahdar AA, Kolahi S, Khojastepour L, Tajeripour F. Diagnosis of tempromandibular disorders using local binary patterns. *J Biomed Phys Eng* 2018; **8**: 87–96. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85042785491&doi=10.22086%2Fjbpe.v0i0.577&partnerID=40&md5=20acd7ffc2268d5985c1f81104195d0c
18. Okada K, Rysavy S, Flores A, Linguraru MG. Noninvasive differential diagnosis of dental periapical lesions in cone-beam CT scans. *Med Phys* 2015; **42**: 1653–65. https://doi.org/10.1118/1.4914418
19. Roberts MG, Graham J, Devlin H. Image texture in dental panoramic radiographs as a potential biomarker of osteoporosis. *IEEE Trans Biomed Eng* 2013; **60**: 2384–92. https://doi.org/10.1109/TBME.2013.2256908
20. Hwang JJ, Lee J-H, Han S-S, Kim YH, Jeong H-G, Choi YJ, et al. Strut analysis for osteoporosis detection model using dental panoramic radiography. *Dentomaxillofac Radiol* 2017; **46**(7): 20170006. https://doi.org/10.1259/dmfr.20170006
21. Ito K, Muraoka H, Hirahara N, Sawada E, Hirohata S, Otsuka K, et al. Quantitative assessment of mandibular bone marrow using computed tomography texture analysis for detect stage 0 medication-related osteonecrosis of the jaw. *Eur J Radiol* 2021; **145**: S0720-048X(21)00511-8: 110030: . https://doi.org/10.1016/j.ejrad.2021.110030
22. Ito K, Muraoka H, Hirahara N, Sawada E, Okada S, Kaneda T. Computed tomography texture analysis of mandibular condylar bone marrow in diabetes mellitus patients. *Oral Radiol* 2021; **37**: 693–99. https://doi.org/10.1007/s11282-021-00517-7
23. Jiang ZY, Lan TJ, Cai WX, Tao Q. Primary clinical study of radiomics for diagnosing simple bone cyst of the jaw. *Dentomaxillofac Radiol* 2021; **50**(7): 20200384. https://doi.org/10.1259/dmfr.20200384
24. Kavitha MS, An S-Y, An C-H, Huh K-H, Yi W-J, Heo M-S, 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**: S2212-4403(14)01393-5: 346–56: . https://doi.org/10.1016/j.oooo.2014.11.009
25. 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. [Internet]. *Eur J Radiol* 2019; **116**: S0720-048X(19)30176-7: 212–18: . https://doi.org/10.1016/j.ejrad.2019.05.009
26. Oda M, Staziaki PV, Qureshi MM, Andreu-Arasa VC, Li B, Takumi K, et al. Using CT texture analysis to differentiate cystic and cystic-appearing odontogenic lesions. *Eur J Radiol* 2019; **120**: S0720-048X(19)30304-3: 108654: . https://doi.org/10.1016/j.ejrad.2019.108654
27. Muraoka H, Ito K, Hirahara N, Ichiki S, Kondo T, Kaneda T. Magnetic resonance imaging texture analysis in the quantitative evaluation of acute osteomyelitis of the mandibular bone. *Dentomaxillofac Radiol* 2022; **51**(1): 20210321. https://doi.org/10.1259/dmfr.20210321
28. Bianchi J, de Oliveira Ruellas AC, Gonçalves JR, Paniagua B, Prieto JC, Styner M, et al. Osteoarthritis of the temporomandibular joint can be diagnosed earlier using biomarkers and machine learning. *Sci Rep* 2020; **10**(1): 8012. https://doi.org/10.1038/s41598-020-64942-0
29. De Rosa CS, Bergamini ML, Palmieri M, Sarmento DJ de S, de Carvalho MO, Ricardo ALF, et al. Differentiation of periapical granuloma from radicular cyst using cone beam computed tomography images texture analysis. *Heliyon* 2020; **6**: e05194. https://doi.org/10.1016/j.heliyon.2020.e05194
30. Marar RFA, Uliyan DM, Al-Sewadi HA. Mandible bone osteoporosis detection using cone-beam computed tomography. *Eng Technol Appl Sci Res* 2020; **10**: 6027–33. https://doi.org/10.48084/etasr.3637
31. Nurtanio I, Purnama I, Hariadi M, Purnomo M. Classifying Cyst and Tumor Lesion Using Support Vector Machine Based on Dental Panoramic Images Texture Features. IAENG Int J Comput Sci. 2013;40(1):29–37.
32. Pociask E, Nurzynska K, Obuchowicz R, Bałon P, Uryga D, Strzelecki M, et al. Differential diagnosis of cysts and granulomas supported by texture analysis of intraoral radiographs. *Sensors (Basel)* 2021; **21**(22): 7481. https://doi.org/10.3390/s21227481
33. Orhan K, Driesen L, Shujaat S, Jacobs R, Chai X. Development and validation of a magnetic resonance imaging-based machine

learning model for TMJ pathologies. *Biomed Res Int* 2021; **2021**: 6656773. https://doi.org/10.1155/2021/6656773
34. Bianchi J, Gonçalves JR, de Oliveira Ruellas AC, Ashman LM, Vimort J-B, Yatabe M, et al. Quantitative bone imaging biomarkers to diagnose temporomandibular joint osteoarthritis. *Int J Oral Maxillofac Surg* 2021; **50**: S0901-5027(20)30163-6: 227–35: . https://doi.org/10.1016/j.ijom.2020.04.018
35. Gonçalves BC, de Araújo EC, Nussi AD, Bechara N, Sarmento D, Oliveira MS, et al. Texture analysis of cone-beam computed tomography images assists the detection of furcal lesion. *J Periodontol* 2020; **91**: 1159–66. https://doi.org/10.1002/JPER.19-0477
36. Moskowitz CS, Welch ML, Jacobs MA, Kurland BF, Simpson AL. Radiomic analysis: study design, statistical analysis, and other bias mitigation strategies. *Radiology* 2022; **304**: 265–73. https://doi.org/10.1148/radiol.211597
37. Tanner C, Khazen M, Kessar P, Leach MO, Hawkes DJ. n.d.).( Classification improvement by segmentation refinement: application to contrast-enhanced MR-mammography. *Lect Notes Comput Sci [Internet] 2004 [Cited 2022 Aug 12];3216(PART 1):184–91 Available From:*. Available from: https://link.springer.com/chapter/10.1007/978-3-540-30135-6_23
38. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 2015; **5**: 13087. https://doi.org/10.1038/srep13087
39. Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT texture analysis: definitions, applications, biologic correlates, and challenges. *Radiographics* 2017; **37**: 1483–1503. https://doi.org/10.1148/rg.2017170056
40. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. *J Nucl Med* 2020; **61**: 488–95. https://doi.org/10.2967/jnumed.118.222893
41. Burges CJC. n.d.).( A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov 1998 22 [Internet] 1998 [Cited 2022 Aug 12];2(2):121–67 Available From:*. Available from: https://link.springer.com/article/10.1023/A:1009715923555
42. Caramella C, Allorant A, Orlhac F, Bidault F, Asselain B, Ammari S, et al. Can we trust the calculation of texture indices of CT images? A phantom study. *Med Phys* 2018; **45**: 1529–36. https://doi.org/10.1002/mp.12809
43. Chalkidou A, O'Doherty MJ, Marsden PK. False discovery rates in PET and CT studies with texture features: A systematic review. *PLoS One* 2015; **10**(5): e0124165. https://doi.org/10.1371/journal.pone.0124165
44. Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. *BMC Med Genomics* 2015; **8**: 33. https://doi.org/10.1186/s12920-015-0108-y
45. Marx V. Biology: the big challenges of big data. *Nature* 2013; **498**: 255–60. https://doi.org/10.1038/498255a
46. Cordeiro MS, Backes AR, Júnior AFD, Gonçalves EHG, de Oliveira JX. Fibrous dysplasia characterization using lacunarity analysis. *J Digit Imaging* 2016; **29**: 134–40. https://doi.org/10.1007/s10278-015-9815-3
47. Luo D, Zeng W, Chen J, Tang W. Deep learning for automatic image segmentation in stomatology and its clinical application. *Front Med Technol* 2021; **3**: 767836. https://doi.org/10.3389/fmedt.2021.767836
48. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* 2017; **28**: 1191–1206. https://doi.org/10.1093/annonc/mdx034
49. Peng Z, Wang Y, Wang Y, Jiang S, Fan R, Zhang H, et al. Application of radiomics and machine learning in head and neck cancers. *Int J Biol Sci* 2021; **17**: 475–86. https://doi.org/10.7150/ijbs.55716
50. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097: e123–30: . https://doi.org/10.1371/journal.pmed.1000097
51. University of York: Centre for Reviews and Dissemination. 2020.