Received: 5 August 2023 | Revised: 10 March 2024 | Accepted: 2 April 2024

DOI: 10.1111/clr.14271

ORIGINAL ARTICLE

CLINICAL ORAL IMPLANTS RESEARCH WILEY

Inter and intra-operator reliability of Lekholm and Zarb classification and proposal of a novel radiomic data-driven clustering for qualitative assessment of edentulous alveolar ridges

Giuseppe Troiano[1] | **Antonio Rapani**[2] | **Francesco Fanelli**[1] | **Federico Berton**[2] | **Marino Caroprese**[1] | **Teresa Lombardi**[3] | **Khrystyna Zhurakivska**[1] | **Claudio Stacchi**[2]

[1]Department of Clinical and Experimental Medicine, University of Foggia, Foggia, Italy

[2]Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, Italy

[3]Department of Health Sciences, University "Magna Græcia", Catanzaro, Italy

Correspondence
Claudio Stacchi, Department of Medical, Surgical and Health Sciences, University of Trieste, Corso Italia 121, 34170 Gorizia, Italy.
Email: claudio@stacchi.it

Abstract

Objectives: The present study was conducted to evaluate the reproducibility of Lekholm and Zarb classification system (L&Z) for bone quality assessment of edentulous alveolar ridges and to investigate the potential of a data-driven approach for bone quality classification.

Materials and Methods: Twenty-six expert clinicians were asked to classify 110 CBCT cross-sections according to L&Z classification (T0). The same evaluation was repeated after one month with the images put in a different order (T1). Intra- and inter-examiner agreement analyses were performed using Cohen's kappa coefficient (CK) and Fleiss' kappa coefficient (FK), respectively. Additionally, radiomic features extraction was performed from 3D edentulous ridge blocks derived from the same 110 CBCTs, and unsupervised clustering using 3 different clustering methods was used to identify patterns in the obtained data.

Results: Intra-examiner agreement between T0 and T1 was weak (CK 0.515). Inter-examiner agreement at both time points was minimal (FK at T0: 0.273; FK at T1: 0.243). The three different unsupervised clustering methods based on radiomic features aggregated the 110 CBCTs in three groups in the same way.

Conclusions: The results showed low agreement among clinicians when using L&Z classification, indicating that the system may not be as reliable as previously thought. The present study suggests the possible application of a reproducible data-driven approach based on radiomics for the classification of edentulous alveolar ridges, with potential implications for improving clinical outcomes. Further research is needed to determine the clinical significance of these findings and to develop more standardized and accurate methods for assessing bone quality of edentulous alveolar ridges.

KEYWORDS
artificial intelligence, big data, bone, clustering, data mining, dental implants, osseointegration, radiomics

1 | INTRODUCTION

In the past decades, oral implantology greatly improved treatment options for partially and totally edentulous patients (Albrektsson et al., 1986; Elani et al., 2018). Accuracy of preoperative diagnosis and treatment planning is crucial to improve both surgical and prosthetic procedures and enhance long-term clinical outcomes (Al-Ekrish et al., 2018; Chrcanovic et al., 2017). In particular, quantitative and qualitative assessments of alveolar bone are fundamental steps to optimize implant placement (Monje et al., 2015). Quantitative evaluation of the bone crest is necessary to determine the feasibility of implant surgery, selecting implants of appropriate dimensions, and planning any necessary bone augmentation procedure (Cicciù et al., 2023). It is typically performed by combining clinical examination and three-dimensional radiographic imaging (with or without computer-assisted tools) to obtain comprehensive information about bone crest dimensions and morphology. With the use of CT (computed tomography) scan is also possible to assess bone density and provide quantitative measurements expressed in Hounsfield units (HU) (Brooks, 1977; Norton & Gamble, 2001). However, it is important to note that bone density measurements obtained through CT scans may not directly correlate with measurements obtained through other methods like dual-energy X-ray absorptiometry (DXA), which is the gold standard for assessing bone density in other areas of the body (Adams, 2009; Ahmad et al., 2023; Kanis et al., 2000). In addition, CBCT (cone beam computed tomography) imaging has become increasingly popular in implantology field due to several advantages over conventional CT scans, including lower radiation dose, dedicated imaging for the oral and maxillofacial region, cost-effectiveness, efficient workflow, and reduced artefacts. CBCT scans do not provide HU like conventional CT scans, but grayscale or grey values that represent the radiodensity of the imaged tissues. However, previous studies highlighted a weak correlation between gray values and HU, together with the low predictive reliability of the grayscale, resulting in its questionable applicability for bone density assessment (Eguren et al., 2022; Gaur et al., 2022).

Consequently, in the absence of objective methods, qualitative assessment of the edentulous bone ridge involves, to some extent, a subjective interpretation. It relies on clinician's expertise, experience, and visual judgment to evaluate the characteristics and features of the alveolar ridge. Various classification systems have been proposed during the years for the qualitative categorization of residual bone crests, based on subjective evaluation of three-dimensional radiographic images (Al-Ekrish et al., 2018; Lekholm & Zarb, 1985; Misch, 1990; Trisi & Rao, 1999; Vercellotti & Vercellotti, 2009; Wang et al., 2023). However, the reproducibility and reliability of these classification systems have been a subject of debate among clinicians (Lindh et al., 2014; Ribeiro-Rotta et al., 2011; Shahlaie et al., 2003).

Radiomics, an advancing area in medical research, involves extracting quantitative metrics, known as radiomic features, from medical images like CT, CBCT, magnetic resonance imaging (MRI), or positron emission tomography (PET). These features encapsulate characteristics such as tissue shape, density, texture, and heterogeneity, providing valuable data that, when combined with other clinical information, may aid in solving complex clinical problems (Mayerhoefer et al., 2020). By analyzing features from medical images, radiomics may create accurate prediction models to support clinical decision-making (Felfli et al., 2023; O'Donnell et al., 2022). The potential applications of radiomics are broad and span various medical fields, including oncology, neurology, and cardiology (Cassinelli Petersen et al., 2022; Huellebrand et al., 2023; Lambin et al., 2017; Santos et al., 2023).

In head and neck area, radiomics has been mainly applied on CT and CBCT imaging for the automated diagnosis, segmentation, and classification of various conditions, including jaw cysts and tumors, cervical lymph node metastasis, salivary gland diseases, temporomandibular joint disorders, maxillary sinus issues, mandibular fractures, and facial deformities (Hung et al., 2022). Nevertheless, there has been a very limited utilization of radiomics potential in the field of oral implantology (Li et al., 2023; Troiano et al., 2023).

Radiomic data can be analyzed using machine learning techniques, which are currently widely used in interpreting complex, high-dimensional data, especially in health sciences research (Bo et al., 2023; Eckhardt et al., 2023). In machine learning, methods are broadly categorized into supervised and unsupervised learning. Supervised learning predicts outcomes based on labeled data, focusing on selecting key features from known outcomes. Conversely, unsupervised learning processes unlabeled data to unearth hidden patterns and similarities. Unsupervised learning includes clustering algorithms aiming to organize data into clusters, making it easier to understand and analyze complex datasets by revealing underlying relationships or categories within the dataset (Gao et al., 2023). K-means clustering, for instance, minimizes within-cluster variation and is used to identify distinct patient subgroups. Hierarchical clustering, on the other hand, does not require presetting the number of clusters but builds them based on dissimilarity measures, providing a flexible approach to data segmentation. Additionally, fuzzy-c-means clustering is a method where each data point belongs to a cluster to a degree specified by a membership level, allowing for a more nuanced grouping of data, particularly beneficial in situations where clear-cut boundaries between clusters are absent (Mingoti & Lima, 2006).

The present study has two main objectives: (i) to evaluate reliability and reproducibility of bone quality assessment performed by different clinicians using the currently most widespread classification system (Lekholm & Zarb, 1985) and (ii) to explore if radiomics application to CBCT images may lead to a data-driven approach for a qualitative assessment of edentulous alveolar ridges.

2 | MATERIALS AND METHODS

2.1 | Ethics

All the procedures of the present study were conducted following the recommendations of the Declaration of Helsinki as revised in Fortaleza (2013) for investigations with human subjects. The study protocol was approved by the Ethical Committee of the University of Trieste (No. 131/2023).

2.2 | Patients and CBCT scans selection

Two authors of the present study (A.R. and C.S.) examined a pool of anonymized CBCT images from partially edentulous patients >18 years old who authorized the use of their data for research purposes. Only patients who had not received any treatment for bone augmentation were included and investigators selected the first consecutive 110 radiographs presenting an acceptable standard for bone quality assessment. One site per patient was included, balancing between maxilla and mandible, resulting in 54 mandibular and 56 maxillary sites for analysis. All CBCT scans were acquired with the same device (NewTom VGi EVO, Cefla, Imola, Italy) and the same field of view (12 × 8). 3D Slicer software (https://www.slicer.org) was used to cut CBCT images at the level of the edentulous sites. The "volume rendering" command was selected, and the "crop" function was enabled to select the edentulous site. A total of 110 middle cross-sectional images of edentulous areas were uploaded to a Powerpoint file.

2.3 | Image analysis and classification

The assessment procedure generally followed the protocol outlined in a prior study evaluating intra- and inter-examiner agreement when assessing radiographic peri-implant bone level (Walton & Layton, 2020). Thirty expert clinicians and researchers active in the implantology field were contacted by phone or e-mail and informed about the objectives of the present investigation and the research protocol: 26 of them (13 academics; 13 private practitioners) agreed to participate in this project. A web-based calibration session was held prior to the beginning of the study. In this session, reference materials (radiographic images illustrating the 4 bone types according to L&Z classification) were provided to the participants, and a set of 8 CBCT cross-sections, chosen as examples, were collectively reviewed and discussed.

Examiners were then asked to assess bone quality of 110 CBCT cross-sectional images according to Lekholm and Zarb (L&Z) classification (T0) (Lekholm & Zarb, 1985). The same images, put in a different order to minimize recall bias, were re-evaluated by the same clinicians after 4 weeks (T1). Participants were asked to perform the two evaluation sessions on the same computer under the same light conditions and to choose only one of the four bone types for each image.

Intra- and inter-examiner agreement analyses were performed by one of the authors (A.R.) using Cohen's kappa coefficient (CK) and Fleiss' kappa coefficient (FK), respectively (McHugh, 2012).

2.4 | Radiomic features extraction

3D-block images were uploaded onto the LIFEx software (Nioche et al., 2018) for subsequent radiomic analysis. Volume of interest (VOI) was designed and manipulated using LIFEx, with a height of 10 mm (starting from the bone crest of the edentulous site), a thickness on the sagittal plane of 4 mm, and an extension on the axial plane including all the tissue between the two cortical laminae, simulating the site for the placement of a standard dental implant. Radiomic features were extracted using the "Texture-Features Extraction" command with default values for "Spatial Re-sampling" of 1 mm on the *X*, *Y*, and *Z* axes. Additionally, "relative resampling" was used for "intensity rescaling." VOI values were automatically re-scaled to between the minimum and maximum value of the ROI content. Extracted features were saved automatically by the software in CSV format in the "VOI-Value" folder.

2.5 | Unsupervised learning algorithms

The extracted radiomic features dataset was analyzed by applying three different clustering algorithms: K-means, hierarchical, and fuzzy C-mean. Random initialization was applied for the initial placement of centroids in K-means clustering. In addition, the maximum number of iterations within each algorithm run was set to 30,000, and the convergence criterion was based on minimizing the within-cluster sum of squares. An average linkage method was applied to hierarchical clustering and Euclidean distance was used as a measure of similarity between different datapoints. For fuzzy C-means clustering the fuzziness parameters was set equal to 2 with a maximum number of iterations 1000. We used the Silhouette method, Calinski-Harabasz index, and Davies-Bouldin index to evaluate the clustering separation, while the thresher and elbow methods were used to determine the optimal number of clusters. All the clustering analyses were performed using the Euclidean distance. The analysis was performed using Jupyter Notebook, which allowed for efficient and reproducible data analysis. By utilizing multiple clustering methods and indices, we aimed to increase the reliability and robustness of the clustering results, thus providing a more comprehensive understanding of the data.

3 | RESULTS

3.1 | Examiner agreement for L&Z classification

Intra- and inter-examiner agreement in assessing bone quality according to L&Z classification was evaluated. Intra-examiner analysis showed weak agreement between T0 and T1 (CK: median 0.515; IQR 0.253). Individual values were reported in Figure 1. Mann–Whitney *U* test showed no significant differences between private practitioners (CK: median 0.447, IQR 0.274) and academics (CK: median 0.556; IQR 0.188) ($p = .92$). Inter-examiner analysis highlighted minimal agreement among examiners at both time points (FK at T0: 0.273; FK at T1: 0.243). No significant differences were present both at T0 and at T1 between private practitioners (FK at T0: 0.274; FK at T1: 0.217) and academics (FK at T0: 0.281; FK at T1: 0.256) (Figure 2).

3.2 | Clustering analysis of radiomic features

Clustering analysis was performed on the radiomic features to identify distinct clusters within the dataset. Using the elbow and thresher

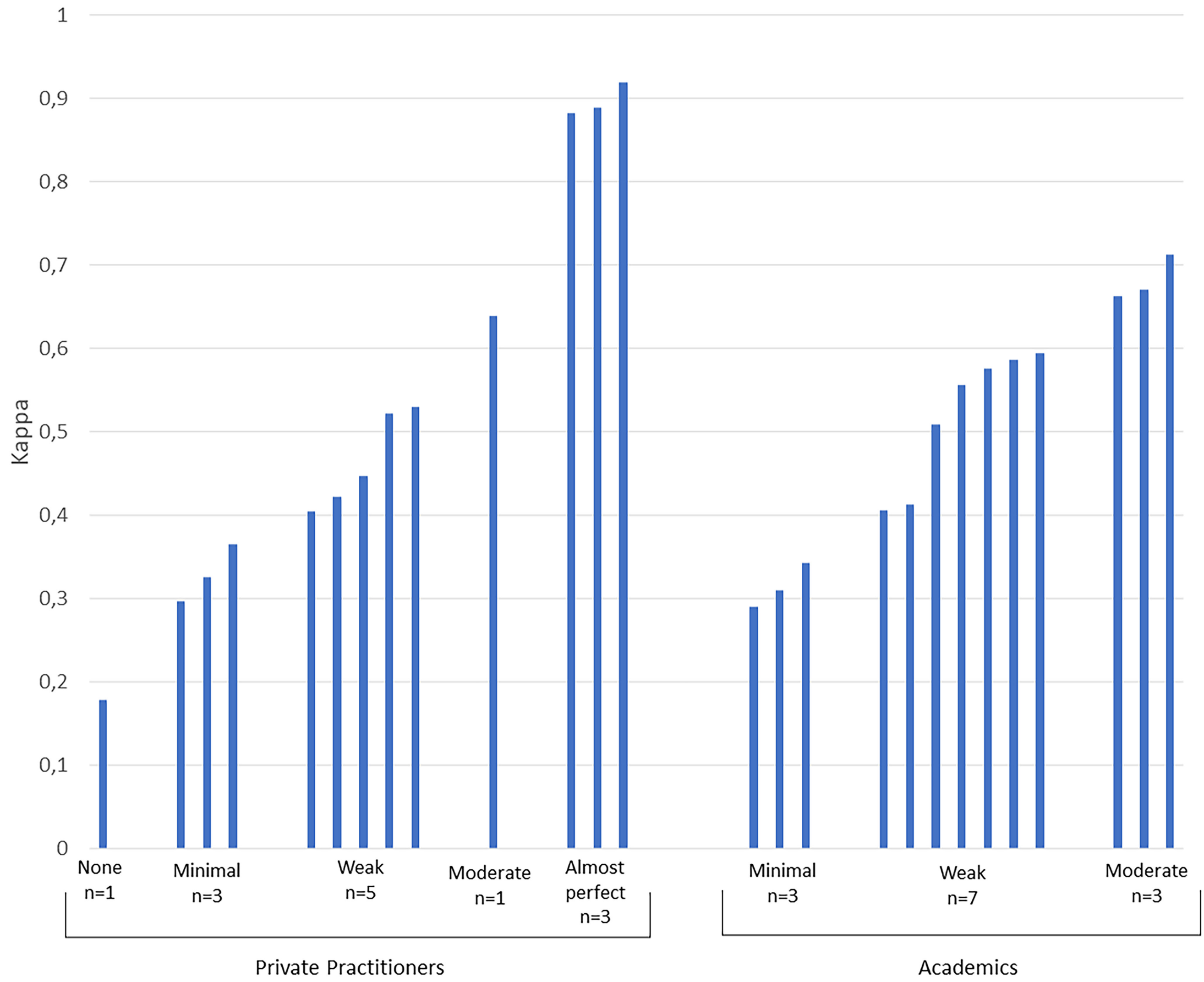

FIGURE 1 Individual intra-examiner consistency assessing bone quality using the Lekholm and Zarb classification, quantified by Cohen's kappa. A higher kappa value indicates greater agreement with repeated assessments, underscoring the evaluator's reliability in consistently interpreting bone quality criteria.

methods (Wang et al., 2018), an optimal number of 3 cluster was chosen for subsequent separation (Figure 3). Subsequently, three clustering methods were applied: K-means (Figure 4), hierarchical (Figure 5), and fuzzy C-means clustering (Petegrosso et al., 2020; Pfeifer & Schimek, 2021; Wu et al., 2015). Interestingly, the three methods aggregated the three groups in the same way. Cluster 1 was mainly represented in the upper jaw, while Cluster 2 and Cluster 3 had a homogeneous distribution in mandible and maxilla (Table 1). The evaluation indices for the clustering methods were consistent, with all three methods producing an average Silhouette score of 0.88, a Calinski-Harabasz Index of 460.66, and a Davies-Bouldin Index of 0.30 (Lovmar et al., 2005; Surangsrirat et al., 2022; Zhang et al., 2022).

4 | DISCUSSION

This study provides important insights into the reliability and reproducibility of Lekholm and Zarb classification system, which is currently the most widespread tool to assess bone quality of edentulous alveolar ridges. The low intra- and inter-examiner agreement observed in our study suggests that L&Z classification may not provide a reliable basis for research purposes, for communicating with colleagues or for treatment planning (Viera & Garrett, 2005). The present results are in agreement with previous studies which, even if conducted with only two examiners, showed significant disagreement in the assessment of bone quality using L&Z classification (Shahlaie et al., 2003; Shapurian et al., 2006). Data of the present study also align with those from prior systematic reviews, which similarly concluded that there is limited evidence supporting the effectiveness of clinical methods for evaluating jawbone tissue before implant placement (Chrcanovic et al., 2017; Ribeiro-Rotta et al., 2007; Ribeiro-Rotta et al., 2011). The observed variability in agreement underscores the subjective nature of current assessment methodologies and the inherent challenge in achieving uniformity across different examiners, even within a controlled study environment. This

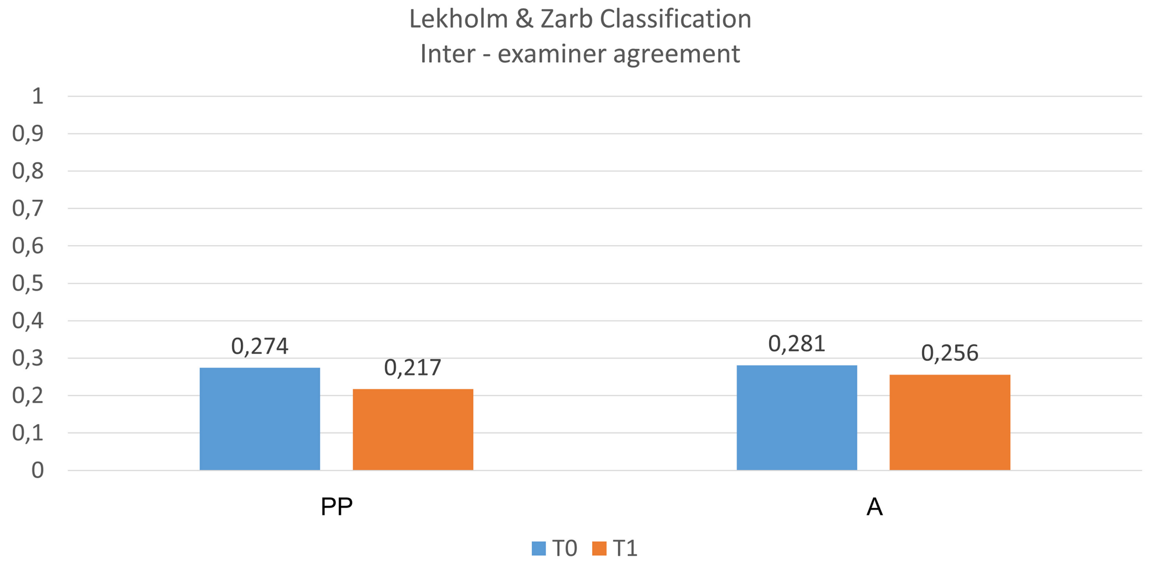

FIGURE 2 Inter-examiner agreement in assessing bone quality at two time points (baseline and 4 weeks later), using Fleiss' kappa. The varying levels of agreement underscore the challenges in achieving a uniform assessment of bone quality across different clinical backgrounds, emphasizing the need for objective diagnostic tools. A, academics; PP, private practitioners; T0, baseline; T1, 4 weeks after T0.

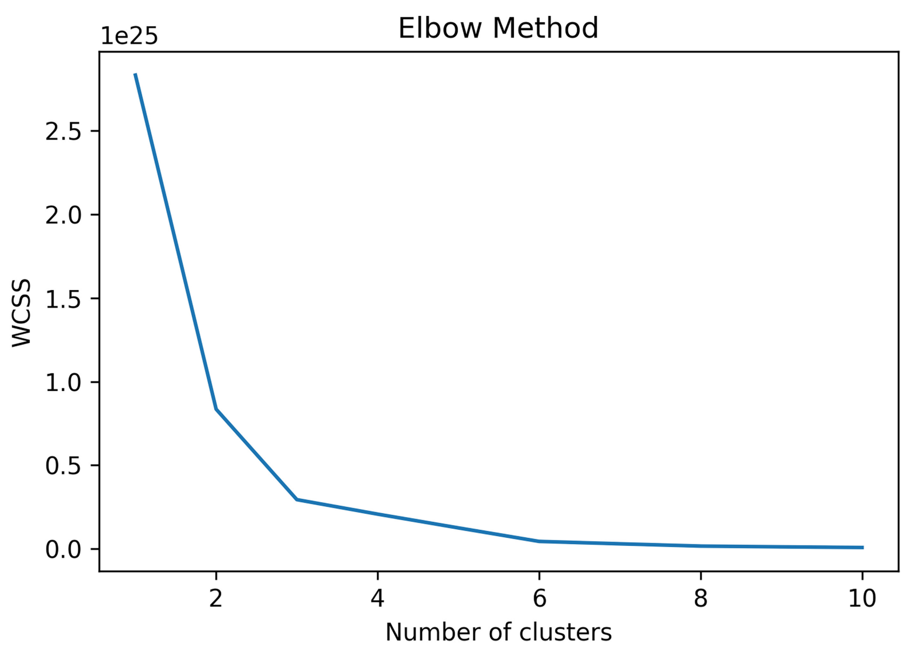

FIGURE 3 Visualization of the Elbow Method analysis to identify the most suitable number of clusters for our dataset. The X-axis shows the number of clusters tested and the Y-axis the within-cluster sum of squares (WCSS), illustrating the efficiency of grouping similar data. The 'elbow' point indicates the optimal number of clusters (in this case $n = 3$), balancing specificity and generalizability of our radiomic feature analysis for clinical application.

variability is not only a reflection of individual interpretation differences but also highlights the critical need for standardized criteria and training in the evaluation process. This is a significant finding, given the increasing use of implantology for the treatment of edentulous patients, where accurate assessment of alveolar bone quality is essential for implant site preparation and subsequent implant placement (French et al., 2019; Heitz-Mayfield et al., 2018; Stacchi et al., 2023). The quality of the surrounding bone has a significant influence on implant primary stability and affects the drilling protocol during implant site preparation (Perez-Pevida et al., 2020). Dense bone requires a slower drilling speed under copious irrigation to prevent overheating, together with the creation of an implant site diameter close to the actual implant diameter to allow firm stabilization of the fixture but without reaching excessive torque during insertion (Toia et al., 2017). Excessive insertional torque can potentially lead to surrounding

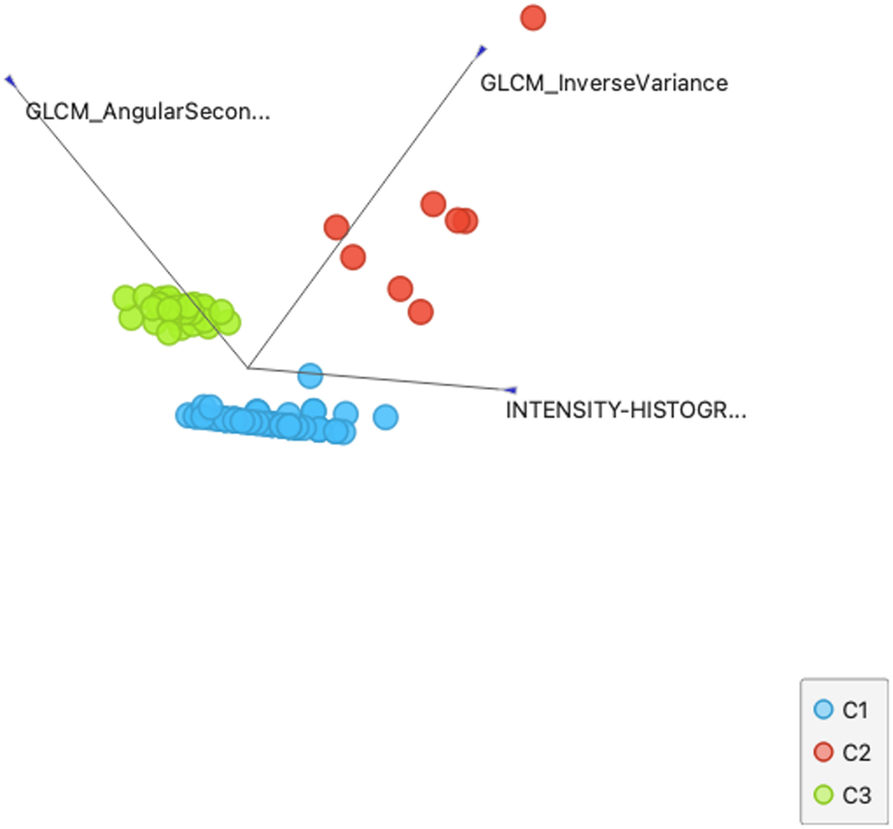

FIGURE 4 Graphic visualization of the results of K-means clustering on the radiomic features dataset using a linear projection. The entire dataset (n = 110 items) was clustered into three distinct groups based on the most informative radiomic features, and each point is colored according to its cluster label (C1: cluster 1; C2: cluster 2; C3: cluster 3).

bone damage and micro-fractures, eliciting bone resorption, and to irreversible mechanical deformation of the implant connection (Cha et al., 2015; Coyac et al., 2019; Teixeira et al., 2015). Conversely, in low-quality bone, an undersized implant site preparation is mandatory to enhance initial implant stability by creating a press-fit between the implant and the surrounding bone (Tabassum et al., 2010). Furthermore, an accurate preoperative assessment of bone quality should guide the clinician in choosing an implant with a suitable macro-geometry (i.e., parallel-walled vs. tapered and shallow vs. deep threads) for the specific situation (Abuhussein et al., 2010; Aleo et al., 2012).

The findings of the present study underline the importance of integrating advanced analytics into clinical practice, facilitating more consistent and reproducible evaluations across practitioners. We believe that addressing this variability is crucial for advancing the reliability of diagnostic processes in oral surgery and implantology. Radiomics, in particular, could represent a possible promising alternative for bone quality evaluation. The analysis of 110 CBCT cross-sections from edentulous alveolar ridges was performed with the application of unsupervised machine learning techniques, focusing on their radiomic features. Unsupervised clustering methods are effective at uncovering latent patterns and structures in data that may not be immediately evident. Techniques like clustering and dimensionality reduction are central to unsupervised learning, simplifying intricate datasets and aiding in the discovery of novel risk factors, advancing prevention strategies, and contributing to personalized healthcare (Lopez et al., 2018).

The 110 examined CBCT cross-sections were aggregated in the same way into three distinct clusters by using three different clustering methods (K-means, hierarchical and fuzzy C-means). Consistent outcomes among diverse unsupervised learning methods contribute to robustness, minimizing bias and improving validation and generalizability. This enhances the reliability and interpretability of the identified clusters, considering the heterogeneity in the approaches used.

Radiomics analysis may offer a more objective and reproducible method for the classification of edentulous alveolar ridges, potentially improving the accuracy of treatment planning for edentulous patients. However, further translational research is absolutely needed prior to apply this classification system to the clinical practice. Bone quality evaluation based on radiomic features should be possibly associated with various characteristics and properties of clinical interest of the alveolar bone (e.g. bone density, osseointegration potential, susceptibility to bone resorption) to allow a meaningful application of this approach to the daily practice.

The results of the present study may have several potential future applications. First, radiomic analysis could enable the development of automated tools based on radiomic features to classify edentulous alveolar ridges in an objective way. This could

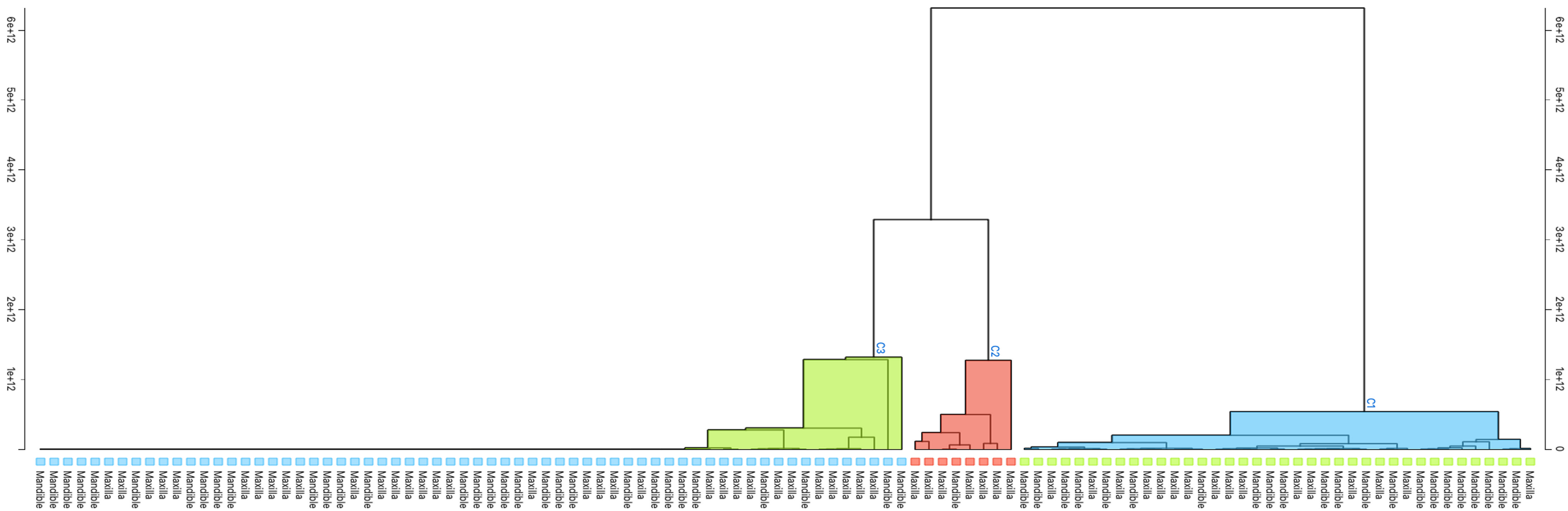

FIGURE 5 Dendrogram illustrating the results of hierarchical clustering, with the Y-axis representing the dissimilarity measure, indicating that clusters merge at variable levels of proximity, which reflects the underlying structure and potential groupings within this dataset. The height of each branch in the dendrogram represents the dissimilarity between clusters, with the longer branches indicating greater dissimilarity. The colors on the branches represent the clusters produced by hierarchical clustering for comparison with the K-means clustering results (colors on the caption Maxilla/Mandible).

TABLE 1 Distribution of clustered edentulous sites in upper and lower arch.

Arch	Cluster 1	Cluster 2	Cluster 3	Total
Mandible	1	22	31	54
	12.5%	57.9%	48.4%	49.1%
Maxilla	7	16	33	56
	87.5%	42.1%	51.6%	50.9%
Total	8	38	64	110
	100%	100%	100%	100%

lead to more efficient and personalized treatment planning for implant-supported rehabilitation of edentulous patients. Second, radiomic analysis may also provide a more comprehensive and detailed assessment of the alveolar ridge compared to traditional clinical assessment methods, enabling clinicians to detect subtle variations and anomalies that could influence implant placement and stability.

In conclusion, the extremely low agreement observed among clinicians when using Lekholm and Zarb classification system highlights the need for more objective and reproducible approaches to evaluate bone quality of edentulous alveolar ridges. The potential applications of radiomic analysis in the development of automated classification tools and the comprehensive assessment of alveolar ridges suggest that this approach has the potential to significantly improve accuracy and personalization of treatment planning for edentulous patients. However, further studies are needed to validate the present findings and to develop standardized and accurate methods to assess bone quality of edentulous alveolar ridges, giving information of clinical relevance.

AUTHOR CONTRIBUTIONS

G.T.:Conceptualization, methodology, formal analysis; writing – original draft; A.R.: investigation, writing – review and editing, and data curation; F.F.: investigation, visualization, and writing – review and editing; F.B.: software, validation, and writing – review and editing; M.C.: investigation, visualization, and writing – review and editing; T.L.: resources, data curation, and writing – review and editing; K.Z.: methodology, validation, and writing – review and editing; C.S.: conceptualization, resources, supervision, and writing – original draft.

ACKNOWLEDGMENTS

Authors wish to thank Roberto Modonutti and Michele Mauro for their precious help in collecting and processing the radiographic images used in the present study.

FUNDING INFORMATION

The present study was self-funded.

CONFLICT OF INTEREST STATEMENT

The authors do not have any financial interests, either directly or indirectly, in the products or information listed in this paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

PATIENT CONSENT STATEMENT

The study was conducted on a pool of anonymized CBCT images of patients who authorized the use of their data for research purposes.

ORCID

Giuseppe Troiano https://orcid.org/0000-0001-5647-4414
Antonio Rapani https://orcid.org/0000-0002-9767-4401
Federico Berton https://orcid.org/0000-0001-6359-9143
Marino Caroprese https://orcid.org/0000-0003-0507-9793
Teresa Lombardi https://orcid.org/0000-0002-0334-6029
Khrystyna Zhurakivska https://orcid.org/0000-0002-2153-0778
Claudio Stacchi https://orcid.org/0000-0003-4017-4980

REFERENCES

Abuhussein, H., Pagni, G., Rebaudi, A., & Wang, H. L. (2010). The effect of thread pattern upon implant osseointegration. *Clinical Oral Implants Research*, *21*(2), 129–136. https://doi.org/10.1111/j.1600-0501.2009.01800.x

Adams, J. E. (2009). Quantitative computed tomography. *European Journal of Radiology*, *71*(3), 415–424. https://doi.org/10.1016/j.ejrad.2009.04.074

Ahmad, A., Crawford, C. H., 3rd, Glassman, S. D., Dimar, J. R., 2nd, Gum, J. L., & Carreon, L. Y. (2023). Correlation between bone density measurements on CT or MRI versus DEXA scan: A systematic review. *N Am Spine Soc J*, *14*, 100204. https://doi.org/10.1016/j.xnsj.2023.100204

Albrektsson, T., Zarb, G., Worthington, P., & Eriksson, A. R. (1986). The long-term efficacy of currently used dental implants: A review and proposed criteria of success. *The International Journal of Oral & Maxillofacial Implants*, *1*(1), 11–25.

Al-Ekrish, A. A., Widmann, G., & Alfadda, S. A. (2018). Revised, computed tomography-based Lekholm and Zarb jawbone quality classification. *The International Journal of Prosthodontics*, *31*(4), 342–345. https://doi.org/10.11607/ijp.5714

Aleo, E., Varvara, G., Scarano, A., Sinjari, B., & Murmura, G. (2012). Comparison of the primary stabilities of conical and cylindrical endosseous dental implants: An in-vitro study. *Journal of Biological Regulators and Homeostatic Agents*, *26*(1), 89–96. https://www.ncbi.nlm.nih.gov/pubmed/22475100

Bo, Z., Chen, B., Zhao, Z., He, Q., Mao, Y., Yang, Y., Yao, F., Yang, Y., Chen, Z., Yang, J., Yu, H., Ma, J., Wu, L., Chen, K., Wang, L., Wang, M., Shi, Z., Yao, X., Dong, Y., … Chen, G. (2023). Prediction of response to lenvatinib monotherapy for unresectable hepatocellular carcinoma by machine learning radiomics: A multicenter cohort study. *Clinical Cancer Research*, *29*(9), 1730–1740. https://doi.org/10.1158/1078-0432.CCR-22-2784

Brooks, R. A. (1977). A quantitative theory of the Hounsfield unit and its application to dual energy scanning. *Journal of Computer Assisted Tomography*, *1*(4), 487–493. https://doi.org/10.1097/00004728-197710000-00016

Cassinelli Petersen, G. I., Shatalov, J., Verma, T., Brim, W. R., Subramanian, H., Brackett, A., Bahar, R. C., Merkaj, S., Zeevi, T., Staib, L. H., Cui, J., Omuro, A., Bronen, R. A., Malhotra, A., & Aboian, M. S. (2022). Machine learning in differentiating gliomas from primary CNS lymphomas: A systematic review, reporting quality, and risk of bias assessment. *AJNR. American Journal of Neuroradiology*, *43*(4), 526–533. https://doi.org/10.3174/ajnr.A7473

Cha, J. Y., Pereira, M. D., Smith, A. A., Houschyar, K. S., Yin, X., Mouraret, S., Brunski, J. B., & Helms, J. A. (2015). Multiscale analyses of the bone-implant interface. *Journal of Dental Research*, *94*(3), 482–490. https://doi.org/10.1177/0022034514566029

Chrcanovic, B., Albrektsson, T., & Wennerberg, A. (2017). Bone quality and quantity and dental implant failure: A systematic review and meta-analysis. *The International Journal of Prosthodontics*, *30*(3), 219–237. https://doi.org/10.11607/ijp.5142

Cicciù, M., Pratella, U., Fiorillo, L., Bernardello, F., Perillo, F., Rapani, A., Stacchi, C., & Lombardi, T. (2023). Influence of buccal and palatal bone thickness on post-surgical marginal bone changes around implants placed in posterior maxilla: A multi-centre prospective study. *BMC Oral Health*, *23*(1), 309. https://doi.org/10.1186/s12903-023-02991-3

Coyac, B. R., Leahy, B., Salvi, G., Hoffmann, W., Brunski, J. B., & Helms, J. A. (2019). A preclinical model links osseo-densification due to misfit and osseo-destruction due to stress/strain. *Clinical Oral Implants Research*, *30*(12), 1238–1249. https://doi.org/10.1111/clr.13537

Eckhardt, C. M., Madjarova, S. J., Williams, R. J., Ollivier, M., Karlsson, J., Pareek, A., & Nwachukwu, B. U. (2023). Unsupervised machine learning methods and emerging applications in healthcare. *Knee Surgery, Sports Traumatology, Arthroscopy*, *31*(2), 376–381. https://doi.org/10.1007/s00167-022-07233-7

Eguren, M., Holguin, A., Diaz, K., Vidalon, J., Linan, C., Pacheco-Pereira, C., & Lagravere Vich, M. O. (2022). Can gray values be converted to Hounsfield units? A systematic review. *Dento Maxillo Facial Radiology*, *51*(1), 20210140. https://doi.org/10.1259/dmfr.20210140

Elani, H. W., Starr, J. R., Da Silva, J. D., & Gallucci, G. O. (2018). Trends in dental implant use in the U.S., 1999-2016, and projections to 2026. *Journal of Dental Research*, *97*(13), 1424–1430. https://doi.org/10.1177/0022034518792567

Felfli, M., Liu, Y., Zerka, F., Voyton, C., Thinnes, A., Jacques, S., Iannessi, A., & Bodard, S. (2023). Systematic review, meta-analysis and radiomics quality score assessment of CT radiomics-based models predicting tumor EGFR mutation status in patients with non-small-cell lung cancer. *International Journal of Molecular Sciences*, *24*(14), 11433. https://doi.org/10.3390/ijms241411433

French, D., Grandin, H. M., & Ofec, R. (2019). Retrospective cohort study of 4,591 dental implants: Analysis of risk indicators for bone loss and prevalence of peri-implant mucositis and peri-implantitis. *Journal of Periodontology*, *90*(7), 691–700. https://doi.org/10.1002/JPER.18-0236

Gao, C. X., Dwyer, D., Zhu, Y., Smith, C. L., Du, L., Filia, K. M., Bayer, J., Menssink, J. M., Wang, T., Bergmeir, C., Wood, S., & Cotton, S. M. (2023). An overview of clustering methods with guidelines for application in mental health research. *Psychiatry Research*, *327*, 115265. https://doi.org/10.1016/j.psychres.2023.115265

Gaur, A., Dhillon, M., Puri, N., Sethi Ahuja, U., & Rathore, A. (2022). Questionable accuracy of CBCT in determining bone density: A comparative CBCT-CT in vitro study. *Dent Med Probl*, *59*(3), 413–419. https://doi.org/10.17219/dmp/143504

Heitz-Mayfield, L. J., Aaboe, M., Araujo, M., Carrion, J. B., Cavalcanti, R., Cionca, N., Cochran, D., Darby, I., Funakoshi, E., Gierthmuehlen, P. C., Hashim, D., Jahangiri, L., Kwon, Y., Lambert, F., Layton, D. M., Lorenzana, E. R., McKenna, G., Mombelli, A., Müller, F., … Yeo, A. (2018). Group 4 ITI consensus report: Risks and biologic complications associated with implant dentistry. *Clinical Oral Implants Research*, *29*(Suppl 16), 351–358. https://doi.org/10.1111/clr.13307

Huellebrand, M., Jarmatz, L., Manini, C., Laube, A., Ivantsits, M., Schulz-Menger, J., Nordmeyer, S., Harloff, A., Hansmann, J., Kelle, S., & Hennemuth, A. (2023). Radiomics-based aortic flow profile characterization with 4D phase-contrast MRI. *Frontiers in Cardiovascular Medicine*, *10*, 1102502. https://doi.org/10.3389/fcvm.2023.1102502

Hung, K. F., Ai, Q. Y. H., Wong, L. M., Yeung, A. W. K., Li, D. T. S., & Leung, Y. Y. (2022). Current applications of deep learning and radiomics on CT and CBCT for maxillofacial diseases. *Diagnostics (Basel)*, *13*, 110. https://doi.org/10.3390/diagnostics13010110

Kanis, J. A., Johnell, O., Oden, A., Jonsson, B., De Laet, C., & Dawson, A. (2000). Risk of hip fracture according to the World Health Organization criteria for osteopenia and osteoporosis. *Bone*, *27*(5), 585–590. https://doi.org/10.1016/s8756-3282(00)00381-1

Lambin, P., Leijenaar, R. T. H., Deist, T. M., Peerlings, J., de Jong, E. E. C., van Timmeren, J., Sanduleanu, S., Larue, R. T. H. M., Even, A. J. G., Jochems, A., van Wijk, Y., Woodruff, H., van Soest, J., Lustberg, T., Roelofs, E., van Elmpt, W., Dekker, A., Mottaghy, F. M., Wildberger, J. E., & Walsh, S. (2017). Radiomics: The bridge between medical imaging and personalized medicine. *Nature Reviews. Clinical Oncology*, *14*(12), 749–762. https://doi.org/10.1038/nrclinonc.2017.141

Lekholm, U., & Zarb, G. A. (1985). Patient selection and preparation. In P. I. Brånemark, G. A. Zarb, & T. Albrektsson (Eds.), *Tissue-integrated prostheses: Osseointegration in clinical dentistry* (pp. 199–209). Quintessence Publishing Co.

Li, J., Jin, F., Wang, R., Shang, X., Yang, P., Zhu, Y., Tsoi, J. K. H., Chan, K., & Wang, S. (2023). Guided bone regeneration in a periodontally compromised individual with autogenous tooth bone graft: A radiomics analysis. *J Funct Biomater*, *14*, 220. https://doi.org/10.3390/jfb14040220

Lindh, C., Oliveira, G. H., Leles, C. R., do Carmo Matias Freire, M., & Ribeiro-Rotta, R. F. (2014). Bone quality assessment in routine dental implant treatment among Brazilian and Swedish specialists. *Clinical Oral Implants Research*, *25*(9), 1004–1009. https://doi.org/10.1111/clr.12221

Lopez, C., Tucker, S., Salameh, T., & Tucker, C. (2018). An unsupervised machine learning method for discovering patient clusters based on genetic signatures. *Journal of Biomedical Informatics*, *85*, 30–39. https://doi.org/10.1016/j.jbi.2018.07.004

Lovmar, L., Ahlford, A., Jonsson, M., & Syvanen, A. C. (2005). Silhouette scores for assessment of SNP genotype clusters. *BMC Genomics*, *6*, 35. https://doi.org/10.1186/1471-2164-6-35

Mayerhoefer, M. E., Materka, A., Langs, G., Häggström, I., Szczypiński, P., Gibbs, P., & Cook, G. (2020). Introduction to radiomics. *Journal of Nuclear Medicine*, *61*(4), 488–495. https://doi.org/10.2967/jnumed.118.222893

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)*, *22*(3), 276–282. https://www.ncbi.nlm.nih.gov/pubmed/23092060

Mingoti, S. A., & Lima, J. O. (2006). Comparing SOM neural network with Fuzzyc-means, K-means and traditional hierarchical clustering algorithms. *European Journal of Operational Research*, *174*(3), 1742–1759.

Misch, C. E. (1990). Density of bone: Effect on treatment plans, surgical approach, healing, and progressive bone loading. *The International Journal of Oral Implantology*, *6*(2), 23–31.

Monje, A., Chan, H. L., Galindo-Moreno, P., Elnayef, B., Suarez-Lopez del Amo, F., Wang, F., & Wang, H. L. (2015). Alveolar bone architecture: A systematic review and meta-analysis. *Journal of Periodontology*, *86*(11), 1231–1248. https://doi.org/10.1902/jop.2015.150263

Nioche, C., Orlhac, F., Boughdad, S., Reuze, S., Goya-Outi, J., Robert, C., Pellot-Barakat, C., Soussan, M., Frouin, F., & Buvat, I. (2018). LIFEx: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Research*, *78*(16), 4786–4789. https://doi.org/10.1158/0008-5472.CAN-18-0125

Norton, M. R., & Gamble, C. (2001). Bone classification: An objective scale of bone density using the computerized tomography scan. *Clinical Oral Implants Research*, *12*(1), 79–84. https://doi.org/10.1034/j.1600-0501.2001.012001079.x

O'Donnell, J. P. M., Gasior, S. A., Davey, M. G., O'Malley, E., Lowery, A. J., McGarry, J., O'Connell, A. M., Kerin, M. J., & McCarthy, P. (2022). The accuracy of breast MRI radiomic methodologies in predicting pathological complete response to neoadjuvant chemotherapy: A systematic review and network meta-analysis. *European Journal of Radiology*, *157*, 110561. https://doi.org/10.1016/j.ejrad.2022.110561

Perez-Pevida, E., Cherro, R., Camps-Font, O., & Pique, N. (2020). Effects of drilling protocol and bone density on the stability of implants according to different macrogeometries of the implant used: Results of an in vitro study. *The International Journal of Oral & Maxillofacial Implants*, *35*(5), 955–964. https://doi.org/10.11607/jomi.8176

Petegrosso, R., Li, Z., & Kuang, R. (2020). Machine learning and statistical methods for clustering single-cell RNA-sequencing data. *Briefings in Bioinformatics*, *21*(4), 1209–1223. https://doi.org/10.1093/bib/bbz063

Pfeifer, B., & Schimek, M. G. (2021). A hierarchical clustering and data fusion approach for disease subtype discovery. *Journal of Biomedical Informatics*, *113*, 103636. https://doi.org/10.1016/j.jbi.2020.103636

Ribeiro-Rotta, R. F., Lindh, C., Pereira, A. C., & Rohlin, M. (2011). Ambiguity in bone tissue characteristics as presented in studies on dental implant planning and placement: A systematic review. *Clinical Oral Implants Research*, *22*(8), 789–801. https://doi.org/10.1111/j.1600-0501.2010.02041.x

Ribeiro-Rotta, R. F., Lindh, C., & Rohlin, M. (2007). Efficacy of clinical methods to assess jawbone tissue prior to and during endosseous dental implant placement: A systematic literature review. *The International Journal of Oral & Maxillofacial Implants*, *22*(2), 289–300.

Santos, G. N. M., da Silva, H. E. C., Ossege, F. E. L., Figueiredo, P. T. S., Melo, N. S., Stefani, C. M., & Leite, A. F. (2023). Radiomics in bone pathology of the jaws. *Dento Maxillo Facial Radiology*, *52*(1), 20220225. https://doi.org/10.1259/dmfr.20220225

Shahlaie, M., Gantes, B., Schulz, E., Riggs, M., & Crigger, M. (2003). Bone density assessments of dental implant sites: 1. Quantitative computed tomography. *The International Journal of Oral & Maxillofacial Implants*, *18*(2), 224–231.

Shapurian, T., Damoulis, P. D., Reiser, G. M., Griffin, T. J., & Rand, W. M. (2006). Quantitative evaluation of bone density using the Hounsfield index. *The International Journal of Oral & Maxillofacial Implants*, *21*(2), 290–297.

Stacchi, C., Lamazza, L., Rapani, A., Troiano, G., Messina, M., Antonelli, A., Giudice, A., & Lombardi, T. (2023). Marginal bone changes around platform-switched conical connection implants placed 1 or 2 mm subcrestally: A multicenter crossover randomized controlled trial. *Clinical Implant Dentistry and Related Research*, *25*(2), 398–408. https://doi.org/10.1111/cid.13186

Surangsrirat, D., Sri-Iesaranusorn, P., Chaiyaroj, A., Vateekul, P., & Bhidayasiri, R. (2022). Parkinson's disease severity clustering based on tapping activity on mobile device. *Scientific Reports*, *12*(1), 3142. https://doi.org/10.1038/s41598-022-06572-2

Tabassum, A., Meijer, G. J., Wolke, J. G., & Jansen, J. A. (2010). Influence of surgical technique and surface roughness on the primary stability of an implant in artificial bone with different cortical thickness: A laboratory study. *Clinical Oral Implants Research*, *21*(2), 213–220. https://doi.org/10.1111/j.1600-0501.2009.01823.x

Teixeira, A. B., Shimano, A. C., Macedo, A. P., Valente, M. L., & dos Reis, A. C. (2015). Influence of torsional strength on different types of dental implant platforms. *Implant Dentistry*, *24*(3), 281–286. https://doi.org/10.1097/ID.0000000000000247

Toia, M., Stocchero, M., Cecchinato, F., Corra, E., Jimbo, R., & Cecchinato, D. (2017). Clinical considerations of adapted drilling protocol by bone quality perception. *The International Journal of Oral & Maxillofacial Implants*, *32*(6), 1288–1295. https://doi.org/10.11607/jomi.5881

Trisi, P., & Rao, W. (1999). Bone classification: Clinical-histomorphometric comparison. *Clinical Oral Implants Research*, *10*(1), 1–7.

Troiano, G., Fanelli, F., Rapani, A., Zotti, M., Lombardi, T., Zhurakivska, K., & Stacchi, C. (2023). Can radiomic features extracted from intraoral radiographs predict physiological bone remodelling around dental implants? *A Hypothesis-Generating Study. J Clin Periodontol*, *50*(7), 932–941. https://doi.org/10.1111/jcpe.13797

Vercellotti, T., & Vercellotti, G. (2009). New bone classification for analysis of the single surgical site. In *Vercellotti, T. Essentials in Piezosurgery: Clinical advantages in dentistry* (pp. 91–94). Quintessence Publishing Co.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, *37*(5), 360–363. https://www.ncbi.nlm.nih.gov/pubmed/15883903

Walton, T. R., & Layton, D. M. (2020). Intra- and inter-examiner agreement when assessing radiographic implant bone levels: Differences related to brightness, accuracy, participant demographics and implant characteristics. *Clinical Oral Implants Research*, *29*(7), 756–771. https://doi.org/10.1111/clr.13290

Wang, M., Abrams, Z. B., Kornblau, S. M., & Coombes, K. R. (2018). Thresher: Determining the number of clusters while removing outliers. *BMC Bioinformatics*, *19*(1), 9. https://doi.org/10.1186/s12859-017-1998-9

Wang, S. H., Hsu, J. T., Fuh, L. J., Peng, S. L., Huang, H. L., & Tsai, M. T. (2023). New classification for bone type at dental implant sites: A dental computed tomography study. *BMC Oral Health*, *23*(1), 324. https://doi.org/10.1186/s12903-023-03039-2

Wu, Y., Duan, H., & Du, S. (2015). Multiple fuzzy c-means clustering algorithm in medical diagnosis. *Technology and Health Care*, *23*(Suppl 2), S519–S527. https://doi.org/10.3233/THC-150989

Zhang, W., Yue, Z., Ye, J., Xu, H., Wang, Y., Zhang, X., & Xi, L. (2022). Modulation format identification using the Calinski-Harabasz index. *Applied Optics*, *61*(3), 851–857. https://doi.org/10.1364/AO.448043

How to cite this article: Troiano, G., Rapani, A., Fanelli, F., Berton, F., Caroprese, M., Lombardi, T., Zhurakivska, K., & Stacchi, C. (2024). Inter and intra-operator reliability of Lekholm and Zarb classification and proposal of a novel radiomic data-driven clustering for qualitative assessment of edentulous alveolar ridges. *Clinical Oral Implants Research*, *35*, 729–738. https://doi.org/10.1111/clr.14271