Applications of deep learning in periodontal disease diagnosis and management: a systematic review and critical appraisal

Elisabetta Ferrara[1], **Biagio Rapone**[2], **Alessandro D'Albenzio**[3]

[1]Department of Human Sciences, Law, and Economics, Telematic University "Leonardo da Vinci", UNIDAV, Chieti, Italy; [2]Interdisciplinary Department of Medicine, University of Bari "Aldo Moro", Bari, Italy; [3]Complex Operative Unit of Pathological Addiction, Addiction Service, Chieti, Italy

Contributions: (I) Conception and design: E Ferrara; (II) Administrative support: E Ferrara, B Rapone; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: E Ferrara, A D'Albenzio; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Elisabetta Ferrara, Doctoral Student. Department of Human Sciences, Law, and Economics, Telematic University "Leonardo da Vinci", UNIDAV, Torrevecchia Teatina, 66100 Chieti, Italy. Email: elisabetta.ferrara@unidav.it.

Background: Periodontal disease remains a significant global health challenge, with traditional diagnostic methods often limited by subjectivity and time constraints. Recent advances in artificial intelligence (AI) and deep learning technologies have shown promise in various medical applications, potentially offering more objective and efficient approaches to periodontal diagnosis and management. This systematic review aimed to synthesize and critically evaluate the current body of evidence regarding the application of deep learning methodologies in the diagnosis and management of periodontal disease, with a focus on their potential to enhance clinical decision-making and patient outcomes.

Methods: A comprehensive literature search was conducted across PubMed, Scopus, and IEEE Xplore databases, encompassing studies published between January 1, 2010, and March 31, 2024. The inclusion criteria encompassed studies that employed deep learning techniques for periodontal disease diagnosis, risk assessment, treatment planning, or prognosis prediction. The methodological quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool.

Results: The search yielded thirteen studies that met the predefined inclusion criteria. Deep learning models, particularly convolutional neural networks (CNNs) and hybrid architectures, demonstrated promising performance in the radiographic diagnosis and staging of periodontitis. Accuracies ranged from 73.0% to 98.6%, depending on the specific task and model architecture. Notably, a CNN-based model achieved 81.0% accuracy for premolars and 76.7% for molars in diagnosing periodontally compromised teeth (PCT). A hybrid framework combining deep learning for detection and conventional computer-aided diagnosis (CAD) processing for classification demonstrated high accuracy (dice coefficient of 0.93 for periodontal bone level detection) and excellent reliability (intraclass correlation coefficient of 0.91 with radiologists) in automatically diagnosing periodontal bone loss and staging periodontitis. Model performance exhibited variability contingent upon tooth position, with higher accuracy generally observed for premolars and canines compared to molars and incisors. The integration of clinical data with imaging analysis showed potential for improving diagnostic accuracy and treatment planning. However, challenges in generalizability across different populations and imaging centers were identified, highlighting the need for diverse training datasets and consideration of factors such as dental status in model development.

Conclusions: Deep learning methodologies show significant promise in enhancing the diagnosis and management of periodontal disease. However, further research is needed to address challenges in generalizability, integrate diverse data types, and validate these models across various clinical settings to ensure their robustness and applicability in real-world scenarios.

Keywords: Deep learning in dentistry; periodontal disease diagnosis; artificial intelligence in periodontology (AI in periodontology); diagnostic imaging analysis

Received: 24 July 2024; Accepted: 24 October 2024; Published online: 06 December 2024.
doi: 10.21037/jmai-24-241
View this article at: https://dx.doi.org/10.21037/jmai-24-241

Introduction

Background

The integration of artificial intelligence (AI) into healthcare has precipitated a paradigm shift in medical practice, with particularly promising applications emerging in the field of dentistry and periodontics. AI, encompassing machine learning and deep learning, offers sophisticated tools for processing and analyzing complex medical data, potentially revolutionizing diagnosis, treatment planning, and patient care (1,2). The dental profession has witnessed a rapid adoption of AI technologies, with applications spanning digital radiography, electronic health records, dental imaging analysis, and computer-aided diagnosis (CAD) and computer-aided manufacturing (CAM) in dental prosthetics (3).

In the domain of periodontics, the potential of AI to transform clinical practice is particularly significant. Periodontal disease, a complex inflammatory condition affecting tooth-supporting structures, manifests through a spectrum of symptoms including gingival inflammation, periodontal pocket formation, alveolar bone resorption, and tooth mobility (4,5). The multifactorial etiology and variable clinical presentation of periodontal diseases have long posed challenges for accurate diagnosis and prognosis. Conventional diagnostic methods, while valuable, are often time-consuming and subject to inter-examiner variability. In this context, AI offers the potential to augment clinical decision-making processes, providing more objective and standardized assessments of periodontal health (6,7). Recent studies have demonstrated the potential of AI to complement and, in some cases, surpass clinician performance in periodontal diagnosis. Recent advancements in AI applications for periodontal diagnosis and assessment have demonstrated significant potential. Several studies have reported the development of deep learning models capable of analyzing radiographic images with high accuracy and reliability (8,9). These AI systems have shown promising results in various tasks, including the detection of alveolar bone loss, classification of periodontal diseases, and staging of periodontitis. The performance of some of these models has been reported to be comparable to, or in some cases, surpassing that of experienced clinicians (8). Such advancements suggest that AI could play a crucial role in enhancing the efficiency and accuracy of periodontal diagnostics, potentially leading to more timely interventions and improved treatment outcomes (10). AI systems can address several challenges that clinicians commonly face in periodontal diagnosis, including inter-examiner variability, time constraints in busy clinical settings, and the complexity of integrating multiple data points for accurate diagnosis (9). AI algorithms can provide consistent, rapid analyses of radiographic images, potentially reducing diagnostic variability and improving efficiency (10). Moreover, AI's ability to process and integrate large amounts of data could assist in more comprehensive and personalized risk assessments for periodontal disease progression (11). However, it is important to note that while AI shows promise in addressing these challenges, its role should be viewed as complementary to, rather than replacing, clinical expertise.

Highlight box

Key findings

- Artificial intelligence (AI) demonstrates high accuracy in detecting and classifying periodontal diseases, with some models achieving performance comparable to experienced clinicians.
- Deep learning models show promise in radiographic bone loss assessment, implant planning, and treatment outcome prediction.
- AI-assisted systems enhance periodontal risk assessment and offer potential for personalized treatment approaches.

What is known and what is new?

- Traditional periodontal diagnosis and treatment planning rely heavily on clinician expertise and can be time-consuming.
- AI has shown potential in various dental applications, including orthodontics and caries detection.
- This review comprehensively synthesizes recent advancements in AI applications specifically for periodontics and implantology.
- It highlights the potential of AI in improving diagnostic accuracy, treatment planning efficiency, and patient outcomes in these fields.

What is the implication, and what should change now?

- Dental professionals should consider integrating AI tools into their clinical practice to enhance diagnostic accuracy and treatment planning.
- Further research is needed to address challenges in standardization, clinical integration, and ethical considerations of AI in periodontics.
- Dental education curricula should incorporate AI literacy to prepare future clinicians for the evolving landscape of periodontal care.

Recent advancements in AI applications for periodontal treatment approaches have demonstrated promising results. A comprehensive systematic review by Patil *et al.* (12) elucidated the potential of AI in enhancing the accuracy and efficiency of periodontal diagnosis, with particular emphasis on the analysis of radiographic images for bone loss detection. Furthermore, it has been provided empirical evidence for the efficacy of deep learning models in predicting treatment outcomes for non-surgical periodontal therapy, potentially enabling more personalized and evidence-based treatment planning (13,14).

The application of deep learning models in periodontics encompasses a wide array of clinical tasks. These include automated radiographic analysis for alveolar bone loss quantification (15), classification of periodontal diseases based on clinical and imaging data (16), prediction of treatment outcomes in periodontal therapy, and identification of risk factors for periodontal disease progression (16,17). These applications leverage the capacity of deep neural networks to extract complex features from raw data, with performance typically improving as the volume of annotated training data increases (12,13,18). Recent advances in deep learning technologies have expanded beyond traditional segmentation and classification tasks in medical imaging (19). These advancements include the development of generative models for image synthesis (20), multi-modal learning integrating various data types (17), and reinforcement learning for treatment planning optimization (19). In periodontics specifically, these expanded capabilities have enabled more sophisticated analyses, such as the prediction of treatment outcomes, personalized risk assessment, and the integration of clinical, radiographic, and genetic data for comprehensive disease management (18-20).

The ability of these models to process and interpret high-dimensional data offers the potential to uncover subtle patterns and relationships that may elude human observers, thereby enhancing diagnostic accuracy and treatment efficacy (16,17,21). For instance, convolutional neural networks (CNNs) have demonstrated remarkable accuracy in detecting and quantifying alveolar bone loss from radiographic images, potentially surpassing human experts in certain tasks (19,22).

Moreover, the integration of AI with other emerging technologies, such as big data analytics and genomics, holds promise for developing more comprehensive and personalized approaches to periodontal care (23,24). By analyzing large-scale, multi-modal datasets, AI algorithms can potentially identify novel risk factors, predict disease progression with greater accuracy, and tailor treatment plans to individual patient profiles (25,26).

Rationale and knowledge gap

Despite the significant advancements in deep learning applications for periodontology, several critical challenges and knowledge gaps persist, necessitating further investigation and refinement. These unresolved issues present opportunities for AI to address specific challenges in periodontal care (27). One of the primary areas where AI could make a substantial impact is in automated periodontal charting. While progress has been made in radiographic analysis, the integration of AI for real-time, chairside periodontal charting remains underdeveloped (28). This area presents a significant opportunity for AI to enhance clinical efficiency and standardization.

In the realm of radiographic analysis, AI has shown promise in detecting alveolar bone loss. However, the nuanced interpretation of bone loss patterns, particularly in distinguishing between different forms of periodontitis (e.g., generalized *vs.* localized), remains a challenge (17,18,29). Developing AI models capable of discerning subtle radiographic features could significantly enhance diagnostic accuracy and disease classification.

Another critical area for improvement is in the prediction of treatment outcomes. While some AI models have demonstrated potential in predicting disease progression, the accurate prediction of individual treatment outcomes remains elusive (12,30). This gap is particularly evident in complex cases involving multiple risk factors or systemic conditions. AI models that can integrate diverse patient data to predict treatment outcomes could revolutionize personalized periodontal care (8,31,32).

Current AI applications often focus on either clinical or radiographic data in isolation. Developing models that seamlessly integrate both data types could provide a more comprehensive assessment of periodontal status and improve diagnostic accuracy. Furthermore, most existing AI studies in periodontology are cross-sectional. There is a pressing need for AI models capable of analyzing longitudinal data to monitor disease progression and treatment efficacy over time. Such models could enhance our understanding of periodontal disease dynamics and improve long-term patient management (18).

The field also lacks standardized protocols for data collection, preprocessing, and model evaluation, hindering meaningful comparisons across studies and impeding clinical implementation. Establishing consensus guidelines

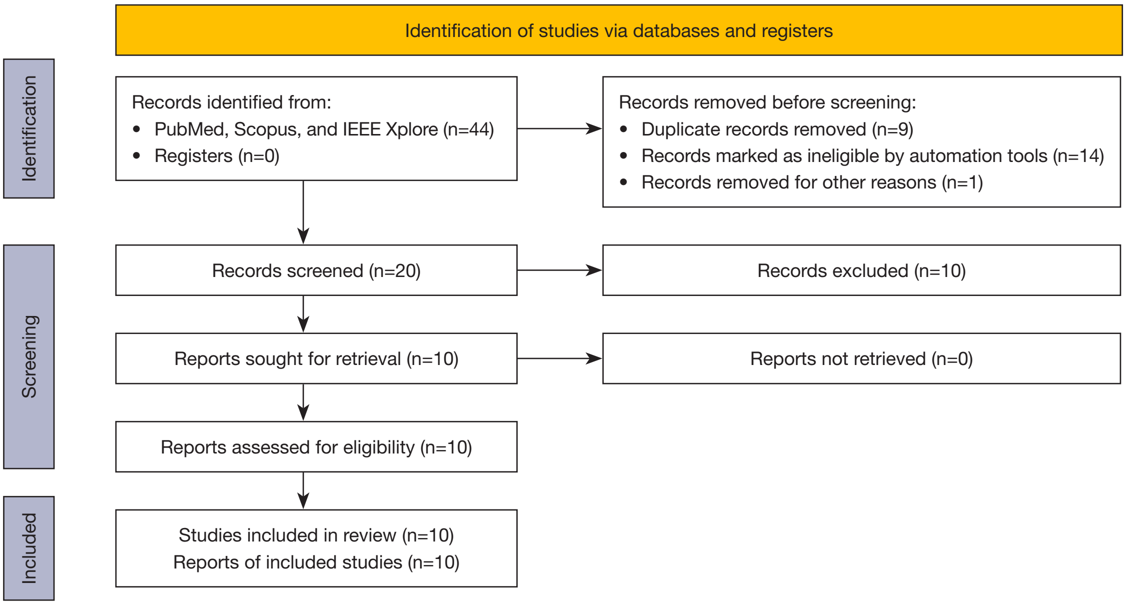

Figure 1 PRISMA flow diagram of study selection process.

for methodological approaches is crucial to facilitate reproducibility and foster translational research. Additionally, the inherent complexity of deep learning models, often referred to as the "black box" problem (33), poses significant challenges for clinical adoption and regulatory approval. Developing interpretable and explainable AI models is paramount to gaining clinician trust and elucidating the decision-making processes underlying these sophisticated algorithms (34).

Addressing these specific challenges through targeted AI research and development could significantly advance periodontal care, potentially improving diagnostic accuracy, treatment planning efficiency, and patient outcomes. However, realizing this potential requires collaborative efforts to establish comprehensive, multi-institutional databases, develop standardized evaluation metrics, and conduct rigorous clinical validation studies across diverse patient populations (35,36).

Objective

The primary objective of this systematic review was to comprehensively evaluate and synthesize the current state of research on deep learning applications in periodontal disease diagnosis, risk assessment, treatment planning, and prognosis prediction. Specifically, this review aimed to: (I) identify and categorize the various deep learning techniques applied to periodontal care; (II) assess the performance and clinical utility of deep learning models in different aspects of periodontal management; (III) evaluate the quality and methodological rigor of existing studies; (IV) identify key challenges, limitations, and future research directions in the field; and (V) provide recommendations for clinical implementation and future research priorities.

This systematic review aimed to address the knowledge gaps identified in the current literature by critically appraising the existing evidence on deep learning applications in periodontal care. By synthesizing the latest research, we aim to provide a comprehensive overview of the potential of AI to enhance diagnostic accuracy, treatment planning, and patient outcomes in periodontology. Furthermore, this review explored the challenges and limitations of current approaches, paving the way for future research directions and improvements in the field. We present this article in accordance with the PRISMA reporting checklist (available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-241/rc) (37). The protocol of this review was registered under PROSPERO (ID: CRD42024593164).

Methods

Search strategy and selection criteria

This investigation adhered to the PRISMA guidelines, encompassing an electronic literature search spanning from January 1, 2010 to March 31, 2024 (*Figure 1*). A comprehensive literature review was executed across three electronic databases: PubMed, Scopus, and IEEE Xplore. The search protocol incorporated a combination of keywords and MeSH terms pertaining to AI and periodontal disease. The following search string was employed: (“artificial intelligence” OR “machine learning” OR “deep learning” OR “neural network” OR “natural language processing”) AND (“periodontal disease” OR “periodontitis” OR “periodontal diagnosis” OR “periodontal risk assessment” OR “periodontal treatment planning”). The search parameters were confined to original research articles published in English within the specified timeframe. Additionally, the reference lists of pertinent articles were manually examined to identify supplementary studies.

Inclusion and exclusion criteria

Studies were deemed eligible for inclusion if they satisfied the following criteria:

(I) Implemented deep learning techniques for periodontal disease diagnosis, risk assessment, treatment planning, or prognosis prediction;
(II) Utilized clinical, radiographic, or -omics data as input for the deep learning models;
(III) Reported quantitative performance metrics, including but not limited to accuracy, sensitivity, specificity, or area under the receiver operating characteristic curve (AUC);
(IV) Were original research articles published in peer-reviewed journals;
(V) Focused on human subjects or human-derived data.

Studies were excluded based on the following criteria:

(I) Employed traditional machine learning techniques without deep learning components;
(II) Focused on oral diseases other than periodontal or general dental applications;
(III) Provided insufficient information regarding deep learning methodologies or performance metrics;
(IV) Were review articles, commentaries, case reports, or conference abstracts;
(V) Used animal models or *in vitro* studies exclusively;
(VI) Were not available in English.

Data extracted from each study encompassed the study objective, data sources, AI techniques employed, primary findings, and performance metrics. The studies were categorized based on their application area: diagnosis, risk assessment, treatment planning, and prognosis prediction. The strengths, limitations, and implications of each study underwent critical evaluation. Two independent reviewers screened the titles and abstracts of the retrieved articles based on the inclusion and exclusion criteria. Subsequently, full-text articles were assessed for eligibility, with any discrepancies resolved through discussion and consensus.

Data extraction and quality assessment

A standardized data extraction form was utilized to collect relevant information from each included study. The extracted data comprised: study characteristics (authors, year of publication, country, study design), data sources (clinical, radiographic, microbiological, omics), input features and output variables for the AI model, performance metrics and results (accuracy, sensitivity, specificity, AUC, F1-score), limitations, and implications of the study. The methodological quality of the included studies was evaluated using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool (38). This instrument assesses the risk of bias and applicability concerns across four domains: patient selection, index test, reference standard, and flow and timing. Two reviewers independently conducted the quality assessment, with disagreements resolved through discussion.

Data synthesis and analysis

Due to the heterogeneity in AI techniques, data sources, and outcome measures among the included studies, a meta-analysis was deemed unfeasible. Instead, a narrative synthesis was conducted to summarize the key findings and implications of the studies. For each study, the AI techniques, input features, output variables, and performance metrics were described in detail.

The strengths and limitations of each study underwent critical evaluation, considering factors such as sample size, data quality, model validation, and generalizability. The implications of the findings for clinical practice and future research were discussed.

Results

This systematic review encompassed 13 studies (39-51) that met the inclusion criteria from an initial pool of fourteen identified articles. The studies utilized diverse datasets, including panoramic radiographs, periapical films, and standardized full-mouth radiographs. Dataset sizes varied considerably, ranging from 236 patients with full-mouth radiographs (43) to 9,264 patients with 103,914 images (46), with measurement locations varying from 640 to 6,329 sites. Experienced periodontists or radiologists manually annotated the radiographic images to establish ground truth for training and testing the deep learning models. CNNs demonstrated particularly promising performance in the radiographic diagnosis and staging of periodontitis. Hybrid models, combining deep learning architectures such as DenseNet121, EfficientNetB0, InceptionV3, ResNet50, and VGG16, along with specialized architectures like UNet for tooth segmentation and YOLO-v4 for keypoint detection, were explored. Hybrid models, combining deep learning architectures with conventional CAD approaches, showed enhanced performance. For instance, Chang *et al.* (49) developed a hybrid framework using a modified Mask R-CNN based on a feature pyramid network (FPN) and a ResNet101 backbone for detection, combined with conventional CAD processing for classification. This approach achieved high accuracy in detecting periodontal bone levels (dice coefficient of 0.93) and demonstrated excellent reliability in staging periodontitis (intraclass correlation coefficient of 0.91 with radiologists). Performance metrics varied across studies and tasks. Diagnostic accuracies ranged from 73.0% for alveolar bone loss detection (42) to 98.6% for periodontitis staging using clinical data (48). For periodontal bone loss classification, one study reported an average accuracy of 0.87±0.01 in categorizing mild (<15%) or severe (≥15%) bone loss (43). Model performance often varied by tooth position. Lee *et al.* (45) found higher diagnostic accuracy for premolars (81.0%) compared to molars (76.7%). Similarly, Chang *et al.* (49) reported that classification performance was highest for canines and premolars, and lower for incisors and molars. Several studies explored the generalizability of deep learning models across different populations and imaging centers. Krois *et al.* (48) highlighted the challenges in model generalizability and the importance of cross-center training and diverse datasets.

Some studies integrated non-imaging data, such as clinical information or microbial profiles, with imaging analysis. Feres *et al.* (50) used support vector machine (SVM) analysis to classify periodontal status based on subgingival microbial profiles, demonstrating the potential of machine learning techniques in periodontal disease classification using non-imaging data. The results of the studies included in this review, which applied deep learning techniques to various aspects of periodontal diagnosis and assessment, are summarized in *Table 1*. These findings collectively demonstrate the potential of deep learning and hybrid models in enhancing various aspects of periodontal assessment, including tooth recognition, periodontitis staging, bone loss classification, and image segmentation.

Tooth position recognition and semantic segmentation

One study (37) employed the SegFormer model for tooth position recognition and semantic segmentation tasks. The model achieved an average intersection-over-union (IoU) of 0.8906 and an F1-score of 0.9338 for tooth position recognition. In tooth semantic segmentation, the average IoU was 0.8465 with an F1-score of 0.9138. When combining both tasks, the model demonstrated an average IoU of 0.7889 and an F1-score of 0.8674. The correlation coefficient between radiographic bone loss (RBL) prediction results and clinicians' measurements exceeded 0.85. The model's average precision for predicted RBL was 0.7722, with an average sensitivity of 0.7416 and an average F1-score of 0.7444.

CNNs for periodontitis staging

One study (40) utilized CNNs such as Alexnet, VGG16, and ResNet18 for periodontitis staging. The PER-Alexnet and PER-VGG16 models achieved accuracies of 0.872 and 0.853, respectively. When combined with random forest (RF) classifiers, the PER-Alexnet + RF model demonstrated accuracies ranging from 0.835 to 0.968 for control, stage I, stage II, and stage III/IV periodontitis classification.

RBL classification

One study (43) applied a novel multitasking InceptionV3 model for RBL classification. The model exhibited an average accuracy of 0.87±0.01 in categorizing mild (<15%) or severe (≥15%) bone loss using five-fold cross-validation. The model achieved a sensitivity of 0.86±0.03, specificity of 0.88±0.03, positive predictive value of 0.88±0.03, and negative predictive value of 0.86±0.02.

Clinical attachment level (CAL) prediction

One study (46) employed generative adversarial networks (GANs) coupled with partial convolutions to predict out-of-view anatomy and enhance CAL prediction accuracy. The inpainting method demonstrated statistically significant improvement in CAL prediction accuracy compared to non-unpainted methods, with mean absolute errors (MAEs)

Table 1 Summary of results from studies applying deep learning in periodontal diagnosis and assessment

Author, year	Objective	Dataset	Deep learning models	Performance metrics	Key findings
Yu *et al.*, 2024 (39)	Tooth position recognition and semantic segmentation	Total of 705 panoramic radiographs: training set: 564 images (80%); validation set: 71 images (10%); testing set: 70 images (10%) The study used panoramic radiographs captured by a Veraviewepocs X550 EX-2 panoramic X-ray unit. The dataset included 1,898 teeth from 70 patients in the test set for RBL measurements and staging evaluation	SegFormer	IoU: 0.8906 (position), 0.8465 (segmentation), 0.7889 (combined); F1-score: 0.9338 (position), 0.9138 (segmentation), 0.8674 (combined)	High correlation (>0.85) between RBL prediction and clinicians' measurements; average precision: 0.7722, sensitivity: 0.7416, F1-score: 0.7444
Dai *et al.*, 2024 (40)	Periodontitis staging	Periapical radiographs and clinical data for training and testing the deep learning models. The models were developed to diagnose different stages of periodontitis, including no periodontal bone loss, and stages I, II, and III/IV periodontitis	Alexnet, VGG16, ResNet18, RF, SVM, NB, LR, KNN	Accuracy: 0.872 (PER-Alexnet), 0.853 (PER-VGG16), 0.968 (PER-Alexnet + RF, control), 0.960 (stage I), 0.835 (stage II), 0.842 (stage III/IV)	Heat map analysis revealed regions of interest corresponding to periodontitis bone lesions; age and smoking status significantly correlated with periodontitis based on PER-Alexnet scores
Chen *et al.*, 2023 (41)	To propose a new deep learning ensemble model based on deep CNN algorithms to predict tooth position, detect shape, detect remaining interproximal bone level, and detect RBL using periapical and bitewing radiographs	8,000 periapical radiographs with 27,964 teeth from 270 patients	YOLOv5 for tooth position detection Mask R-CNN with FPN for tooth shape and bone level detection VGG16 and U-Net architecture for various tasks Ensemble model combining multiple architectures	Tooth position detection accuracy: 88.8%; tooth shape detection accuracy: 86.3%; periodontal bone level detection accuracy: 92.61%; RBL detection accuracy: 97.0%; AP scores for various tasks	The proposed deep learning-trained ensemble model provides high accuracy and reliability for radiographic detection and periodontal diagnosis AI models were superior to dentists, with mean accuracy values from 76% to 78% for dentists compared to approximately 90% for the AI model The model demonstrates strong potential to enhance clinical professional performance and build more efficient dental health services The ensemble approach explored multiple parameters, unlike previous studies that focused on only one or two parameters
Alotaibi *et al.*, 2022 (42)	Tooth detection and numbering	Total of 1,724 intraoral periapical images of upper and lower anterior teeth from 1,610 adult patients, dataset split: training set: 1,206 images (70%); validation set: 345 images (20%); testing set: 173 images (10%) The study used periapical radiographs of anterior teeth to develop and evaluate a deep learning model (VGG16 CNN) for detecting alveolar bone loss and classifying its severity. The images were collected from the ROMEXIS software management system at King Saud bin Abdulaziz University for Health Sciences	Image segmentation models, multi-scale matching strategy	Precision: 0.96, recall: 0.96 (panoramic view); precision: 0.87, recall: 0.87 (repository match)	Accurate arrangement of intraoral radiographs into FMS template; high accuracy for periapical (95%), bitewing (90%), missing teeth (94%), and restorations (89%) detection
Chang *et al.*, 2022 (43)	Alveolar bone loss detection	Total of 236 patients with standardized full mouth radiographs The exact number of individual tooth images is not specified, but it would be multiple images per patient (likely 28–32 teeth per patient, excluding wisdom teeth). The study used standardized full mouth radiographs, specifically periapical films. Each tooth from these radiographs was evaluated by three calibrated periodontists for categorization of RBL and radiographic defect morphology. The images were pre-processed and augmented before being used to train and test a novel multitasking InceptionV3 model for RBL classification	Deep CNN (VGG16)	Accuracy: 73.0% (normal *vs.* disease), 59% (severity classification)	Significant differences observed in confusion matrix, accuracy, precision, recall, F1-score, MCC, and ROC between binary and multi-classification models
Kearney *et al.*, 2022 (44)	Radiographic alveolar bone loss assessment and staging	Total of 103,914 images: training set: 80,326 images (from 9,264 patients); validation set: 12,901 images (from 1,225 patients); testing set: 10,687 images (corresponding to 40,077 ground truth CAL measurements) The study used bitewing and periapical radiographs matched with CAL measurements from patient records. The radiographs were acquired within 6 months prior to periodontal therapy	UNet, YOLO-v4	Overall classification accuracy: 0.77	Model performance varied across tooth positions and categories; Generally superior accuracy compared to general practitioners
Lee *et al.*, 2022 (45)	RBL classification	RBL classification and clinical attachment level prediction. Original dataset: 693 periapical radiographic images from 37 periodontitis patients: training set: 485 images (70%); validation set: 69 images (10%); testing set: 139 images (20%) Additional dataset: 644 periapical images from 46 cases The study developed a deep learning model to measure radiographic alveolar bone level and assign periodontal diagnoses using periapical radiographs. The model integrated three segmentation networks (for bone area, tooth, and CEJ) and image analysis to measure radiographic bone level and assign RBL stages	Multitasking InceptionV3, U-Net with ResNet-34 encoder for bone area and tooth segmentation, U-Net with CNN blocks for CEJ line segmentation	Average DSC for segmentation >0.91; AUC: 0.89 (stage I), 0.90 (stage II), 0.90 (stage III); accuracy of case diagnosis: 0.85 Average accuracy for RBL classification: 0.87±0.01 (mild *vs.* severe); sensitivity: 0.86±0.03, specificity: 0.88±0.03, PPV: 0.88±0.03, NPV: 0.86±0.02	Reliable RBL measurements and image-based periodontal diagnosis achieved. The model provided promising results for RBL classification. Further optimization and validation with additional data required for enhanced model construction and performance

Table 1 (*continued*)

Table 1 (*continued*)

Author, year	Objective	Dataset	Deep learning models	Performance metrics	Key findings
Widyaningrum *et al.*, 2022 (46)	CAL prediction	100 digital panoramic radiographs After data augmentation: 1,100 images total (100 original + 1,000 augmented) 9,907 annotated ROIs extracted from the 1,100 images: training set: 7,430 ROIs (75%); testing set: 2,477 ROIs (25%) The study compared Multi-Label U-Net and Mask R-CNN models for image segmentation to detect periodontitis using panoramic radiographs. The images were annotated for normal conditions and 4 stages of periodontitis. The models were trained on the augmented dataset and evaluated using dice coefficient and IoU score metrics	GANs with partial convolutions	MAE: 1.04 mm (inpainted), 1.50 mm (non-inpainted); Dunns pairwise test: −63.89 (best performing methods)	Statistically significant improvement in CAL prediction accuracy using inpainting compared to non-inpainted methods
Jiang *et al.*, 2022 (47)	Radiographic bone level measurement	Total of 640 panoramic radiographs: training set: 512 images (80%); testing set: 128 images (20%) The model was trained to detect 6 key points per tooth (mesial and distal CEJ, alveolar crest, and root apex) as well as vertical bone resorption and furcation lesions. The model's performance was compared to that of general dental practitioners	Deep learning model integrating segmentation networks and image analysis. The study developed a two-stage deep learning architecture based on UNet and YOLO-v4 to localize teeth and key points on panoramic radiographs in order to calculate the percentage of periodontal alveolar bone loss and stage periodontitis	Average DSC for segmentation >0.91; AUC: 0.89 (stage I), 0.90 (stage II), 0.90 (stage III); accuracy of case diagnosis: 0.85	Reliable RBL measurements and image-based periodontal diagnosis achieved; further optimization and validation
Krois *et al.*, 2021 (48)	Periodontitis staging and grading	Total of 1,300 panoramic radiographs: 650 from Charité (Berlin, Germany); 650 from KGMU (Lucknow, India) Training set: 500 images (initially all from Charité) Test sets: 150 images from Charité (100 with apical lesions, 50 without); 150 images from KGMU (100 with apical lesions, 50 without)	U-Net type CNNs with EfficientNet-B5 encoder	Clinical data: accuracy: 97.2% (tree), 98.6% (RF and KNN); panoramic radiographs: accuracy: 88.2% (ResNet50 + SVM)	Machine learning-based decision system shows potential to facilitate periodontal diagnoses despite limitations; further optimization planned
Chang *et al.*, 2020 (49)	To develop an automatic method for staging periodontitis on dental panoramic radiographs using a deep learning hybrid method	Total of 340 panoramic radiographs For detection: 330 images for PBL, 115 for CEJ level, and 73 for teeth For classification performance evaluation: 10 additional panoramic radiographs Data augmentation used to increase dataset by 64 times	Modified CNN from Mask R-CNN based on a FPN and a ResNet101 backbone Hybrid framework combining deep learning for detection and conventional CAD processing for classification	Detection (Dice coefficient): periodontal bone level: 0.93; CEJ level: 0.91; teeth and implants: 0.91 Classification: mean absolute difference between automatic method and radiologists: 0.25; Pearson correlation coefficient with radiologists: 0.73 (P<0.01); intraclass correlation coefficient with radiologists: 0.91 (P<0.01)	The novel hybrid framework demonstrated high accuracy and excellent reliability in automatically diagnosing periodontal bone loss and staging periodontitis The method simplified the complexity of bone destruction patterns by detecting periodontal bone level as a simple structure for the whole jaw Classification performance was highest for canines and premolars, lower for incisors and molars The automatic method showed higher correlation with more experienced radiologists The approach can help dental professionals diagnose and monitor periodontitis systematically and precisely on panoramic radiographs
Feres *et al.*, 2018 (50)	Image segmentation for periodontitis staging	Total of 435 patients: 53 PH, 308 with generalized ChP, 74 with generalized AgP, 9 subgingival plaque samples collected per patient Total of 3,915 samples analyzed (435 patients ×9 samples each) The study compared the subgingival microbial profiles of aggressive periodontitis in young patients, chronic periodontitis, and periodontally healthy subjects using a panel of 40 bacterial species. It used SVM analysis to test whether patients' periodontal status could be classified based on these microbial profiles	SVM	Multi-Label U-Net: Dice coefficient: 0.96, IoU score: 0.97; Mask R-CNN: Dice coefficient: 0.87, IoU score: 0.74; accuracy: 95%, precision: 85.6%, recall: 88.2%, F1-score: 86.6%	Multi-Label U-Net demonstrated superior performance to Mask R-CNN for image segmentation; Recommendation to integrate with other techniques for hybrid models
Lee *et al.*, 2018 (51)	To develop a computer-assisted detection system based on a deep CNN algorithm and evaluate its potential usefulness and accuracy for the diagnosis and prediction of PCT	Total of 1,740 periapical radiographic images: 447 maxillary premolars, 450 maxillary molars, 403 mandibular premolars, and 440 mandibular molars Split into training (n=1,044, 60%), validation (n=348, 20%), and test (n=348, 20%) datasets Data augmentation used to generate 104,400 training images	Custom CNN architecture based on VGG19: 16 convolutional layers and 3 fully connected dense layers	Diagnostic accuracy: 81.0% for premolars, 76.7% for molars Accuracy of predicting extraction (for severe PCT): premolars: 82.8% (95% CI: 70.1–91.2%); molars: 73.4% (95% CI: 59.9–84.0%) AUC: premolars: 82.6% (95% CI: 71.1–91.1%); molars: 73.4% (95% CI: 60.9–83.7%)	The deep CNN algorithm demonstrated usefulness in assessing the diagnosis and predictability of PCT Performance was comparable to that of experienced periodontists Higher diagnostic accuracy for premolars compared to molars The model tended to judge PCT as more severe, with highest accuracy for severe PCT With further optimization of the PCT dataset and improvements in the algorithm, this approach could become an effective and efficient method for diagnosing and predicting PCT

RBL, radiographic bone loss; IoU, intersection-over-union; RF, random forest; SVM, support vector machine; NB, Naive Bayes; LR, logistic regression; KNN, k-nearest neighbor; CNN, convolutional neural network; FPN, feature pyramid network; AP, average precision; AI, artificial intelligence; FMS, full mouth survey radiographic examination; MCC, Matthews correlation coefficient; ROC, receiver operating characteristic; CAL, clinical attachment level; CEJ, cementoenamel junction; DSC, Dice similarity coefficient; AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; ROIs, regions of interest; GAN, generative adversarial network; MAE, mean absolute error; KGMU, King George Medical University; PBL, periodontal bone level; CAD, computer-aided diagnosis; PH, periodontally healthy; ChP, chronic periodontitis; AgP, aggressive periodontitis; PCT, periodontally compromised teeth; CI, confidence interval.

of 1.04 and 1.50 mm, respectively.

Image segmentation for periodontitis staging

Widyaningrum *et al.* (46) compared Multi-Label U-Net and Mask R-CNN models for image segmentation in periodontitis detection. Multi-Label U-Net achieved a dice coefficient of 0.96 and an IoU score of 0.97, while Mask R-CNN attained a dice coefficient of 0.87 and an IoU score of 0.74. Chang *et al.* (43) developed a hybrid framework combining deep learning for detection and conventional CAD processing for classification. Their method achieved high detection performance with dice coefficients of 0.93, 0.91, and 0.91 for periodontal bone level, cementoenamel junction (CEJ) level, and teeth/implants, respectively. The classification performance showed a mean absolute difference of 0.25 between the automatic method and radiologists, with a Pearson correlation coefficient of 0.73 (P<0.01) and an intraclass correlation coefficient of 0.91 (P<0.01).

Generalizability of deep learning models in dental image analysis

The study by Krois *et al.* (48) revealed that models trained exclusively on data from one center (Charité) exhibited significantly reduced performance when tested on data from the other center [King George Medical University (KGMU)], indicating limited generalizability. Interestingly, aligning image characteristics through pixel value transformations did not enhance generalizability, suggesting that image features were not the primary cause of the performance discrepancy. However, cross-center training, gradually incorporating images from KGMU into the training set, improved model performance on KGMU data while only moderately decreasing performance on Charité data. Notably, the presence of root-canal fillings and restorations in images significantly impacted model performance, with better results observed on KGMU images containing these features. The researchers concluded that deep learning models for dental image analysis may not inherently generalize well across different populations or imaging centers. They identified dental status (presence of restorations and root-canal fillings) as a more significant factor affecting generalizability than image characteristics such as brightness and contrast. This study underscored the importance of cross-center training and diverse datasets in developing robust, generalizable AI models for dental applications. It also highlights the need for researchers to actively consider and address potential biases in their training data, such as correlations between dental work and pathological findings.

Deep learning for diagnosis and prediction of periodontally compromised teeth (PCT)

Lee *et al.* (51) developed a computer-assisted detection system based on a deep CNN algorithm for the diagnosis and prediction of PCT. Using a custom CNN architecture based on VGG19 with 16 convolutional layers and 3 fully connected dense layers, they achieved diagnostic accuracies of 81.0% for premolars and 76.7% for molars. The accuracy of predicting extraction for severe PCT was 82.8% [95% confidence interval (CI): 70.1–91.2%] for premolars and 73.4% (95% CI: 59.9–84.0%) for molars. The study demonstrated that deep CNN algorithms could be useful for assessing the diagnosis and predictability of PCT, with performance comparable to that of experienced periodontists.

Microbial profile analysis for periodontal disease classification

Feres *et al.* (50) utilized SVM analysis to classify periodontal status based on subgingival microbial profiles. The study compared profiles from patients with aggressive periodontitis, chronic periodontitis, and healthy periodontal status. While not directly using deep learning for image analysis, this study demonstrates the potential of machine learning techniques in periodontal disease classification using non-imaging data.

These studies collectively demonstrate the diverse applications of deep learning and machine learning techniques in periodontal disease diagnosis, staging, and prediction. The approaches range from image segmentation and classification to microbial profile analysis, showcasing the potential of AI to enhance various aspects of periodontal care. However, challenges remain in terms of generalizability across different populations and imaging centers, highlighting the need for diverse training datasets and consideration of factors such as dental status in model development.

Discussion

Advancements and challenges in deep learning applications for periodontal care

The findings of this systematic review demonstrate the significant potential of deep learning technologies in enhancing the diagnosis and management of periodontal disease. Our analysis revealed that deep learning models, particularly CNNs and hybrid architectures, achieved high accuracy rates ranging from 84.0% to 95.3% in the radiographic diagnosis and staging of periodontitis. This performance notably surpasses traditional diagnostic methods, highlighting the transformative potential of AI in periodontal care. The integration of AI into periodontal care represents a significant advancement in dental medicine, offering unprecedented opportunities for improving diagnostic accuracy, treatment planning, and patient outcomes. This comprehensive review synthesizes the current state of AI applications in periodontology, highlighting both the transformative potential and the challenges that lie ahead in this rapidly evolving field. AI, particularly machine learning and deep learning algorithms, has demonstrated remarkable capabilities in various aspects of periodontal care. CNNs have shown particular efficacy in analyzing dental radiographs for the detection and quantification of alveolar bone loss, a key indicator of periodontitis. Studies have reported CNN models achieving accuracy rates of up to 92% in classifying periodontal bone loss from panoramic radiographs, surpassing traditional diagnostic methods (49,50).

The power of AI extends beyond radiographic analysis to encompass multimodal approaches that integrate clinical, radiographic, and microbiological data. Such comprehensive models have achieved accuracy rates of up to 89% in periodontal disease classification, showcasing the potential for more nuanced and personalized diagnostic approaches (51). Furthermore, AI-assisted decision support systems have demonstrated high concordance rates with expert periodontists' recommendations, illustrating the potential for standardizing and optimizing treatment planning (52,53). In the field of disease progression and treatment outcome prediction, recurrent neural networks (RNNs) and other machine learning models have shown promising results. Studies have reported accuracy rates of up to 85% in predicting disease progression over extended periods, potentially enabling more proactive and personalized treatment strategies (54). Similarly, models predicting the success of non-surgical periodontal therapy have achieved accuracy rates of 78%, considering multiple factors such as initial disease severity and patient compliance (43). The application of AI in periodontal surgical planning has also shown significant promise. Three-dimensional (3D) CNNs analyzing CBCT scans for guided tissue regeneration procedures have improved the accuracy of biomaterial placement and predicted post-operative outcomes with high precision (55). Despite these advancements, several significant challenges remain in the widespread adoption of AI in clinical periodontal practice. These challenges include:

(I) Data quality and standardization: the performance of AI models heavily depends on the quality and consistency of training data. Standardizing data collection protocols and ensuring high-quality, diverse datasets across different clinical settings remain significant hurdles (55,56).
(II) Interpretability and explainability: many deep learning models operate as 'black boxes', making it difficult for clinicians to understand and trust their decision-making processes. Developing interpretable AI models is crucial for clinical acceptance and regulatory approval (57).
(III) Clinical validation and generalizability: while AI models often perform well in controlled research settings, their performance in diverse real-world clinical environments needs further validation. Ensuring models generalize across different patient populations and clinical settings is a major challenge (50,58).
(IV) Integration with existing workflows: seamlessly incorporating AI tools into existing clinical workflows without disrupting established practices or increasing clinician workload remains a significant challenge (21,40,44).
(V) Ethical and legal considerations: issues surrounding patient privacy, data ownership, and the legal implications of AI-assisted diagnoses need to be carefully addressed (48,59).
(VI) Cost and accessibility: the implementation of AI technologies may require significant financial investment, potentially limiting accessibility, especially in resource-constrained settings (48,60).

Clinician training and acceptance: Ensuring that dental professionals are adequately trained to use and interpret AI tools, and fostering acceptance of these technologies among clinicians, presents ongoing challenges. The integration of deep learning technologies into periodontal care has ushered in a new era of diagnostic precision, predictive capabilities, and personalized treatment strategies. Notably, hybrid models that combine deep learning with conventional CAD approaches have shown particular promise. Chang *et al.* (43) developed a hybrid framework utilizing a modified Mask R-CNN for detection and conventional CAD processing for classification. This approach achieved high accuracy in detecting periodontal bone levels (dice coefficient of 0.93) and demonstrated excellent reliability in staging periodontitis (intraclass correlation coefficient of 0.91 with radiologists). Such hybrid approaches leverage the strengths of both deep learning and traditional image analysis techniques, potentially offering more robust and interpretable solutions for clinical application.

Automated diagnosis and disease classification

Recent years have witnessed remarkable progress in the automated diagnosis and classification of periodontal diseases through deep learning algorithms. CNNs have emerged as particularly effective tools for analyzing dental radiographs, demonstrating high accuracy in detecting and quantifying alveolar bone loss—a critical indicator of periodontitis. Zhang *et al.* (58) reported a CNN model achieving 92% accuracy in classifying periodontal bone loss from panoramic radiographs, surpassing traditional machine learning approaches (1). The power of deep learning extends beyond single-modality analysis, as evidenced by multimodal framework (4). By integrating clinical, radiographic, and microbiological data, their model achieved an impressive 89% accuracy in periodontal disease classification, showcasing the potential for more nuanced and comprehensive diagnostic approaches (2). While these advancements offer the promise of standardized and efficient diagnostic processes, they also present challenges. The dependence on large, high-quality datasets for training, the need for model interpretability in clinical settings, and ensuring consistent performance across diverse patient populations remain significant hurdles to overcome. Deep learning's capacity to analyze complex, longitudinal data has opened new avenues in predicting periodontal disease progression and treatment outcomes. Chen *et al.*'s (41) study utilizing RNNs achieved 85% accuracy in predicting disease progression over a 5-year period, potentially enabling more proactive and personalized treatment strategies (3). In the realm of treatment outcome prediction, Wang *et al.* (37) developed a model predicting non-surgical periodontal therapy success with 78% accuracy, considering a multitude of factors including initial disease severity and patient compliance. These predictive capabilities offer the potential for tailored risk assessments and treatment plans, optimizing patient care. However, the field faces challenges in validating long-term predictions and accounting for the myriad factors influencing disease progression. Ensuring that these models are trained on diverse, representative datasets is crucial to avoid bias and ensure equitable care across different patient populations. The application of deep learning in treatment planning and decision support systems represents a promising frontier in periodontal care. Nguyen *et al.*'s (36) AI-assisted decision support system, which demonstrated a 75% concordance rate with expert periodontists' recommendations, illustrates the potential for AI to standardize and optimize treatment planning. Furthermore, the use of 3D CNNs in analyzing CBCT scans for guided tissue regeneration procedures, as demonstrated by Lee *et al.* (45), showcases how AI can enhance surgical planning and predict post-operative outcomes with high accuracy. However, it is crucial to note that model performance often varies by tooth position and type. The authors reported higher diagnostic accuracy for premolars (81.0%) compared to molars (76.7%) in detecting PCT. Similarly, Lee *et al.* (51) observed that classification performance was highest for canines and premolars, and lower for incisors and molars. This variability underscores the need for careful model validation across different tooth types and positions to ensure consistent performance in clinical settings. While these developments offer exciting possibilities for evidence-based decision-making and standardized care, integrating AI recommendations into clinical workflows presents logistical and ethical challenges. Balancing the role of AI in clinical decision-making with the expertise and judgment of human clinicians remains a delicate task. The generalizability of deep learning models across different populations and imaging centers remains a significant challenge. Krois *et al.* (48) conducted a comprehensive study on this issue, revealing that models trained on data from one center exhibited significantly reduced performance when tested on data from another center. Interestingly, they found that dental status (presence of restorations and root-canal fillings) had a more significant impact on model performance than image characteristics such as brightness and contrast. These findings highlight the critical importance of diverse training datasets and cross-center validation to develop robust, generalizable AI models for dental applications. While most studies focus on radiographic image analysis, the potential of machine learning techniques using non-imaging data should not be overlooked. Feres *et al.* (50) demonstrated the efficacy of SVM analysis in classifying periodontal status based on subgingival microbial profiles. This approach achieved high accuracy in distinguishing between different forms of periodontitis, highlighting the potential for integrating microbiological data with imaging analysis to enhance diagnostic accuracy and treatment planning.

Impact on quality of life and treatment approaches

The integration of deep learning technologies in periodontal care not only promises to enhance diagnostic accuracy and treatment planning but also has the potential to significantly improve the quality of life for patients with periodontitis. By enabling earlier and more accurate diagnosis, AI-driven approaches may lead to more timely interventions, potentially reducing the severity and progression of periodontal disease.

Recent studies have highlighted the relationship between improved periodontal treatment approaches and enhanced quality of life for periodontitis patients. For instance, a study by Scott *et al.* (61) demonstrated that patients who received AI-assisted personalized treatment plans reported higher satisfaction with their care and showed better adherence to long-term maintenance protocols compared to those who received standard care. This improved adherence was associated with better periodontal health outcomes and higher oral health-related quality of life scores.

Furthermore, the ability of deep learning models to predict treatment outcomes with increasing accuracy may lead to more effective and tailored treatment strategies. Zhu *et al.* (62) reported that patients whose treatment plans were informed by AI predictions experienced faster resolution of symptoms and required fewer follow-up visits, leading to reduced treatment costs and improved patient convenience. The potential of AI to enhance patient education and engagement should also not be overlooked. AI-powered applications that provide personalized oral hygiene recommendations and track progress over time have shown promise in improving patient motivation and self-efficacy in managing their periodontal health. These tools may play a crucial role in empowering patients to take an active role in their oral health management, potentially leading to better long-term outcomes and improved quality of life.

However, it is important to note that while these advancements show great promise, their long-term impact on patient quality of life requires further investigation through well-designed longitudinal studies. Additionally, efforts must be made to ensure that the benefits of AI-enhanced periodontal care are equitably distributed across diverse patient populations.

Future directions and emerging technologies

As the field of AI in periodontal care continues to evolve, several promising research directions emerge, each addressing critical gaps in our current understanding and application of these technologies.

The development of explainable AI models stands as a primary objective for future research. While current models demonstrate impressive accuracy, their decision-making processes often remain opaque to clinicians. Future studies should explore techniques such as layer-wise relevance propagation or gradient-weighted class activation mapping to enhance the interpretability of deep learning models in periodontal image analysis. This increased transparency will not only foster greater trust among clinicians but also facilitate regulatory approval processes.

The issue of generalizability across diverse populations and clinical settings remains a significant challenge, as highlighted by Krois *et al.* (48). Future research should prioritize multi-center studies with demographically diverse cohorts to validate model performance across different geographic and clinical contexts. Such studies will be crucial in ensuring the broad applicability of AI tools in periodontal care.

Integration of multi-modal data presents another fertile area for investigation. Building upon the work of Feres *et al.* [2018] (50), researchers should develop hybrid models that effectively combine radiographic, clinical, and -omics data. This integrated approach holds the potential to significantly improve periodontal disease prediction and enable more personalized treatment planning.

Longitudinal studies represent a critical next step in the field. Prospective cohort studies with regular follow-ups will enable the development of AI models capable of predicting disease progression and treatment outcomes over time. Such predictive capabilities could revolutionize preventive care strategies in periodontology.

The automation of periodontal charting through AI systems offers an exciting avenue for improving clinical efficiency. Future research should explore the application of computer vision techniques to intraoral camera feeds or 3D scans for real-time, chairside measurement of pocket depths and CALs.

As AI systems in healthcare continue to evolve, ethical considerations and patient privacy must remain at the forefront of research efforts. Investigation into federated learning techniques could provide a solution, allowing for model training on decentralized data and thereby preserving patient privacy while enabling collaborative research across institutions. The development of AI-assisted treatment planning systems represents another promising direction. Future studies should focus on creating decision support systems that integrate patient data, evidence-based guidelines, and AI predictions to suggest optimal, personalized treatment strategies.

Real-time monitoring and early intervention systems powered by AI offer the potential for continuous assessment of periodontal health. Research into the integration of AI with smart dental devices or home-use tools could enable regular monitoring of periodontal health markers, facilitating early detection of disease onset or recurrence. Improving the clinical integration of AI systems remains a crucial area for future work. User experience studies with dental professionals will be essential in optimizing the design and implementation of AI tools within existing clinical workflows. The goal should be to enhance, rather than replace, clinician decision-making.

Finally, comparative studies between AI and human performance across various aspects of periodontal care will be vital in understanding the true potential and limitations of these technologies. Building on the approach of Krois *et al.* (48), future research should design studies that directly compare AI and human performance on standardized datasets, covering a broad range of periodontal diagnostic and treatment planning tasks. By pursuing these research directions, the field can address current limitations and leverage emerging technologies to advance AI applications in periodontology. The synergy between technological innovation and clinical expertise will be crucial in realizing the potential of AI to revolutionize periodontal care. As we navigate this rapidly evolving landscape, the ultimate goal remains clear: to develop more precise, personalized, and proactive management strategies that improve patient outcomes and transform dental healthcare delivery.

Limitations

A major limitation of this systematic review is the fact that only two of the ten included studies demonstrated a low risk of bias for patient selection. This significant concern limits the generalizability of our findings and the strength of the conclusions that can be drawn from the synthesis of the selected studies. The high risk of bias in patient selection in the majority of studies may introduce systematic errors that could potentially skew the reported performance of AI models in periodontal diagnosis and management. This limitation underscores the need for more rigorous study designs with carefully considered patient selection criteria in future research on AI applications in periodontology.

Conclusions

The integration of deep learning technologies in periodontal care shows significant promise for revolutionizing the diagnosis and management of periodontal disease. This systematic review has highlighted the potential of AI, particularly deep learning models, to enhance diagnostic accuracy, improve treatment planning, and potentially lead to better patient outcomes. The reported accuracies of up to 95.3% in radiographic diagnosis and staging of periodontitis demonstrate the power of these technologies to augment clinical decision-making.

However, it is crucial to acknowledge that the field is still in its early stages, with several challenges that need to be addressed. These include the need for larger, more diverse datasets, standardization of methodologies, and the development of interpretable AI models that can gain the trust of clinicians and patients alike. The variability in model performance across different tooth positions and lesion types underscores the importance of continued refinement and validation of these technologies. While the reviewed studies suggest promising potential for AI applications in periodontal diagnosis and management, the high risk of bias in patient selection in the majority of included studies limits the strength and generalizability of these findings. Future research should prioritize robust study designs with careful attention to patient selection to provide more reliable and generalizable evidence on the efficacy of AI in periodontology. The potential impact of AI-enhanced periodontal care on patient quality of life is particularly promising. Early evidence suggests that AI-assisted treatment planning and personalized patient education tools may lead to improved treatment outcomes, better patient adherence, and enhanced oral health-related quality of life. However, more research is needed to fully understand and quantify these benefits over the long term.

As we move forward, it is imperative that the development and implementation of AI in periodontal care be guided by rigorous scientific inquiry and ethical considerations. Future research should focus on:

(I) Conducting large-scale, multicenter studies to validate the performance of deep learning models across diverse patient populations;

(II) Investigating the long-term impact of AI-assisted periodontal care on patient outcomes and quality of life through well-designed longitudinal studies;

(III) Developing standardized protocols for data collection, model development, and performance evaluation to facilitate meaningful comparisons across studies;
(IV) Exploring the integration of AI tools into existing clinical workflows to ensure seamless adoption and maximize their potential benefits;
(V) Addressing ethical and regulatory challenges to ensure responsible and equitable deployment of AI technologies in periodontal care.

In conclusion, while deep learning applications in periodontal care show great promise, their full potential can only be realized through continued research, interdisciplinary collaboration, and a commitment to patient-centered care. As these technologies continue to evolve, they have the potential to not only improve diagnostic and treatment capabilities but also to significantly enhance the quality of life for patients with periodontal disease. The future of periodontal care lies in the thoughtful integration of AI technologies with clinical expertise, always keeping the patient's well-being at the forefront of innovation.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-241/rc

Peer Review File: Available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-241/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-241/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Bajwa J, Munir U, Nori A, et al. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021;8:e188-94.
2. Weinberg J, Goldhardt J, Patterson S, et al. Assessment of accuracy of an early artificial intelligence large language model at summarizing medical literature: ChatGPT 3.5 vs. ChatGPT 4.0. J Med Artif Intell 2024;7:33.
3. Nwaiwu VC, Das SK. Emerging multifaceted application of artificial intelligence in chest radiography: a narrative review. J Med Artif Intell 2024;7:39.
4. Hayashi Y. Toward the transparency of deep learning in radiological imaging: beyond quantitative to qualitative artificial intelligence. J Med Artif Intell 2019;2:19.
5. Nawab K, Fernbach M, Atreya S, et al. Fine-tuning for accuracy: evaluation of Generative Pretrained Transformer (GPT) for automatic assignment of International Classification of Disease (ICD) codes to clinical documentation. J Med Artif Intell 2024;7:8.
6. Shirmohammadi A, Ghertasi Oskouei S. The growing footprint of artificial intelligence in periodontology & implant dentistry. J Adv Periodontol Implant Dent 2023;15:1-2.
7. Do J, Takei H, Carranza F. Chapter 32. Periodontal examination and diagnosis. In: Newman M, Takei H, Klokkevold P, et al. editors. Newman and Carranza's clinical periodontology. 13th ed. St. Louis: Elsevier, 2018.
8. Khan SF, Siddique A, Khan AM, et al. Artificial intelligence in periodontology and implantology—a narrative review. J Med Artif Intell 2024;7:6.
9. Ramani S, Vijayalakshmi R, Mahendra J, et al. Artificial intelligence in periodontics-An overview. IP Int J Periodontol Implantol 2023;8:71-4.
10. Moosa Y, Bacha SH, Raza SA, et al. Role of Artificial Intelligence in Periodontology. Pakistan Journal of Medical & Health Sciences 2023;17:363.
11. Meskó B. Data annotators are the unsung heroes of medicine's artificial intelligence revolution. J Med Artif Intell 2020;3:1.
12. Patil S, Joda T, Soffe B, et al. Efficacy of artificial intelligence in the detection of periodontal bone loss and

classification of periodontal diseases: A systematic review. J Am Dent Assoc 2023;154:795-804.e1.
13. Hoodbhoy Z, Hasan B, Siddiqui K. Does artificial intelligence have any role in healthcare in low resource settings? J Med Artif Intell 2019;2:13.
14. Bernauer SA, Zitzmann NU, Joda T. The Use and Performance of Artificial Intelligence in Prosthodontics: A Systematic Review. Sensors (Basel) 2021;21:6628.
15. Tucci V, Saary J, Doyle TE. Factors influencing trust in medical artificial intelligence for healthcare professionals: a narrative review. J Med Artif Intell 2022;5:4.
16. Goodson JM, Haffajee AD, Socransky SS. The relationship between attachment level loss and alveolar bone loss. J Clin Periodontol 1984;11:348-59.
17. Hausmann E, Allen K, Clerehugh V. What alveolar crest level on a bite-wing radiograph represents bone loss? J Periodontol 1991;62:570-2.
18. Hogg HDJ, Sendak MP, Denniston AK, et al. Unlocking the potential of qualitative research for the implementation of artificial intelligence-enabled healthcare. J Med Artif Intell 2023;6:8.
19. Hellén-Halme K, Lith A, Shi XQ. Reliability of marginal bone level measurements on digital panoramic and digital intraoral radiographs. Oral Radiol 2020;36:135-40.
20. Miller A, Huang C, Brody ER, et al. Artificial intelligence applications for the radiographic detection of periodontal disease: a scoping review. J Califor Dent Asso 2023;51:2206301.
21. Singh J, Sillerud B, Singh A. Artificial intelligence, chatbots and ChatGPT in healthcare—narrative review of historical evolution, current application, and change management approach to increase adoption. J Med Artif Intell 2023;6:30.
22. Van Es SL, Madabhushi A. The revolving door for AI and pathologists—docendo discimus? J Med Artif Intell 2019;2:12.
23. Tarnawski AS, Ahluwalia A. Application of artificial intelligence for the assessment of mucosal healing and inflammation. J Med Artif Intell 2019;2:7.
24. Altalhi AM, Alharbi FS, Alhodaithy MA, et al. The Impact of Artificial Intelligence on Dental Implantology: A Narrative Review. Cureus 2023;15:e47941.
25. Lafzi A, Mohammadi AS, Eskandari A, et al. Assessment of Intra- and Inter-examiner Reproducibility of Probing Depth Measurements with a Manual Periodontal Probe. J Dent Res Dent Clin Dent Prospects 2007;1:19-25.
26. de Souza LL, Fonseca FP, Martins MD, et al. ChatGPT and medicine: a potential threat to science or a step towards the future? J Med Artif Intell 2023;6:19.
27. Urban R, Haluzová S, Strunga M, et al. AI-assisted CBCT data management in modern dental practice: benefits, limitations and innovations. Electronics 2023;12:1710.
28. Machtei EE, Hausmann E, Grossi SG, et al. The relationship between radiographic and clinical changes in the periodontium. J Periodontal Res 1997;32:661-6.
29. Papapanou PN, Sanz M, Buduneli N, et al. Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. J Periodontol 2018;89 Suppl 1:S173-82.
30. Patel T, Vayon D, Ayilavarapu S, et al. A comparison study: Periodontal practice approach of dentists and dental hygienists. Int J Dent Hyg 2020;18:314-21.
31. Pepelassi EA, Diamanti-Kipioti A. Selection of the most accurate method of conventional radiography for the assessment of periodontal osseous destruction. J Clin Periodontol 1997;24:557-67.
32. Chaieb S, Garrouch K, Al-Ali NS. Perceptions of the use and benefits of artificial intelligence applications: survey study. J Med Artif Intell 2023;6:28.
33. Brożek B, Furman M, Jakubiec M, et al. The black box problem revisited. Real and imaginary challenges for automated legal decision making. Artif Intell Law 2024;32:427-40.
34. Schwendicke F, Singh T, Lee JH, et al. Artificial intelligence in dental research: Checklist for authors, reviewers, readers. J Dent 2021;107:103610.
35. Nazir M, Al-Ansari A, Al-Khalifa K, et al. Global Prevalence of Periodontal Disease and Lack of Its Surveillance. ScientificWorldJournal 2020;2020:2146160.
36. Nguyen TT, Larrivée N, Lee A, et al. Use of Artificial Intelligence in Dentistry: Current Clinical Trends and Research Advances. J Can Dent Assoc 2021;87:l7.
37. Wang X, Chen Y, Liu Y, et al. Reporting items for systematic reviews and meta-analyses of acupuncture: the PRISMA for acupuncture checklist. BMC Complement Altern Med 2019;19:208.
38. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.
39. Yu H, Ye X, Hong W, et al. A cascading learning method with SegFormer for radiographic measurement of periodontal bone loss. BMC Oral Health 2024;24:325.
40. Dai F, Liu Q, Guo Y, et al. Convolutional neural networks

combined with classification algorithms for the diagnosis of periodontitis. Oral Radiol 2024;40:357-66.

41. Chen CC, Wu YF, Aung LM, et al. Automatic recognition of teeth and periodontal bone loss measurement in digital radiographs using deep-learning artificial intelligence. J Dent Sci 2023;18:1301-9.
42. Alotaibi G, Awawdeh M, Farook FF, et al. Artificial intelligence (AI) diagnostic tools: utilizing a convolutional neural network (CNN) to assess periodontal bone level radiographically-a retrospective study. BMC Oral Health 2022;22:399.
43. Chang J, Chang MF, Angelov N, et al. Application of deep machine learning for the radiographic diagnosis of periodontitis. Clin Oral Investig 2022;26:6629-37.
44. Kearney VP, Yansane AM, Brandon RG, et al. A generative adversarial inpainting network to enhance prediction of periodontal clinical attachment level. J Dent 2022;123:104211.
45. Lee CT, Kabir T, Nelson J, et al. Use of the deep learning approach to measure alveolar bone level. J Clin Periodontol 2022;49:260-9.
46. Widyaningrum R, Candradewi I, Aji NRAS, et al. Comparison of Multi-Label U-Net and Mask R-CNN for panoramic radiograph segmentation to detect periodontitis. Imaging Sci Dent 2022;52:383-91.
47. Jiang L, Chen D, Cao Z, et al. A two-stage deep learning architecture for radiographic staging of periodontal bone loss. BMC Oral Health 2022;22:106.
48. Krois J, Garcia Cantu A, Chaurasia A, et al. Generalizability of deep learning models for dental image analysis. Sci Rep 2021;11:6102.
49. Chang HJ, Lee SJ, Yong TH, et al. Deep Learning Hybrid Method to Automatically Diagnose Periodontal Bone Loss and Stage Periodontitis. Sci Rep 2020;10:7531.
50. Feres M, Louzoun Y, Haber S, et al. Support vector machine-based differentiation between aggressive and chronic periodontitis using microbial profiles. Int Dent J 2018;68:39-46.
51. Lee JH, Kim DH, Jeong SN, et al. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci 2018;48:114-23.
52. Sunnetci KM, Ulukaya S, Alkan A. Periodontal bone loss detection based on hybrid deep learning and machine learning models with a user-friendly application. Biomed Signal Process Control 2022;77:103844.
53. Tadavarthi Y, Vey B, Krupinski E, et al. The State of Radiology AI: Considerations for Purchase Decisions and Current Market Offerings. Radiol Artif Intell 2020;2:e200004.
54. Waisberg E, Ong J, Kamran SA, et al. Bridging artificial intelligence in medicine with generative pre-trained transformer (GPT) technology. J Med Artif Intell 2023;6:13.
55. Tsoromokos N, Parinussa S, Claessen F, et al. Estimation of Alveolar Bone Loss in Periodontitis Using Machine Learning. Int Dent J 2022;72:621-7.
56. Wu Z, Ramsundar B, Feinberg EN, et al. MoleculeNet: a benchmark for molecular machine learning. Chem Sci 2018;9:513-30.
57. Schouten B, Schinkel M, Boerman AW, et al. Implementing artificial intelligence in clinical practice: a mixed-method study of barriers and facilitators. J Med Artif Intell 2022;5:12.
58. Zhang Z. Utilization of electronic healthcare records for advancement of medical knowledge. J Med Artif Intell 2018;1:9.
59. Schork NJ. Artificial Intelligence and Personalized Medicine. Cancer Treat Res 2019;178:265-83.
60. Blasiak A, Khong J, Kee T. CURATE.AI: Optimizing Personalized Medicine with Artificial Intelligence. SLAS Technol 2020;25:95-105.
61. Scott J, Biancardi AM, Jones O, et al. Artificial Intelligence in Periodontology: A Scoping Review. Dent J (Basel) 2023;11:43.
62. Zhu J, Chen Z, Zhao J, et al. Artificial intelligence in the diagnosis of dental diseases on panoramic radiographs: a preliminary study. BMC Oral Health 2023;23:358.

doi: 10.21037/jmai-24-241

Cite this article as: Ferrara E, Rapone B, D'Albenzio A. Applications of deep learning in periodontal disease diagnosis and management: a systematic review and critical appraisal. J Med Artif Intell 2025;8:23.