Journal of Clinical Periodontology

WILEY

Journal of Clinical Periodontology

SYSTEMATIC REVIEW

Emerging Applications of Digital Technologies for Periodontal Screening, Diagnosis and Prognosis in the Dental Setting

Roberto Farina[1,2] | Anna Simonelli[1,2] | Leonardo Trombelli[1,2] | Johanna B. Ettmayer[3] | Jan L. Schmid[3] | Christoph A. Ramseier[3]

[1]Research Centre for the Study of Periodontal and Peri-Implant Diseases, University of Ferrara, Ferrara, Italy | [2]Operative Unit of Dentistry, Azienda Unità Sanitaria Locale (AUSL), Ferrara, Italy | [3]Department of Periodontology, School of Dental Medicine, University of Bern, Bern, Switzerland

Correspondence: Christoph A. Ramseier (christoph.ramseier@unibe.ch)

Received: 28 October 2024 | **Revised:** 24 January 2025 | **Accepted:** 7 March 2025

Funding: The authors received no specific funding for this work.

Keywords: clinical trial | diagnosis | periodontal diseases | systematic review

ABSTRACT

Aim: To comprehensively review digital technologies (including artificial intelligence, AI) for periodontal screening, diagnosis and prognosis in the dental setting, focusing on accuracy metrics.

Materials and Methods: Two separate literature searches were conducted for periodontal screening and diagnosis (part I, scoping review) and prognosis (part II, systematic approach). PubMed, Scopus and Embase databases were searched.

Results: In part I, 40 studies evaluated AI and advanced imaging on different substrata. The combination of AI with 2D radiographs was the most frequently investigated and demonstrated a high level of periodontitis detection and stage definition. In part II, eight studies, identified as having a high risk of bias, tested supervised machine learning models using 6–74 predictors. The models demonstrated variable predictive accuracy, often outperforming traditional risk assessment tools and classical statistical models in the few studies evaluating such comparisons.

Conclusions: AI and advanced imaging techniques are promising for periodontal screening, diagnosis and prognosis in the dental setting, although the evidence remains inconsistent and inconclusive. In addition, AI-driven analysis of 2D radiographs (for diagnosis and staging of periodontitis), neural networks and the aggregation of multiple algorithms (for predicting tooth-related outcomes) appear to be the most promising approaches entering clinical application.

1 | Introduction

In periodontology, early screening (i.e., the preliminary, basic examination to estimate periodontal conditions and the treatment required) and accurate diagnosis (i.e., subject classification according to specific case definitions) are essential for effective disease management (Ramseier 2024). Traditional screening and diagnostic methods such as periodontal probing and conventional radiography, however, are characterised by a certain level of invasiveness and often lack the precision required to detect early or subtle changes in periodontal health. Despite technological advances, significant challenges remain in achieving early detection and accurate diagnosis of periodontal diseases.

The assessment of periodontitis risk, that is, the estimation of the probability of an individual/tooth to develop a negative

20th European workshop on periodontology 'Periodontal Diagnosis—Contemporary and Emerging Technologies' Group 3, Review #1.

event—the *label*—ranging from periodontitis onset/progression to periodontitis-related tooth loss, over a predefined prediction period, is another key component of the periodontal visit (Lang et al. 2015; Sanz et al. 2020). Risk stratification is clinically relevant to inform and motivate patients in the context of primary prevention strategies based on behaviour change (Asimakopoulou et al. 2015; Chapple et al. 2015; Jepsen et al. 2017; Tonetti et al. 2017; Farina et al. 2023, 2024), but it is also indicated when structuring secondary prevention strategies and monitoring their effectiveness by assessing their impact on disease risk indicators (Farina et al. 2021; Tonetti et al. 2017). In this context, several methods and tools have been proposed and partly validated to support oral health professionals (Chow et al. 2024; Du et al. 2018). These instruments rely on simple algorithms that combine a limited number of well-established risk factors, either previously selected for their documented prognostic impact or identified as candidate predictors in statistical models (e.g., multivariate analysis) (Farina et al. 2023; Lang et al. 2015; Trombelli and Farina 2020). However, these methods and tools can lead to oversimplification, especially in complex diseases including periodontitis, which are characterised by intricate interactions between variables and outcomes (Chiarito et al. 2022; Harrell et al. 2004; Shameer et al. 2018).

Tools, systems and devices that can generate, create, store or process data (which will be considered here under the broad term 'digital technologies') are increasingly used in dentistry. Among these, artificial intelligence (AI) has shown promise in improving the accuracy and efficiency of periodontal diagnostics (Pitchika et al. 2024). However, the integration of these technologies into clinical practice for screening and diagnosis is still in its early stages, and there is a lack of comprehensive reviews that synthesise current evidence across the range of technologies available. AI also holds promise as a tool for generating predictive models in medicine (Chiarito et al. 2022). However, a systematic review from 2018 focusing on predictive models for periodontitis incidence and progression found no studies applying AI-based models for this purpose (Du et al. 2018). A more recent systematic review also included studies of AI-based models but focused only on the prediction of tooth loss, and did not present comparative data to assess whether AI could be a valuable alternative to traditional risk assessment tools and statistical methods (Chow et al. 2024).

This review focuses on digital tools in the dental setting and aims to (i) summarise evidence regarding their application for periodontal disease screening and diagnosis (part I), and (ii) evaluate the efficacy of AI-based models in assessing the risk of periodontitis incidence/progression and tooth loss (part II).

2 | Materials and Methods

A scoping review was performed according to the framework proposed by Arksey and O'Malley (Arksey and O'Malley 2005) to thoroughly explore the current state of knowledge and preliminarily evaluate the feasibility of a systematic review for both parts (I and II). Based on the results, part I was maintained as a scoping review, while part II was managed by conducting a systematic review. For both parts, an a priori protocol was developed, assessed and approved by the Scientific Committee of the 20th European Workshop on Periodontology.

2.1 | Part I—Emerging Applications of Digital Technologies for Screening and Diagnosis of Periodontal Diseases in the Dental Setting

2.1.1 | Research Question

"How effective are digital technologies for the screening and diagnosis of periodontal diseases?"

2.1.2 | Identification of Relevant Studies

A broad, inclusive literature search was conducted across multiple databases for articles published in English up until 10 July 2024, without restrictions on study design or publication type. The search terms used in each database are detailed in Appendix S1.

2.1.3 | Study Selection

References were initially managed using EndNote (21.4, Clarivate, Philadelphia, PA). After removing duplicates, the records were screened (abstract first, then full text) to select studies addressing the use of digital technologies, including AI-based systems, advanced radiographic techniques and biomarker analysis in saliva and gingival crevicular fluid (GCF) for the screening and diagnosis of gingival inflammation/gingivitis and periodontitis in the dental practice.

2.1.4 | Data Extraction, Mapping and Synthesis

Three authors (C.A.R., J.L.S., J.B.E.) captured key information such as study design, population characteristics, index text (i.e., the type of digital technology used), reference test (i.e., the comparison) and outcomes measured from the selected studies. Metrics of interest included accuracy, sensitivity, specificity and clinical relevance of the digital technologies in the context of periodontal care. Given the expected heterogeneity of the included studies, data was narratively summarised using evidence tables and thematic analysis by type of emerging technology and its application in periodontal screening and diagnosis.

2.2 | Part II—Emerging Applications of Digital Technologies for the Evaluation of Periodontal Prognosis in the Dental Setting

2.2.1 | Focused Question (FQ)

"What is the accuracy of methods based on AI for the estimation of the risk of periodontal deterioration and tooth loss in the dental setting as compared to individual prognostic factors, traditional periodontitis risk assessment methods, traditional statistical models or any alternative AI-based intervention?"

2.2.2 | Protocol Development and Registration

The systematic review was registered in PROSPERO (ID: CRD42024573820) and follows the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) (Liberati et al. 2009; Moher et al. 2009).

2.2.3 | Study Selection Criteria

Study selection was based on the following inclusion criteria (PICOTS; Moons et al. 2019):

- Population: studies were included if they were conducted on patients who received a prognostic assessment and were enrolled in a follow-up study that measured tooth loss and/or periodontal attachment. To ensure focus on AI-based technologies for periodontal diseases, only studies with clearly defined periodontal conditions and/or tooth loss reasons related to periodontitis were included;
- Intervention (index test): any type of prognostic assessment (e.g., score, code, class) generated by digital technologies based on AI models, including supervised/unsupervised/reinforcement machine learning (ML), deep learning (DL) or other digital tools designed for periodontal care applied in a dental setting (due to the inclusion of at least one predictor to be assessed or collected by an oral health professional);
- Comparison (reference test): any prognostic assessment alternative to the intervention, including (i) individual prognostic factors; (ii) traditional periodontitis risk assessment methods; (iii) traditional statistical models (e.g., multilevel, multivariate models); or (iv) any of the above AI-based interventions;
- Outcome measures: outcome measures were related to the period between prognostic assessment and the last visit at which tooth loss was recorded and/or periodontal status was reassessed. Assessment of periodontal status during follow-up could be based on tooth loss (for any reason or for periodontal reasons), change in clinical attachment level (CAL), furcation lesions and radiographic measurements of bone levels. Primary outcome measures included: metrics describing the predictive accuracy of the model (e.g., sensitivity, specificity, positive and negative predictive values) or comparing the performance of different tools/models/algorithms (discrimination, i.e., the ability of a model to classify individuals to the correct outcome; calibration, i.e., the agreement between predicted and observed outcomes; net reclassification improvement, i.e., how well a model predicts an individual's outcome compared with a previous model; measures of overall model fit, e.g. proportion of time to event occurrence);
- Timing: studies were included if they evaluated prognostic models for use at the first periodontal visit (i.e., before step I of therapy), at the completion of steps II/III or during supportive periodontal care (SPC) (Sanz et al. 2020) for the prediction of a predefined label over a period of at least 6 months;
- Study design: retrospective or prospective longitudinal clinical trials. For studies evaluating a population with a heterogeneous periodontal diagnosis or undefined/unclassified baseline periodontal status, a further restriction was applied to studies evaluating periodontitis-related outcomes (e.g., periodontitis-related tooth loss, periodontal attachment loss due to periodontitis). Cross-sectional (association-type) studies evaluating the level of agreement between digital technologies and other methods (e.g., traditional periodontitis risk assessment tools) were also included. Studies were considered eligible regardless of the statistical unit (patient or tooth).

Additionally, exclusion criteria related to the Population, Intervention, Outcomes, Timing and Study design were applied to refine the selection of pertinent studies (Appendix S1).

2.2.4 | Literature Search

2.2.4.1 | Electronic and Hand Searches. According to the strategy reported in Appendix S1, Medline, Elsevier Scopus and Embase databases were searched for relevant literature from inception up until 15 August 2024. A separate search of the Cochrane Oral Health Group Specialty Trials' Register was conducted, and the reference lists of relevant systematic reviews were screened for the presence of eligible studies. Grey literature was retrieved through the Proquest database. The *Journal of Clinical Periodontology*, the *Journal of Periodontology*, the *Journal of Periodontal Research* and the *Journal of Dental Research—Clinical and Translational Research* were hand-searched. The reference lists of the selected publications were screened for the presence of eligible studies. Only full-text articles written in English were included. Titles and abstracts from the electronic searches were managed using Zotero (v7.0.7; Corporation for Digital Scholarships).

2.2.4.2 | Screening Methods. Duplicates were removed using Rayyan software (Cambridge, MA). Two authors (A.S. and R.F.) performed the primary search by independently screening the titles and abstracts. The same reviewers selected full manuscripts for studies meeting the inclusion criteria. After identifying the studies to be included, the authors resolved disagreements by discussion. If consensus was not reached, disagreements were resolved by discussion with a third author (L.T.).

2.2.4.3 | Data Extraction: Intervention Characterisation. Two reviewers (A.S. and R.F.) extracted data in duplicate and resolved disagreements by discussion. For each included study, data were retrieved and recorded using specially designed forms.

2.2.4.4 | Assessment of Reporting Quality and Risk of Bias in Individual Studies. Adherence to reporting standards was recorded for each study. The risk of bias (RoB) was assessed using PROBAST (Moons et al. 2019).

2.2.4.5 | Statistical Analysis. A meta-analysis was planned but was not possible due to the heterogeneity of study designs, populations and prediction models/outcomes/labels examined. Therefore, data from the included studies were reported narratively and summarised in Tables 8 and 9.

3 | Results

The digital technologies, as grouped by field of application (i.e., periodontal screening, diagnosis or prognosis), data source (e.g., radiographs, health records) and year of publication of each study, are illustrated in Figure 1. The results are presented separately for parts I and II.

3.1 | Part I—Emerging Applications of Digital Technologies for Screening and Diagnosis of Periodontal Diseases in the Dental Setting

3.1.1 | Study Identification and Categorisation

The study selection process resulted in the inclusion of 40 studies (Figure 2a) that focused on periodontal screening, diagnosis or both.

Studies were grouped according to the data source(s) used with digital technologies for periodontal screening and diagnosis (Table 1). The most common materials were radiographs, either 2D ($n=14$) or 3D, such as micro-CT or cone beam computed tomography ($n=7$), followed by photographs (primarily intraoral but including also those made by mobile phone; $n=7$), clinical data ($n=6$), oral fluid samples (particularly saliva and GCF; $n=5$) and electronic health records ($n=1$). The most common digital technologies were AI-based methods, particularly DL and ML algorithms.

3.1.2 | Technologies Using Radiographs

Digital technologies applied to bidimensional (2D) and tridimensional (3D) radiographs include AI-driven models, such as convolutional neural networks (CNNs) and other DL methods, which are evaluated for their periodontal diagnostic accuracies (Tables 1 and 2). For 2D radiographs, these AI models consistently demonstrate high accuracy in classifying periodontal conditions, such as alveolar bone loss and periodontitis stages. In particular, studies by Alotaibi et al. (2022) and Bayrakdar et al. (2020) illustrate how CNNs can be effectively used to detect alveolar bone loss with high sensitivity and specificity, potentially outperforming conventional diagnostic approaches (Alotaibi et al. 2022; Bayrakdar et al. 2020). Furthermore, the study by Dai et al. (2024) illustrates the potential of combining CNNs with classification algorithms to accurately diagnose the stages of periodontitis (Dai et al. 2024).

Digital technologies applied to tridimensional radiographs (i.e., micro-CT, CBCT) consisted of AI-driven models and advanced imaging techniques and were evaluated for periodontal diagnosis (Tables 1 and 3). The combination of these technologies with CBCT demonstrated high diagnostic accuracy in several studies. In particular, the study by Yusof et al. (2021) found that CBCT had better sensitivity compared with periapical radiographs in detecting furcation defects (Yusof et al. 2021). Additionally, Padmanabhan et al. (2017) and Qiao et al. (2014) both demonstrated that measurements performed with advanced imaging on CBCT at mandibular and maxillary molar furcation defects were in close agreement with surgical findings, demonstrating the trueness of the radiographic assessments (Padmanabhan et al. 2017; Qiao et al. 2014). Meanwhile, Kazimierczak et al. (2024) highlighted the integration of AI with CBCT, showing that AI can significantly improve the sensitivity

FIGURE 1 | Number of studies per year of publication, as grouped according to digital technology and data source. (a) Periodontal screening and diagnosis in the dental setting (part I, scoping review). (b) Periodontal prognosis in the dental setting (part II, systematic review).

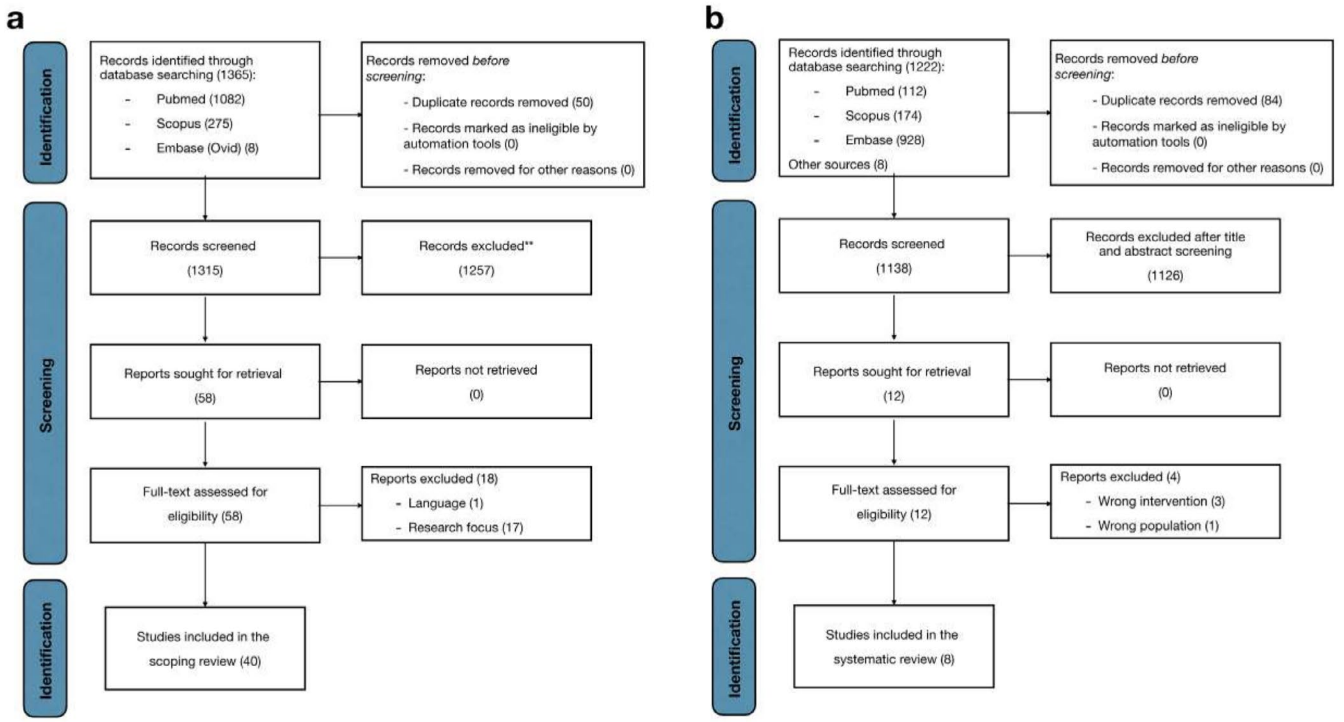

FIGURE 2 | Study selection flow (PRISMA 2020 flow diagram). (a) Emerging applications of digital technologies for periodontal screening and diagnosis in the dental setting. (b) Emerging applications of digital technologies for periodontal prognosis in the dental setting.

of early periodontitis detection (Kazimierczak et al. 2024). The only study conducted on micro-CT correlated 3D bone loss measurements with periodontitis stages, suggesting a strong potential of advanced imaging for detailed periodontal assessment (Hong et al. 2020).

3.1.3 | Technologies Using Photographs

The use of AI-driven models to analyse intraoral photographs for periodontal screening is an emerging trend (Tables 1 and 4). Studies such as Alalharith et al. (2020) and Chau et al. (2023) demonstrated how AI, particularly CNN-based models, can effectively detect early signs of gingival inflammation (Alalharith et al. 2020; Chau et al. 2023). For example, Alalharith et al. (2020) reported 100% accuracy in detecting teeth and 77.12% accuracy in detecting gingival inflammation (Alalharith et al. 2020). These findings suggest that DL models may have potential for early, non-invasive diagnosis, potentially allowing for more proactive periodontal care. In addition, research by Kurt-Bayrakdar et al. highlighted the versatility of AI in tooth numbering and detecting various gingival conditions, achieving high performance in tooth numbering and moderate performance in detecting signs of gingival overgrowth and inflammation (Kurt-Bayrakdar et al. 2023). This may suggest that AI models could significantly accelerate the digital transformation of dental diagnostics by automating the analysis of intraoral photographs.

As demonstrated in the reviewed studies, AI models can detect subtle changes in gingival health that may be missed during routine visual inspection. For example, Li et al. (2021) applied a multi-task learning CNN model to RGB photographs and achieved an area under the curve (AUC) of 87.11% for the detection of gingival inflammation, illustrating the potential of these models to support large-scale dental screening programmes (Li et al. 2021).

3.1.4 | Technologies Using Oral Fluid Samples

The use of oral fluid samples, such as saliva and GCF, in combination with advanced technologies, such as ML and infrared spectroscopy, represents an emerging approach to periodontal screening and diagnosis (Tables 1 and 5). The application of ML to salivary biomarkers and genomic data has shown high accuracy in discriminating between different severities of periodontal disease. Among others, Deng et al. (2023) reported an AUC greater than 0.94 in the classification of periodontal health, gingivitis and periodontitis, indicating that ML models can be effective in non-invasive periodontal screening and diagnosis (Deng et al. 2023). Similarly, Kim et al. (2020) demonstrated that ML models could achieve 93% accuracy in classifying 'healthy' versus 'moderate/severe' periodontitis based on salivary bacterial profiles (Kim et al. 2020).

Another notable approach is the use of Fourier transform infrared spectroscopy (FT-IR) combined with ML, as evaluated by da Silva et al. (2024) (da Silva et al. 2024). This method achieved classification accuracies above 80% for both periodontitis and diabetes, suggesting that FT-IR could serve as a rapid and non-invasive screening tool.

Analysis of GCF using antibody arrays and ML, as demonstrated by Huang et al. (2020), also showed promising results with an AUC of 0.984 for predicting severe periodontitis (Huang et al. 2020).

TABLE 1 | Emerging applications of digital technologies for periodontal screening and diagnosis in the dental setting.

Author	Year	Digital technology	Category	Data source	Screening	Diagnosis
Bidimensional radiographs ($n = 14$)						
Alotaibi et al.	2022	Deep Convolutional Neural Network (CNN)	AI	Periapical radiographs		X
Bayrakdar et al.	2020	Convolutional Neural Networks (CNNs)	AI	Panoramic radiography		X
Chang et al.	2020	DL Hybrid Method	AI	Panoramic radiography		X
Chang et al.	2022	ML using Inception V3 Model	AI	Radiographs		X
Dai et al.	2024	Convolutional Neural Networks (CNNs), Classification Algorithms	AI	Periapical radiographs		X
Ertaş et al.	2022	ML Algorithms using Python, Hybrid Network Model with ResNet50	AI	Panoramic radiographic images		X
Guler Ayyildiz et al.	2024	DL, Global Average Pooling (GAP), Global Max Pooling (GMP), Flatten Layers (FL)	AI	Panoramic radiographs		X
Hoss et al.	2023	Convolutional Neural Networks (CNNs)	AI	Periapical radiographs		X
Kabir et al.	2022	Guided Framework for Dental Diagnosis	AI	Panoramic and intraoral radiographs		X
Kurt-Bayrakdar et al.	2024	Convolutional Neural Networks (CNNs), U-Net Architecture	AI	Panoramic radiographs		X
Lee et al.	2022	Deep Convolutional Neural Network (CNN) for Alveolar Bone Level Measurement	AI	Periapical radiographic images		X
Li et al.	2020	Deetal-Perio, Mask R-CNN	AI	Panoramic radiographs		X
Liu et al.	2023	Convolutional Neural Networks (CNNs), Gradient-weighted Class Activation Mapping	AI	Panoramic radiographs (PARs)		X
Tsoromokos et al.	2022	Convolutional Neural Networks (CNN) for Alveolar Bone Loss Estimation	AI	Periapical radiographs		X

(Continues)

TABLE 1 | (Continued)

Author	Year	Digital technology	Category	Data source	Screening	Diagnosis
Tridimensional radiographs ($n=7$)						
Hong et al.	2020	3D Clinical Attachment Loss (3D-CAL), Micro-Computed Tomography (Micro-CT)	Advanced imaging	Radiograph		X
Kazimierczak et al.	2024	Cone Beam Computed Tomography (CBCT), Diagnocat Software	AI	CBCT images and panoramic radiography (OPG)		X
Manavella et al.	2017	Cone Beam Computed Tomography (CBCT)	Advanced imaging	Radiograph		X
Mohan et al.	2014	Cone Beam Computed Tomography (CBCT)	Advanced imaging	Radiograph		X
Padmanabhan et al.	2017	Cone Beam Computed Tomography (CBCT)	Advanced imaging	Radiograph		X
Qiao et al.	2014	Cone Beam Computed Tomography (CBCT)	Advanced imaging	Radiograph		X
Yusof et al.	2021	Cone Beam Computed Tomography (CBCT), Periapical Radiograph	Advanced imaging	Radiographs		X
Photographs ($n=7$)						
Alalharith et al.	2020	Faster Region-Based Convolutional Neural Networks (R-CNNs) using ResNet-50	AI	Intraoral photographs	X	
Chau et al.	2023	AI-based System for Gingivitis Detection	AI	Intraoral photographs	X	
Chen and Chen et al.	2020	Grey-Level Co-Occurrence Matrix (GLCM), Artificial Neural Network (ANN), Genetic Algorithms (GA)	AI	Intraoral photographs	X	
Kurt-Bayrakdar et al.	2023	Convolutional Neural Networks (CNNs), YOLOv5x Architecture	AI	Intraoral photographs	X	

(Continues)

TABLE 1 | (Continued)

Author	Year	Digital technology	Category	Data source	Screening	Diagnosis
Li et al.	2018	Contrast-Limited Adaptive Histogram Equalisation (CLAHE), Grey-Level Co-Occurrence Matrix (GLCM), Extreme Learning Machine (ELM)	AI	Intraoral photographs	X	
Li et al.	2019	Emerging Technologies Investigated: Multichannel Grey-Level Co-Occurrence Matrix (MGLCM), Particle Swarm Optimization Neural Network (PSONN)	AI	Intraoral photographs	X	
Li et al.	2021	Multichannel Grey-Level Co-Occurrence Matrix (MGLCM), Particle Swarm Optimization Neural Network (PSONN)	AI	Intraoral photographs	X	
Oral fluid samples ($n=5$)						
da Silva et al.	2024	Fourier Transform Infrared Spectroscopy (FT-IR), ML	AI	Saliva samples	X	
Deng et al.	2023	ML, Random Forest, Logistic Regression, Salivary Biomarkers	AI	Saliva samples	X	
Huang et al.	2020	Antibody Array, ML Classifiers	AI	Gingival crevicular fluid (GCF)	X	
Kim et al.	2020	ML Models, Salivary Bacterial Copy Number	AI	Saliva samples	X	
Shen et al.	2022	AI-Assisted Dental Monitoring Application	AI	Smartphone camera intra-oral scanning	X	
Health records ($n=1$)						
Tastan Eroglu et al.	2024	ChatGPT Language Model	AI	Baseline digital data of periodontitis patients		X
Clinical data ($n=6$)						
Alqahtani et al.	2022	Random Forest, Classification and Regression Tree Analysis (CART)	AI	Clinical data		X

(Continues)

TABLE 1 | (Continued)

Author	Year	Digital technology	Category	Data source	Screening	Diagnosis
Bashir et al.	2022	ML Algorithms for Predictive Modelling	AI	Clinical data		X
Farhadian et al.	2020	Support Vector Machine (SVM)	AI	Clinical data		X
Lakshmi and Dheeba	2023	ML, DL, Cross-Validation, Ensemble Learning	AI	Clinical data, patient records		X
Papantonopoulos et al.	2014	Artificial Neural Networks (ANNs)	AI	Clinical and immunologic datasets	X	X
Snider et al.	2024	Dental Monitoring (DM) Technology	AI	Oral scans, clinical oral examinations		X

Abbreviations: 3D-CAL, 3D Clinical attachment loss; AI, artificial intelligence; ANN, artificial neural network; CART, classification and regression tree analysis; CBCT, cone beam computed tomography; CLAHE, contrast-limited adaptive histogram equalisation; CNN, convolutional neural network; DL, deep learning; DM, dental monitoring; ELM, extreme learning machine; FL, flatten layers; FT-IR, Fourier transform infrared spectroscopy; GA, genetic algorithms; GAP, global average pooling; GCF, gingival crevicular fluid; GMP, global max pooling; MGLCM, multichannel grey-level co-occurrence matrix; Micro-CT, micro-computed tomography; OPG, orthopantomogram (panoramic radiography); PSONN, particle swarm optimization neural network; R-CNN, region-based convolutional neural network; SVM, support vector machine.

3.1.5 | Technologies Using Health Records

The integration of electronic health records with AI-based technologies shows emerging potential in periodontal diagnostics (Tables 1 and 6). Health records typically consist of demographic data, medical history and administrative data collected through healthcare services. These may include general patient characteristics (e.g., age, sex, medical history) and broader healthcare information (e.g., insurance data, medications, diagnoses) rather than direct clinical measurements specific to a periodontal exam. Tastan Eroglu et al. (2024) evaluated the AI tool ChatGPT for periodontitis classification, with accuracy rates of 59.5% for stage and 84% for extent classification (Tastan Eroglu et al. 2024).

3.1.6 | Technologies Using Clinical Data

The use of clinical data in combination with ML and other AI-driven models is demonstrating potential in the screening and diagnosis of periodontal diseases (Tables 1 and 7). Clinical data refers to direct, patient-specific health data collected during clinical assessments, including periodontal parameters and physical exam findings. For example, studies by Alqahtani et al. (2022) and Bashir et al. (2022) identified significant predictors of periodontitis, such as age and comorbidities, and achieved high accuracy during internal validation (AUC >0.95) (Alqahtani et al. 2022; Bashir et al. 2022).

Support vector machines (SVMs) have also been used in studies such as Farhadian et al. (2020) and Lakshmi and Dheeba (2023), showing an accuracy of around 88.7% in classifying the severity of periodontal disease (Farhadian et al. 2020; Lakshmi and Dheeba 2023). Artificial neural networks (ANNs), as used by Papantonopoulos et al. (2014), showed high accuracy (90%–98%) in distinguishing between aggressive and chronic periodontitis, although these results are based on small sample sizes and require further validation (Table 7) (Papantonopoulos et al. 2014).

Additionally, AI-based remote monitoring, evaluated by Snider et al. (2024), demonstrated moderate accuracy but low sensitivity in detecting conditions such as gingival inflammation (Table 7) (Snider et al. 2024).

3.2 | Part II—Emerging Applications of Digital Technologies for the Evaluation of Periodontal Prognosis in the Dental Setting

3.2.1 | Search Results and Description of the Included Studies

The study selection resulted in the inclusion of eight articles/studies (Krois et al. 2019; Troiano et al. 2023; Lee et al. 2023, 2024; Nagarajan et al. 2019; Patel et al. 2023; Santamaria et al. 2024; Schwendicke et al. 2021) (Figure 2b) from six patient populations.

The main characteristics of the included studies are presented in Table 8. Studies consisted of retrospective analyses of either

TABLE 2 | Emerging applications of digital technologies using bidimensional radiographs for periodontal screening and diagnosis in the dental setting.

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Alotaibi et al. (2022)	Develop a CNN-based model for tooth numbering and abnormality detection	Convolutional Neural Network (CNN)	Retrospective study with 1610 patients and 1724 images	High diagnostic accuracy for alveolar bone loss detection (73%) and severity classification (59%)	Moderate accuracy in classifying the severity of bone loss; limited to anterior teeth	CNNs could enhance diagnostic workflows in identifying and classifying dental abnormalities in radiographs	Optimization and validation of CNN models across more varied dental conditions and image types are necessary
Bayrakdar et al. (2020)	Detect alveolar bone loss using AI from panoramic radiographs	Convolutional Neural Network (CNN), Google Net Inception V3, Transfer learning	2276 panoramic radiographs; training ($n = 1856$), validation ($n = 210$), testing ($n = 210$)	The CNN detected bone loss with 0.94 sensitivity, 0.88 specificity, 0.89 precision, 0.91 accuracy, and 0.91 F1 score	Limited to panoramic radiographs; need for broader application and validation	AI systems could enhance periodontal diagnostic accuracy and efficiency	Broader validation across different imaging types and clinical environments is necessary
Chang et al. (2020)	Develop an automatic method to stage periodontitis on panoramic radiographs	DL Hybrid Method	340 panoramic radiographs analysed with hybrid DL architecture, training (90%), testing (10%)	High intra-class correlation (0.91); accurate periodontitis staging (Pearson correlation 0.73, $p < 0.01$)	Limited to panoramic radiographs; may not generalise to other imaging modalities	DL methods could be integrated into routine periodontal assessments for more accurate staging of periodontitis	Broader validation on different types of radiographs and in varied clinical environments is required
Chang et al. (2022)	Apply DL for radiographic diagnosis of periodontitis	DL	236 patients' full-mouth radiographs analysed; model tested for bone loss detection	Model accuracy 87% for categorising mild (< 15%) or severe (≥ 15%) bone loss	Limited to standardised radiographic settings; real-world generalisability not fully tested	DL models could support periodontal diagnosis and treatment planning by assessing radiographic bone loss	Additional data needed to improve model construction and performance for clinical implementation

(Continues)

TABLE 2 | (Continued)

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Dai et al. (2024)	Develop a DL model combining CNN and classification algorithms to assist in diagnosing periodontitis stages	Convolutional Neural Networks (CNN) combined with Classification Algorithms (CA)	Periapical radiographs and clinical data, models from 834 (training) and 278 (testing) individuals trained with Alexnet, VGG16, ResNet18, and classification algorithms (RF, SVM, NB, LR and KNN)	PER-Alexnet + RF model achieved highest accuracy with 0.968 for control, 0.960 for stage I, 0.835 for stage II, and 0.842 for stage III/IV; significant relation of age and smoking with periodontitis	Limited to the specific dataset used	The model can assist dentists in quickly and accurately diagnosing periodontitis stages	Broader validation and testing on diverse datasets needed
Ertaş et al. (2022)	Evaluate DL models to determine and facilitate the staging and grading of periodontitis	DL	Clinical data from 144 individuals, processed panoramic radiographs	97.2%–98.6% accuracy in staging using clinical data; 88.2% accuracy in staging with hybrid model on radiographs	Limited performance on maxillary roots; requires improvement in model precision	DL models could assist in detecting apical periodontitis, improving diagnostic speed and accuracy	Model refinement to improve detection precision, particularly in maxillary regions
Guler Ayyildiz et al. (2024)	Assess the accuracy of computer-assisted periodontal classification using DL on panoramic radiographs and compare different models and layers	DL methods including ResNet50, DenseNet121, InceptionV3 combined with ML models	2533 panoramic radiographs classified into 'healthy', 'Stage 1/2', and 'Stage 3/4'; features extracted using GAP, GMP, FL and reduced with mRMR; input into 8 ML models	DenseNet121 + GAP + mRMR-based ML techniques outperformed other models, achieving higher performance in periodontal bone loss classification	Limited to specific dataset and radiographic types used	The model enhances periodontal bone loss classification without the need for manual feature selection	Broader validation and testing with diverse datasets and radiographic images necessary

(Continues)

TABLE 2 | (Continued)

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Hoss et al. (2023))	Assess effectiveness of AI for detecting periodontal bone loss	Convolutional Neural Networks (CNNs)	Diagnostic study using 21,819 radiographs	Accuracy (91%), sensitivity (94%), specificity (88%)	High variability in performance across different types of bone loss	CNNs could augment traditional diagnostic tools in periodontal assessments	Further refinement of CNN algorithms is needed for consistent performance across various clinical scenarios
Kabir et al. (2022)	DL approach to measure alveolar bone level	Deep Convolutional Neural Network	Tested on 1240 intraoral radiographs (different from the training and internal validation cohort)	The framework achieved high precision and recall (0.96 for panoramic view, 0.87 for repository match) and arranged intraoral radiographs into FMS templates with high accuracy (95% for periapical, 90% for bitewing). It performed well even in challenging conditions (94% accuracy with missing teeth, 89% with restorations)	Limited to a specific type of radiographic image; potential for bias in training data	DL models could be integrated into clinical practice to support accurate assessment of alveolar bone levels	Further optimisation and validation in broader clinical environments are essential for widespread adoption

(Continues)

TABLE 2 | (Continued)

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Kurt-Bayrakdar et al. (2024)	To develop a DL algorithm for interpreting panoramic radiographs and evaluating its performance in detecting periodontal bone losses and furcation defects	DL, Convolutional Neural Network (CNN) with U-Net architecture	1121 panoramic radiographs; labelled for total alveolar bone loss (2251), interdental bone loss (25303), furcation defects (2815); further divided into horizontal ($n = 21{,}839$) and vertical ($n = 3464$) bone losses	The algorithm performed best in detecting total alveolar bone loss (AUC = 0.951) and worst in detecting vertical bone losses (AUC = 0.733)	Lower performance in detecting vertical bone losses	AI systems can aid in detailed periodontal diagnosis and treatment planning	Improve vertical bone loss detection and validate in diverse datasets
Lee et al. (2022)	Measure alveolar bone level using DL in panoramic radiographs	DL	Radiographic image analysis from 693 periapical radiographs (37 periodontitis patients), 70% training, 10% validation, 20% testing	The DL model achieved an average Dice Similarity Coefficient (DSC) over 0.91. No significant difference was found between DL and examiner measurements for RBL percentage. AUROC for RBL stage assignment was 0.89–0.90 for stages I-III. Diagnostic accuracy was 0.85	Limited to specific radiographic views; needs broader validation	DL models could enhance periodontal diagnosis in routine radiographic examinations	Expansion of validation studies to include different types of radiographic images

(Continues)

TABLE 2 | (Continued)

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Li et al. (2020)	To develop an AI-based method for diagnosing chronic gingivitis using MGLCM and PSONN	Multichannel Grey-Level Co-Occurrence Matrix (MGLCM), Particle Swarm Optimization Neural Network (PSONN)	800 images (400 gingivitis, 400 healthy); various training algorithms compared	Accuracy, sensitivity, specificity, precision and F1 score all around 78.2%; outperformed NBC, WN + SVM, ELM and CLAHE + ELM	Moderate accuracy, further improvement needed	AI methods like MGLCM and PSONN can improve efficiency in diagnosing gingivitis, especially where medical resources are limited	Future work should focus on improving model accuracy and testing in diverse clinical environments
Liu et al. (2023)	Use DL and CAD to assess periodontal bone loss and disease staging	DL and Computer-Aided Design (CAD)	Panoramic radiographs from 1882 periodontitis cases and 1537 healthy controls; model training and validation	The model achieved AUCs of 0.843 and 0.793; accuracy comparable to periodontal experts; faster reading time	Slightly lower accuracy than experts; needs further validation	Hybrid DL and CAD systems could improve diagnostic accuracy and efficiency in clinical settings	Need for external validation and testing with different imaging types and settings
Tsoromokos et al. (2022)	Examine diagnostic performance of CNNs for periodontal bone loss	Convolutional Neural Networks (CNNs)	1546 approximal sites from 54 participants; training set (1308 sites), validation set (98 sites), test set (140 sites)	CNN scored a mean of 23.1% ABL with moderate reliability (ICC = 0.601). Performance was better on non-molar teeth (ICC = 0.763)	Variation in accuracy across different tooth locations; limited generalizability due to specific training data	CNNs can potentially support clinical decision-making in diagnosing and staging periodontal bone loss	Need for improved models that can handle variations in tooth location and different types of bone loss

Abbreviations: ABL, alveolar bone loss; AI, artificial intelligence; AlexNet, a convolutional neural network model; AUC, area under the curve; CAD, computer-aided design; CLAHE, contrast-limited adaptive histogram equalisation; CNN, convolutional neural network; DL, deep learning; DSC, dice similarity coefficient; ELM, extreme learning machine; FMS, full mouth series; Google Net Inception V3: A deep learning architecture developed by Google; ICC, intra-class correlation coefficient; KNN, k-nearest neighbour; MGLCM, multichannel grey-level co-occurrence matrix; NBC, naïve Bayes classifier; PSONN, particle swarm optimization neural network; RBL, radiographic bone loss; RF, random forest; U-Net, a convolutional network architecture designed for biomedical image segmentation; VGG16, a deep learning model with 16 layers; WN + SVM, weighted network plus support vector machine.

TABLE 3 | Emerging applications of digital technologies using tridimensional radiographs for periodontal screening and diagnosis in the dental setting.

First author	Objective	Data source	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Hong et al. (2020)	Correlate 3D bone loss and CRR with periodontitis classification	Micro-computed Tomography (micro-CT)	Analysis of 90 premolars with 3D SBL, CAL and CRR	Strong correlation found; 3D-CAL and STA parameters significant for evaluating periodontitis stages	Small sample size; limited to single-rooted premolars	3D imaging techniques could enhance diagnostic accuracy in assessing and classifying periodontitis	Expansion of research to multi-rooted teeth and different stages of periodontitis needed
Kazimierczak et al. (2024)	Evaluate the effectiveness of AI in detecting early signs of periodontitis	CBCT and orthopantomograms (OPG)	Study included 49 patients, totalling 1223 teeth; OPG and CBCT images analysed by AI and three clinicians, compared with human consensus using CBCT	AI sensitivity for OPG images was 33.33% with an F1 score of 32.73%; for CBCT images, sensitivity was 77.78% with an F1 score of 84.00%. Specificity was over 98% for both image types	Limited to specific imaging conditions; performance varies across different patient demographics	AI could be integrated into routine dental check-ups for early detection of periodontitis	Validation in a broader clinical setting with more diverse populations is required
Manavella et al. (2017)	Present and validate a volumetric analysis method for resorbed sockets	CBCT combined with image processing techniques	9 severely resorbed sockets analysed with different segmentation techniques	Automated segmentation (Mimics) had the smallest error (1.5%) compared with manual methods	Small sample size; only tested on severely resorbed sockets	Automated segmentation can enhance the accuracy of volumetric analysis in periodontal assessments	Validation on a larger sample size and different types of resorptions is necessary
Mohan et al. (2014)	Assess accuracy of CBCT for aggressive periodontitis	CBCT	Case study; comparison of CBCT with direct surgical measurements ($n=1$)	CBCT accurately measured osseous defects; measurements identical to surgical findings	Single patient case, not generalisable	CBCT can be a valuable diagnostic tool for detailed assessment of osseous defects	Larger studies needed to confirm findings across different patient populations

(Continues)

TABLE 3 | (Continued)

First author	Objective	Data source	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Padmanabhan et al. (2017)	Compare CBCT with direct measurements in mandibular molar furcation	CBCT	14 patients with 25 M sites; comparison with intra-surgical measurements	CBCT measurements closely matched surgical measurements with no significant difference	Limited to mandibular molars; small sample size	CBCT is an effective adjunctive tool for assessing furcation defects in mandibular molars	Broader studies across different tooth types and furcation grades are recommended
Qiao et al. (2014)	Investigate accuracy of CBCT in assessing maxillary molar furcation	CBCT	15 patients with chronic periodontitis; CBCT vs. surgical findings ($n = 20$ M)	CBCT confirmed 82.4% of surgical findings; weighted kappa of 0.917	Limited sample size; underestimation of bone loss by CBCT in some cases	CBCT can be a reliable tool for assessing periodontal tissue loss, especially in complex cases	Further studies should explore ways to improve CBCT accuracy in detecting horizontal and vertical bone loss
Yusof et al. (2021)	Evaluate accuracy of radiographic techniques for furcation defects	CBCT and Periapical Radiograph	Longitudinal RCT with 22 periodontitis patients ($n = 22$)	CBCT provided better diagnostic sensitivity compared with periapical radiograph (sensitivity: 62.8% vs. 56.9%)	Small sample size; limited to specific clinical parameters	CBCT should be considered over periapical radiographs in complex furcation involvement cases	Additional parameters should be investigated to enhance the diagnostic utility of CBCT

Abbreviations: 3D-CAL, 3D clinical attachment loss; AI, artificial intelligence; CBCT, cone beam computed tomography; CRR, crown-to-root-ratio; OPG, orthopantomogram (panoramic radiography); RCT, randomised controlled trial; SBL, subtraction bone level; STA, soft tissue attachment.

TABLE 4 | Emerging applications of digital technologies using photographs for periodontal screening and diagnosis in the dental setting.

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Alalharith et al. (2020)	Detect early gingivitis in orthodontic patients using DL	DL, Faster R-CNN, ResNet-50 CNN	134 intraoral images; training ($n = 107$), test ($n = 27$)	Teeth detection model achieved 100% accuracy; inflammation detection model had 77.12% accuracy	Limited sample size, moderate recall for gingival inflammation detection	DL models show promise for early non-invasive diagnosis of gingivitis	Further refinement and larger sample validation needed to improve diagnostic performance
Chau et al. (2023)	Assess AI-based photographic detection of gingivitis	AI	567 intraoral photographs analysed, 80% training, 20% validation	Sensitivity 0.92, specificity 0.94, mean intersection-over-union 0.60	Limited to frontal view photographs; needs validation across diverse datasets	AI could monitor plaque control effectiveness, supporting professional advice	Validation with diverse populations and inclusion of various intraoral angles and conditions
Chen and Chen (2020)	Improve gingivitis diagnosis using GLCM and ANN	Grey-Level Co-Occurrence Matrix (GLCM) and ANN	180 oral images (90 gingivitis, 90 healthy); K-fold cross-validation	The proposed method outperformed CLAHE and GLCM + ELM in identification performance	Limited dataset; requires comparison with other diagnostic methods	GLCM and ANN combination could be used as an adjunctive diagnostic tool for gingivitis in clinical settings	Expanding dataset and comparison with other diagnostic tools and populations needed
Kurt-Bayrakdar et al. (2023)	Develop AI model for tooth numbering and gingival condition detection	AI model using YOLOv5x architecture	654 intraoral photographs; labelled by three periodontists (16,795 teeth, 2493 frenulum attachments, 1211 gingival overgrowth areas, 2956 gingival inflammation signs)	High performance in tooth numbering (AUC = 0.989) and good performance in detecting other conditions (AUC = 0.774–0.827)	Variable performance in detecting gingival overgrowth and inflammation signs	AI can accelerate digital transformation in dental diagnostics by automating analysis of intraoral photographs	Further refinement needed for detecting gingival overgrowth and inflammation, and validation on larger datasets

(Continues)

TABLE 4 | (Continued)

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Li et al. (2019)	Explore the impact of AI-enabled toothbrush on periodontal treatment outcomes	CLAHE, Grey-Level Co-Occurrence Matrix (GLCM), Extreme Learning Machine (ELM)	93 images (58 gingivitis, 35 healthy)	Sensitivity 75%, specificity 73%, precision 74%, accuracy 74%	Small dataset; moderate accuracy	Offers a more accurate and sensitive alternative for gingivitis diagnosis compared with other methods	Further research needed to validate findings on larger datasets and improve model accuracy
Li et al. (2019)	Develop an AI-based method to diagnose chronic gingivitis using multichannel grey-level co-occurrence matrix and particle swarm optimisation neural network	Multichannel Grey-Level Co-Occurrence Matrix (MGLCM) and Particle Swarm Optimization Neural Network (PSONN)	800 images: 400 chronic gingivitis and 400 healthy gingiva, comparison with other training algorithms	MGLCM + PSONN achieved specificity, sensitivity, precision, accuracy and F1 Score all around 78.20%, outperforming NBC, WN + SVM, ELM and CLAHE + ELM	Limited by moderate accuracy and the specific dataset used	The method offers a more accurate and efficient alternative for gingivitis diagnosis compared with traditional approaches	Further validation and improvement of model accuracy with larger and more diverse datasets needed
Li et al. (2021)	Screen gingivitis in RGB photos using DL	DL	625 patients' photos analysed; Multi-Task Learning CNN model applied	AUC for detecting gingivitis 87.11%, for calculus 80.11%, and for soft deposits 78.57%	Specific to RGB photos; moderate accuracy	DL models could support cost-effective screening for dental issues in large populations	Broader validation across different populations and environments needed

Abbreviations: AI, artificial intelligence; ANN, artificial neural network; AUC, area under the curve; CLAHE, contrast-limited adaptive histogram equalisation; CNN, convolutional neural network; DL, deep learning; ELM, extreme learning machine; GLCM, grey-level co-occurrence matrix; MGLCM, multichannel grey-level co-occurrence matrix; NBC, naïve Bayes classifier; PSONN, particle swarm optimization neural network; R-CNN, region-based convolutional neural network; ResNet-50, residual network with 50 layers; RGB, red, green, blue (colour model used for digital images); WN + SVM, weighted network plus support vector machine; YOLOv5x, you only look once version 5x (a type of object detection algorithm).

TABLE 5 | Emerging applications of digital technologies using oral fluid samples for periodontal screening and diagnosis in the dental setting.

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
da Silva et al. (2024)	Evaluate infrared spectroscopy for screening diabetes and periodontitis	Fourier Transform Infrared Spectroscopy (FT-IR)	80 patients, saliva samples analysed with FT-IR and ML ($n=80$)	True positive rates 78%–93% across groups, classification accuracy over 80% for periodontitis and diabetes	Small sample size; needs further validation and data analysis method refinement	FT-IR combined with ML could be a non-invasive, rapid screening tool in clinical settings	Larger studies to validate findings and refine data analysis methods
Deng et al. (2023)	Develop ML tool for screening periodontal health using salivary biomarkers	ML	Cross-sectional study with 408 subjects, using ML to classify periodontal status	High accuracy (AUC > 0.94) in discriminating between health, gingivitis and periodontitis	Limited sample size; requires external validation	ML-based screening tools using salivary biomarkers could support non-invasive periodontal health assessment	Larger, externally validated studies needed to confirm efficacy and generalisability
Huang et al. (2020)	Develop a periodontal disease antibody array for predicting severe periodontitis	Antibody Array and ML	Gingival crevicular fluid (GCF) samples from 50 patients (25 healthy, 25 severe periodontitis)	ROC analysis showed AUC of 0.984 for IL-1B; linear discriminant analysis had the highest classification accuracy	Limited sample size, specific to severe periodontitis	Antibody arrays combined with ML can potentially diagnose and predict periodontitis severity	Larger, more diverse datasets needed for validation
Kim et al. (2020)	Assess ML models for predicting periodontitis severity using salivary bacteria	ML models	692 subjects (144 healthy, 548 generalised periodontitis); ML models trained on genomic data	Accuracy 93% for classifying 'healthy' vs. 'moderate/severe' periodontitis, AUC 0.96	Limited validation with external samples; specificity lower than sensitivity	ML models could potentially identify periodontitis severity, supporting early intervention strategies	External validation with more diverse and larger sample sets needed

(Continues)

TABLE 5 | (Continued)

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Shen et al. (2022)	Evaluate AI-assisted dental monitoring in periodontitis treatment	AI-based Dental Monitoring	Randomised controlled trial with 53 patients; three groups: AI ($n = 16$), AI + human counselling ($n = 17$), control ($n = 20$)	AI and AI + human counselling groups showed greater improvement in PPD, CAL and plaque index over CG	Small sample size, limited follow-up period	AI-based monitoring can enhance treatment outcomes when combined with human intervention	Larger trials with longer follow-up periods to validate and refine AI-assisted monitoring interventions

Abbreviations: AI, artificial intelligence; AUC, area under the curve; CAL, clinical attachment loss; CG, control group; FT-IR, Fourier transform infrared spectroscopy; GCF, gingival crevicular fluid; IL-1B, interleukin-1 beta; ML, machine learning; PPD, probing pocket depth; ROC, receiver operating characteristic.

single (Lee et al. 2024, 2023; Patel et al. 2023) or multiple (Krois et al. 2019; Santamaria et al. 2024; Schwendicke et al. 2021; Troiano et al. 2023) cohorts of patients with longitudinal follow-up or prospective trials (Nagarajan et al. 2019).

In seven articles, the study population consisted of periodontitis patients who were followed up after active periodontal therapy, either within an SPC programme (Schwendicke et al. 2021; Krois et al. 2019; Troiano et al. 2023; Santamaria et al. 2024) or without explicit mention of SPC (Lee et al. 2023, 2024; Nagarajan et al. 2019), and the label consisted of an outcome/event assessed at the patient level (i.e., yearly rate of tooth loss for any reason; number of teeth lost for any reason; loss of at least one tooth for any reason; or periodontitis progression, defined as the combination of CAL loss ≥ 2 mm and BoP in ≥ 1 site) or at the tooth level (i.e., tooth loss for any reason; or tooth/molar loss due to periodontitis). In one article (Patel et al. 2023), information on the baseline periodontal conditions of the patient sample was not available, and the model was trained to predict periodontal status at the end of the prediction period.

In all studies, the index test consisted of one or more prognostic models generated by supervised ML after training with a number of predictors ranging from 6 (Santamaria et al. 2024; Schwendicke et al. 2021) to 74 (Patel et al. 2023).

Supervised ML models were compared for their prognostic performance with reference tests including traditional periodontitis risk assessment methods (Patel et al. 2023; Santamaria et al. 2024), traditional statistical models (e.g., multilevel, multivariate models) (Lee et al. 2024, 2023), different ML models either alone or aggregated (Nagarajan et al. 2019; Schwendicke et al. 2021), or the same model with different complexity, sample size, prediction period or validation scheme (Krois et al. 2019).

The mean length of follow-up was 18.2 ± 5.6 years and 6.6 ± 2.9 years for the two cohorts of the study by Krois et al. (2019), 6.7 ± 3.0 years, 9.4 ± 2.0 years, 10.1 ± 0.5 years and 18.2 ± 5.5 years for the four cohorts contributing to the study by Schwendicke et al. (2021), 10 years (Santamaria et al. 2024; Troiano et al. 2023), 2.41 ± 2.17 years (Lee et al. 2023, 2024), 28 weeks (Nagarajan et al. 2019) and 5 years (Patel et al. 2023).

3.2.2 | Summary of Main Findings

The main findings of the included studies are summarised in Table 9 and reported in the following paragraphs.

3.2.2.1 | Prognostic Performance of Supervised ML Models Versus Traditional Periodontitis Risk Assessment Tools. In the study by Patel et al. (2023), the accuracy of predicting periodontal status (classified as 'healthy', 'mild periodontal disease' or 'severe periodontal disease') at the end of a 5-year prediction period was comparatively evaluated for supervised ML and four traditional periodontitis risk assessment methods, including two validated tools, that is, the Periodontal Risk Assessment (PRA; Lang and Tonetti 2003) and the PreViser (Page et al. 2002, 2003). The level of agreement between the risk outputs generated by each and the label was 70% for the supervised ML model,

TABLE 6 | Emerging applications of digital technologies using health records for periodontal screening and diagnosis in the dental setting.

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Tastan Eroglu et al. (2024)	Assess ChatGPT's ability to classify periodontitis using the 2018 classification	AI	200 untreated periodontitis patients, compared with standardised reference diagnoses	Stage accuracy 59.5%, grade accuracy 50.5%, extent accuracy 84%	ChatGPT's performance varies significantly, with moderate accuracy in extent classification	AI tools like ChatGPT might be useful for preliminary diagnosis but need improvement for full clinical utility	Enhancements required in AI training to improve staging and grading accuracy in periodontal classification

Abbreviation: AI, artificial intelligence.

followed by PreViser (55%), PRA and Phillips (35%), and Cigna (25%). The Cohen's Kappa value between supervised ML and label was 0.6 (indicating moderate to high agreement), followed by PreViser (0.4), PRA (0.3), Phillips (0.3) and Cigna (0.2, no to low agreement).

At the tooth level, two ML models (i.e., logistic regression, LR; and neural network, NN) were compared with a traditional risk assessment method based on 6 clinical and radiographic parameters (Nibali et al. 2017) in the prediction of tooth loss due to periodontitis over a 10-year period (Santamaria et al. 2024). For the traditional risk assessment method, high specificity (99.96%), moderate positive predictive value (PPV = 50.0%) and very low sensitivity (0.85%) were reported. For both AI models, moderate specificity (NN = 52.26%, LR = 67.59%), high sensitivity (NN = 98.29%, LR = 91.45%) and high PPV (NN = 89.1%, LR = 88.6%) were found. Overall, the results of the study suggest that the two investigated AI-based models may perform at least as well as traditional risk assessment tools in predicting tooth loss due to periodontitis (Santamaria et al. 2024).

3.2.2.2 | Prognostic Performance of Supervised ML Models Versus Traditional Statistical Models. In the study by Lee et al. (2023), the ML model and a classical statistical model (i.e., count regression) achieved an average root mean square error (RMSE) of 2.71 and 3.88, respectively, when applied to the prediction of tooth loss count, indicating that the former performed slightly better than the latter.

The study by Lee et al. (2024) compared the ML model and a classical statistical model (i.e., logistic regression model) for their accuracy in predicting the binary tooth loss phenotype. On average, the ML models achieved an AUROC of 0.71 and an AUPRC of 0.66 (sensitivity: 0.67, specificity: 0.63), whereas the final model of the classical statistical approach achieved an AUROC of 0.70 and an AUPRC of 0.64. The sensitivity and specificity associated with the optimal cut-off were 0.65 and 0.67, respectively.

3.2.2.3 | Comparative Evaluation of the Prognostic Performance of Different Supervised ML Models. The study by Krois et al. (2019) modulated some parameters (i.e., model complexity, sample size used as training dataset, prediction period and validation scheme) and evaluated their impact on the prognostic accuracy of each of four ML models. While some parameters (i.e., model complexity and prediction period) had a limited impact on model outcome, all models showed a tendency to lose accuracy (to a similar extent across models) when trained on smaller datasets. Interestingly, no model showed higher accuracy than the no-information rate, suggesting that despite high accuracy, none of the models developed would be useful in a clinical setting (Krois et al. 2019). Also, in the study by Schwendicke et al. (2021), the centre used to test the ML model was found to influence predictive performance, with the available predictors being used differently (in some cases even in opposite directions) and the models often performing better when tested in different centres than when tested on data from their own centre. These findings limit the ability to generalise the applicability of the ML models developed.

Four studies reported the results of a comparative evaluation of the prognostic accuracy of two or more different ML models using a patient-based (Nagarajan et al. 2019; Schwendicke

TABLE 7 | Emerging applications of digital technologies using clinical data for periodontal screening and diagnosis in the dental setting.

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Alqahtani et al. (2022)	Identify factors associated with periodontitis using ML models	ML	Cross-sectional study with 4555 participants from NHANES dataset	Age, education, alcohol use, hypertension and diabetes were significant predictors	Study limited to US adults; may not generalise to other populations	ML models can assist in identifying risk factors for periodontitis, aiding in targeted prevention strategies	Need for validation across different populations and inclusion of additional predictors
Bashir et al. (2022)	Compare ML algorithms for developing periodontitis prediction models	ML algorithms	National survey data from Taiwan ($n = 3453$) and the United States ($n = 3685$); ML models compared	Internal validation AUC > 0.95, accuracy $> 95\%$; external validation showed significant performance drop	Significant drop in performance during external validation; need for more complex predictors	ML models show promise but require larger sample sizes and more complex predictors for accurate periodontitis prediction	Larger, more diverse sample sizes and additional predictors necessary for reliable ML model performance
Farhadian et al. (2020)	Design an SVM-based decision support system for periodontal disease diagnosis	Support Vector Machine (SVM)	300 patients categorised into gingivitis, localised, or generalised periodontitis	Classification accuracy 88.7%, hypervolume under the manifold (HUM) 0.912	Small sample size; lack of external validation	SVM models could enhance diagnostic accuracy for periodontal disease in clinical settings	External validation and testing on more varied patient populations necessary
Lakshmi and Dheeba (2023)	Predict periodontal disease severity using SVM and feature selection	Support Vector Machine (SVM)	300 periodontal cases ($n = 300$) analysed with SVM and feature selection techniques	SVM achieved accuracy of 88.7%; feature selection improved classification performance	Requires validation with larger, more diverse datasets	SVM could assist in accurate early diagnosis and severity prediction of periodontal diseases	Broader validation and testing needed across different populations

(Continues)

TABLE 7 | (Continued)

First author	Objective	Digital technology	Methods (sample size)	Results	Limitations	Implications for practice	Future research directions
Papantonopoulos et al. (2014)	Evaluate the use of ANNs for diagnosing aggressive periodontitis	Artificial Neural Networks (ANNs)	Clinical data from four samples, totalling 347 patients with AgP and CP; cross-validation with ANN	ANN achieved 90%–98% accuracy in classifying AgP vs. CP based on immunologic parameters	Small sample size; lack of external validation	ANN could be used to support clinical decision-making for distinguishing between AgP and CP	Larger and more diverse sample sizes needed for validation
Snider et al. (2024)	Evaluate AI-driven remote monitoring in orthodontic treatment	AI	24 patients monitored over 232 clinical time points	Sensitivity 0.53–0.22, specificity 0.94–0.99, accuracy 0.60–0.72	Low sensitivity for detecting gingivitis and recession, moderate accuracy	AI monitoring can enhance treatment by identifying oral hygiene and mucogingival conditions remotely	Further research needed to improve sensitivity and validate in larger populations

Abbreviations: AgP, aggressive periodontitis; AI, artificial intelligence; AUC, area under the curve; CP, chronic periodontitis; HUM, hypervolume under the manifold; ML, machine learning; NHANES, national health and nutrition examination survey; NNs, artificial Neural networks; SVM, support vector machine.

et al. 2021) or tooth-based (Troiano et al. 2023; Santamaria et al. 2024) label. Three studies reported similar performance for the tested models in predicting periodontitis progression and non-progression (Nagarajan et al. 2019), the number of teeth lost for any reason (Schwendicke et al. 2021), and tooth loss due to periodontitis (Santamaria et al. 2024). In the study by (Troiano et al. 2023), model performance (assessed by AUC) was better for the neural network (AUC = 0.702) compared with the random forest (AUC = 0.683), naïve Bayes (AUC = 0.649), logistic regression (AUC = 0.611), K-nearest neighbour (AUC = 0.565), gradient boosting (AUC = 0.527) and support vector machine (AUC = 0.512) for predicting molar loss due to periodontitis.

Two studies included an evaluation of multiple classification algorithms aggregated within an ensemble predictive modelling framework (Nagarajan et al. 2019; Troiano et al. 2023). In the study by Nagarajan et al. (2019), the sensitivity of the classification approaches within the ensemble framework was on average better than that of single classifier systems for predicting periodontitis progression/non-progression, with the ensemble naïve Bayes classifier having the highest average sensitivity (~89%) with a specificity of ~46%. In the study by (Troiano et al. 2023), the average performance of the ensemble model was equal to that of the neural network (AUC = 0.702), followed by the other ML models investigated. The performance of both the ensemble and neural network models was most stable when applied to the three different cohorts used for external validation, with their performance never falling below AUC = 0.70 (Troiano et al. 2023).

3.2.3 | Reporting Standards and Risk of Bias

Four studies (Krois et al. 2019; Troiano et al. 2023; Schwendicke et al. 2021; Santamaria et al. 2024) explicitly reported their adherence to reporting standards of the TRIPOD (Collins et al. 2015), two studies (Lee et al. 2023, 2024) adhered to the reporting standards of the STROBE (von Elm et al. 2008), while no explicit mention of reporting standards was found in two studies (Nagarajan et al. 2019; Patel et al. 2023).

Based on PROBAST (Moons et al. 2019), all included studies were assigned an overall 'high risk' of bias and 'high concern' related to the applicability of their findings (Table 10). Due to the retrospective nature of the majority of studies, which often led to a biased inclusion of periodontitis patients complying partially or completely with a SPC program, participant selection was one of the most penalised PROBAST domains. Also, in several studies, the low frequency of events (e.g., tooth loss during SPC) in relation to the high number of variables included in the prognostic model would have often required a larger sample size (Table 10, Appendix S1).

4 | Discussion

4.1 | Emerging Applications of Digital Technologies for Periodontal Screening, Diagnosis and Prognosis: Key Findings and Strengths

Based on 40 studies, the scoping review presented in part I described a group of AI-driven technologies and advanced imaging techniques for periodontal screening and/or diagnosis.

TABLE 8 | Characteristics of the included studies on AI-based models for the estimation of the risk of tooth loss and change in periodontal disease condition in the dental setting.

Study (author, year)	Study design	Population for model training and testing (country)	Label (statistical unit)	Follow-up and/or prediction period	Intervention (no. and type of predictors included in the model)	Intervention (i.e., the index test based on AI)	Comparison (i.e., the reference test)	Performance measures
Krois et al. (2019)	Retrospective analysis of data from 2 independent patient cohorts with longitudinal follow-up	11,651 teeth in 627 periodontitis patients under SPC (Europe)	Tooth loss for any reason during SPC (Tooth)	Kiel cohort: 18.2 ± 5.6 years Greifswald cohort: 6.6 ± 2.9 years Prediction period: 10, 15, or 20 years or uncensored	10 predictors extrapolated from patient demographics, medical and dental history, periodontal charting and radiographs.	4 binary classification models with different complexity: • logistic regression • recursive partitioning • random forest • extreme gradient boosting	Any of the AI-based models listed as 'Intervention' (but modulated in terms of model complexity, sample size, prediction periods, and training and validation strategies)	Calculated for all models in 6 validation scenarios • AUC • Specificity • Sensitivity • Accuracy • No-information rate
Nagarajan et al. (2019)	Prospective trial	114 patients with moderate/severe periodontitis receiving only oral hygiene instruction ($n = 73$) or conventional quadrant-wise full-mouth scaling and root planing ($n = 41$) (USA)	≥ 1 site with CAL loss ≥ 2 mm and BoP over experimental period (patient)	28 weeks	27 predictors extrapolated from biomarkers in serum, saliva and subgingival plaque.	3 single ML classifier systems: • •Naïve Bayes (NBC) • Support vector machine with linear kernel (SVM) • Linear Discriminant Analysis (LDA) Ensemble predictive modelling framework: • NBC^E, SVM^E, LDA^E	Any of the AI-based models listed as 'Intervention'	• Sensitivity • Specificity • Positive predictive value • Negative predictive value
Schwendicke et al. (2021)	Retrospective analysis of data from 4 independent patient cohorts with longitudinal follow-up	897 periodontitis patients under SPC (Europe)	Tooth loss rate per year of SPC (patient)	Duration of SPC: 10 ± 1 years	6 predictors extrapolated from patient demographics, medical and dental history, and periodontal charting.	• elastic net regression model • stochastic gradient boosting model • multivariable linear regression model	Any of the AI-based models listed as 'Intervention' (but tested in different patient cohorts)	10-fold cross-validation of the root-mean-squared-error (RMSE), that is, the mean deviation in error in predicting annual tooth loss per patient

(Continues)

TABLE 8 | (Continued)

Study (author, year)	Study design	Population for model training and testing (country)	Label (statistical unit)	Follow-up and/or prediction period	Intervention (no. and type of predictors included in the model)	Intervention (i.e., the index test based on AI)	Comparison (i.e., the reference test)	Performance measures
Lee et al. (2023)	Retrospective analysis of data from a patient cohort with longitudinal follow-up	7840 periodontitis patients who completed an initial therapeutic phase based on non-surgical periodontal therapy and were followed up for at least 6 months (USA)	No. of teeth—excluding third molars- lost after initial visit (patient)	Mean follow-up: 2.41 ± 2.17 years	18 predictors extrapolated from patient demographics, medical and dental history, and periodontal charting.	Rule- Fit algorithm followed by count regression model	Classical statistics (count regression model)	Rule-fit algorithm Root-mean-squared error (RMSE), (i.e., the mean deviation in error in predicting tooth loss count per patient) Regression analysis *p*-value
Patel et al. (2023)	Retrospective analysis of data from a patient cohort with longitudinal follow-up	Training: 27,138 patients Testing: 20 dental patients with heterogeneous baseline periodontal conditions (USA)	Change in periodontal status (among 'healthy', 'mild periodontal disease', and 'severe periodontal disease') (patient)	5 years	74 candidate predictors extrapolated from patient demographics, medical and dental history, periodontal charting, clinical notes, social determinants of health, and radiographs.	Extreme gradient boosting	Clinical prognostic systems (PreViser, PRA, Sonicare, Cigna)	The AI-based Intervention provided the most accurate prediction (70%), followed by Previser (55%), PRA (35%), Phillips (35%) and Cigna (25%)
Troiano et al. (2023)	Retrospective analysis of data from patient cohorts with longitudinal follow-up	3157 M in 515 periodontitis patients under SPC (Europe, USA, China)	First/s molar loss for any reasons (during follow-up) First/s molar loss due to periodontitis (during follow-up) (tooth)	10 years	9 predictors extrapolated from patient demographics, medical and dental history, prosthetic aspects, periodontal charting and radiographs.	• Logistic regression • Support vector machine • K-nearest neighbours • Random forest • Neural network • Gradient boosting • Naïve Bayes • Combinations of the above by means of the 'ensembled stacking method'	Any of the AI-based models listed as 'Intervention'	Calculated for each predictive model (including the ensembled model) and each outcome • AUROC • Sensitivity • Specificity • Positive predictive value (PPV) • Negative predictive value (NPV) • Harmonic mean for sensitivity and specificity

(Continues)

TABLE 8 | (Continued)

Study (author, year)	Study design	Population for model training and testing (country)	Label (statistical unit)	Follow-up and/or prediction period	Intervention (no. and type of predictors included in the model)	Intervention (i.e., the index test based on AI)	Comparison (i.e., the reference test)	Performance measures
Lee et al. (2024)	Retrospective analysis of data from a patient cohort with longitudinal follow-up	7840 periodontitis patients who completed an initial therapeutic phase based on non-surgical periodontal therapy and were followed up for at least 6 months (USA)	Loss of ≥ 1 tooth (excluding third molars) after initial visit (patient)	Mean follow-up: 2.41 ± 2.17 years	19 predictors extrapolated from patient demographics, medical and dental history, periodontal charting, and non-periodontal aspects (i.e., presence of caries).	Rule-Fit algorithm followed by logistic regression model	Classical statistics (logistic regression model)	Rule-Fit algorithm and classical statistical approaches • AUROC • AUPRC • Sensitivity • Specificity
Santamaria et al. (2024)	Retrospective analysis of data from patient cohorts with longitudinal follow-up	4418 teeth in 185 periodontitis patients under SPC (Europe, USA)	Tooth loss due to periodontitis during follow-up (tooth)	10 years	6 predictors extrapolated from periodontal charting and radiographs. In the absence of radiographs, tooth restorability was evaluated.	• Logistic regression • Neural network	• Any of the AI-based models listed as 'Intervention' • Clinical prognostic system (Nibali et al. 2017)	Area under the cure of receiver operating characteristic (AUC ROC) curve was used to assess discrimination; in addition, predicted probabilities were calculated from the model and applied in the ROC analysis to assess the best threshold probability and then calculated sensitivity, specificity, positive predictive values (PPVs), negative predictive values (NPVs) and accuracy for both the LR and NN models. Calibration was assessed by building calibration plots and by calculating calibration performance metrics (calibration-in-the-large and Brier score). Specificity, sensitivity, PPVs, NPVs and accuracy were estimated on the basis of the predicted probability from each category of the clinical prognostic system using contingency table built with the hypothesis 'more likely to be retained' for 'good' and 'fair' prognosis and hypothesis 'more likely to be lost' for the 'questionable' and 'unfavourable' categories

Abbreviations: AI, artificial intelligence; BoP, bleeding on probing; CAL, clinical attachment level; ML, machine learning; SPC, supportive periodontal care.

TABLE 9 | Main findings of the included studies on supervised ML models for the estimation of the risk of tooth loss and change in periodontal disease condition in the dental setting.

		Prognostic efficacy of the Intervention (i.e., the index test based on AI) when evaluated in relation to the following comparison (i.e., the reference test)				
Label	**Study (author, year)**	**Alternative 'Intervention' generated by modulating the AI-based model in terms of model complexity, sample size, prediction periods, training and validation strategies**	**Alternative 'Intervention' resulting from testing the AI-based model in a different patient cohort**	**Other AI-based models**	**Traditional statistical methods**	**Traditional risk assessment tools**
Tooth loss (for any reason) in periodontitis patients	Krois et al. (2019)	– Model complexity and prediction period had a limited impact on model outcome – Models tended to lose accuracy when trained on smaller data sets – No model showed higher accuracy than the no-information rate (thus indicating that none of the developed models would be useful in a clinical setting, despite high accuracy)				
	Schwendicke et al. (2021)		– No model showed useful predictive values – The centre/cohort used for testing the model had an impact on the accuracy of predictions, indicating low generalisability			

(Continues)

TABLE 9 | (Continued)

Label	Study (author, year)	Prognostic efficacy of the Intervention (i.e., the index test based on AI) when evaluated in relation to the following comparison (i.e., the reference test)				
		Alternative 'Intervention' generated by modulating the AI-based model in terms of model complexity, sample size, prediction periods, training and validation strategies	Alternative 'Intervention' resulting from testing the AI-based model in a different patient cohort	Other AI-based models	Traditional statistical methods	Traditional risk assessment tools
	Lee et al. (2023)				– The ML model showed a slightly better predictive performance than the classical statistical model (i.e., count regression) – Both ML and classical statistical models did not achieve excellent performance	
	Lee et al. (2024)				– The ML model showed a slightly better predictive performance than the classical statistical model (i.e. logistic regression) – Both ML and classical statistical models did not achieve excellent performance	
	Troiano et al. (2023)			– The ensembled model combining neural network and logistic regression consistently ranked the highest (or almost) for predictive performance in the three validation cohorts		

(Continues)

TABLE 9 | (Continued)

Label	Study (author, year)	Prognostic efficacy of the Intervention (i.e., the index test based on AI) when evaluated in relation to the following comparison (i.e., the reference test)				
		Alternative 'Intervention' generated by modulating the AI-based model in terms of model complexity, sample size, prediction periods, training and validation strategies	Alternative 'Intervention' resulting from testing the AI-based model in a different patient cohort	Other AI-based models	Traditional statistical methods	Traditional risk assessment tools
Tooth loss (due to periodontitis)	Troiano et al. (2023)			– The ensembled model combining neural network and logistic regression consistently ranked the highest (or almost) for predictive performance in the three validation cohorts		
	Santamaria et al. (2024)			– In the external validation cohort, logistic regression model and neural network model performed similarly in terms of positive predictive value, negative predictive value and accuracy		– The two investigated AI-based models showed a similar overall predictive performance compared with the traditional risk assessment tool
Disease progression (CAL loss)	Nagarajan et al. (2019)			– The sensitivity of the classification approaches within the ensemble frameworks was on average better than those of single classifier systems, with NBC exhibiting the highest average sensitivity (~89%)		

(Continues)

TABLE 9 | (Continued)

Label	Study (author, year)	Prognostic efficacy of the Intervention (i.e., the index test based on AI) when evaluated in relation to the following comparison (i.e., the reference test)				
		Alternative 'Intervention' generated by modulating the AI-based model in terms of model complexity, sample size, prediction periods, training and validation strategies	Alternative 'Intervention' resulting from testing the AI-based model in a different patient cohort	Other AI-based models	Traditional statistical methods	Traditional risk assessment tools
Periodontal condition at the end of the prediction period	Patel et al. (2023)					The level of agreement between the risk outputs that were generated by the investigated tools and the label was 70% for the ML model, followed by PreViser (55%), PRA and Phillips (35%) and Cigna (25%). The Cohen's Kappa value between ML model and the label was 0.6 (indicating moderate to high agreement), followed by Previser (0.4, low to moderate agreement), PRA 0.3 (low agreement), Phillips (0.3, low agreement) and Cigna (0.2, no to low agreement)

Note: Studies are grouped by label, and their findings are reported for each comparison (i.e., the reference test) used within each study.
Abbreviations: AI, artificial intelligence; CAL, clinical attachment level; ML, machine learning; NBC, naïve Bayes classifier; RMSE, root-mean-squared error.

TABLE 10 | Risk of bias (as evaluated according to PROBAST; Moons et al. 2019) for included studies on emerging applications of digital technologies for periodontal prognosis in the dental setting.

Author (year)	Risk of bias				Applicability			Overall	
	1. Participants	2. Predictors	3. Outcome	4. Analysis	1. Participants	2. Predictors	3. Outcome	Risk of bias	Applicability
Krois et al. (2019)	–	+	?	+	–	+	+	–	–
Nagarajan et al. (2019)	–	+	–	–	+	+	–	–	–
Schwendicke et al. (2021)	–	+	?	–	–	+	+	–	–
Troiano et al. (2023)	–	+	+	+	–	+	+	–	–
Lee et al. (2023)	–	+	+	+	–	+	+	–	–
Patel et al. (2023)	–	?	–	–	–	–	–	–	–
Lee et al. (2024)	–	+	+	+	–	+	+	–	–
Santamaria et al. (2024)	–	+	+	+	–	+	+	–	–

The accuracy of such technologies was tested on different types of data sources among radiographs, photographs, oral fluid samples, electronic health records and clinical data. Bidimensional radiographs were the most used data source, with AI methods showing high accuracy in diagnosing periodontitis and alveolar bone loss. Also, micro-CT and CBCT combined with AI improved the detection of bone loss and other periodontal conditions. Photographs, particularly intraoral images, were also frequently studied, with AI models showing promise for non-invasive detection of periodontal disease, particularly gingival inflammation. Data on the combination of AI and clinical photographs suggest that this strategy may soon surpass traditional visual assessment of gingival inflammation in both speed and accuracy. Oral fluid samples were analysed using advanced techniques such as FT-IR and ML models, achieving high accuracy in the discrimination of disease severity. Electronic health records and clinical data were used in combination with AI models, achieving moderate to high classification accuracy. Overall, these results highlight a trend towards the integration of AI and advanced imaging techniques in periodontal diagnostics, offering improved accuracy and non-invasive options for both screening and diagnosis.

The systematic review presented in part II identified 8 studies evaluating the prognostic performance models generated by supervised ML, a potent tool for building prediction models (Moons et al. 2019). The models were fed with a varying number (between 6 and 74) and types (e.g., age, sex, GCF and serum biomarkers, clinical and radiographic parameters, systemic conditions/diseases) of predictors that are all available at the intended moment of prediction, thus making the investigated models candidates for the real-world setting (Moons et al. 2019). Diverse accuracy in predicting tooth loss and change in periodontal condition was reported, with some models showing robust discrimination and calibration index. Interestingly, data from single, comparative studies indicated that supervised ML can be used to generate prognostic models either (i) able to predict the patient's periodontal status over a 5-year prediction period with a higher accuracy compared with validated, traditional risk assessment tools such as the PRA and the PreViser, or (ii) characterised by a better predictive performance than classical statistical models in predicting tooth loss (evaluated as either binary variable or count) in periodontitis patients.

4.2 | Periodontal Screening and Diagnosis With Digital Technologies in the Dental Setting: Current Limitations

Despite their promising results, the retrieved studies also highlight some limitations of the digital technologies that were evaluated for periodontal screening and diagnosis in the dental setting.

On bidimensional radiographs, AI-driven models may show reduced diagnostic accuracy in certain areas (e.g., severity classification of anterior bone loss) and need broader validation across different imaging modalities and patient populations (Alotaibi et al. 2022; Guler Ayyildiz et al. 2024). Although the enhancement of tridimensional radiographs (e.g., CBCT) with AI and advanced imaging showed promising diagnostic metrics, several studies were conducted on small patient samples and used specific imaging conditions. Also, variability in performance across different patient demographics was observed, thus limiting consistency across different clinical settings.

Variability in performance across different datasets, the specificity of certain models to certain photographic angles (Chau et al. 2023; Li et al. 2021) and the limited performance in detecting complex conditions such as gingival overgrowth (Kurt-Bayrakdar et al. 2023) point to areas where the combination of AI and clinical photographs still faces challenges. Further development to improve accuracy and validate the results on diverse and larger datasets will be needed to match the diagnostic skills of experienced clinicians.

Studies on the application of AI on oral fluids highlighted limitations such as small sample sizes and the need for external validation (Huang et al. 2020; Kim et al. 2020). The specificity and generalisability of these models remain areas of concern, particularly when applied to different populations and clinical settings. For example, while ML models have shown high accuracy in specific cohorts, their performance can vary significantly when tested on external samples, highlighting the need for larger and more diverse studies to confirm their efficacy. Such strategies are promising not only in diagnosis, but also in ongoing patient management (Shen et al. 2022) through a continuous, real-time monitoring of periodontal health, but these benefits will depend on overcoming current limitations related to follow-up periods and sample sizes, which are often insufficient to fully validate the long-term efficacy of these tools.

Electronic health records provide a rich source of patient data, including demographic information, clinical history and radiographic findings, which can be used to develop predictive models for periodontal disease management. The accuracy of AI models to be applied on electronic health records can be significantly influenced by the quality and completeness of radiology data, as well as the need for external validation in different clinical settings. This is particularly important given the variability in data quality across healthcare settings, which may affect the generalisability of these models. Furthermore, while ChatGPT demonstrated moderate accuracy in classifying the extent of periodontitis (84%), its performance in staging and grading was less reliable (Tastan Eroglu et al. 2024), suggesting that while AI can assist in preliminary assessments, significant improvements are needed before these tools can be fully integrated into clinical practice.

The use of ML models based on clinical data is in line with the broader trend towards personalised medicine; however, the generalisability of these results is often limited, as most studies highlight the challenges of external validation. The decline in performance during external validation highlights the need for more complex models and larger, more diverse datasets to ensure that these tools can be reliably applied to different populations. Support vector machines and artificial neural networks are also being used to improve diagnostic accuracy (Farhadian et al. 2020; Papantonopoulos et al. 2014) but, similarly to ML models, these approaches require further validation on more diverse and larger datasets to confirm their effectiveness in real-world scenarios.

4.3 | Assessment of Periodontal Prognosis With Digital Technologies in the Dental Setting: Current Limitations

Among studies included in the present review, some failed to generate supervised ML models with higher accuracy than the no-information rate, thus indicating that these models cannot be usefully applied to a clinical setting (Krois et al. 2019; Schwendicke et al. 2021), while some reported very low performance metrics for the tested models. For the other studies suggesting a potential translation of the tested models to clinical practice, a major limitation consisted of the lack of a discrimination between 'black-box' and transparent algorithms. Explainable AI makes clinical decision-making explainable for the practitioner and, consequently, for the patient. For example, diagnosis based on a black-box model or prognosis based on metadata to predict risk is not compatible with evidence-based dentistry, and therefore with better quality of care. Only two studies used supervised ML to capture the probability for a patient to change his/her periodontal condition in the future, while most studies focused on tooth loss as expressed at either patient- or tooth-level. However, it must also be considered that tooth loss is partly dependent on clinical decisions, and limited information on clinical decisions leading to tooth extraction could be retrieved in the included studies. Therefore, it is possible that some teeth with poor prognosis were extracted before the study, with an impact on tooth loss count as observed during the follow-up period.

4.4 | Limitations of the Present Review

A first, general limitation of the present review comes from the fact that it focuses mainly on the performance of digital technologies in disease detection and prediction, which may underestimate the practical challenges of integrating these technologies into everyday clinical practice. In this respect, some important elements such as costs, training, data protection and interpretation, or the patient perspective were all considered beyond the scope of the present review. In addition, while the review identifies emerging trends, it may not fully capture the rapidly evolving nature of AI and digital technologies, meaning that some of the findings could quickly become outdated as newer studies emerge.

Additional specific limitations related to the quantity and quality of the collected evidence must also be acknowledged. Firstly, the reliance on a relatively small number of studies in certain categories, as well as the variability in study designs and methodologies, makes it difficult to directly compare the effectiveness/efficacy of different technologies. This led to a qualitative rather than quantitative synthesis of the results and limits the generalisability of the findings. Some of the included studies had a small sample size, which may affect the robustness of the conclusions drawn. Together with other issues such as the retrospective nature of the majority of studies (which often led to a biased patient selection) and an unfavourably low frequency of events (e.g., tooth loss during SPC) in relation to the high number of variables incorporated into the models, the limited sample size contributed a high risk of bias and high concern related to the applicability of their findings for all included studies on emerging applications of digital technologies for the evaluation of periodontal prognosis. For this latter group of studies, additional limitations consisted of frequent incorporation bias (determined by the inclusion of predictors in the model that form part of the definition or assessment of the label that the model predicts; Moons et al. 2019), which may result in optimistic estimates of model performance; and lack of robust reference standards, since the traditional tools/methods that were used as control treatments are far from being considered the standard of care due to less than optimal prognostic performance and some persisting issues and limitations (Farina et al. 2023).

4.5 | Conclusions

When considered collectively, the results of the present review indicate that:

- Digital technologies for periodontal screening/diagnosis consist of AI and advanced imaging techniques as applied to different data sources, including radiographs, photographs, clinical data, oral fluid samples and electronic health records. Although all investigated options have limitations, including reduced diagnostic accuracy in certain areas of dentition and variability in performance across different datasets, the AI-driven analysis of 2D radiographs for periodontitis diagnosis and staging seems to be the option that is most closely approaching applicability to practice.
- Supervised ML is a promising tool for the generation of prognostic models to be applied in the field of Periodontology. Among the tested ML models, neural networks and ensemble predictive modelling frameworks resulting from the aggregation of multiple classification algorithms and tooth-related labels seem to be the scenario that is closer than others to translation into practice. However, the current evidence is not sufficiently consistent and, therefore, conclusive for clinical applicability and generalisability to practice.

Author Contributions

R.F., A.S., L.T. and C.A.R. conceptualised and designed the review. All Authors were involved in the literature search and/or article selection. R.F., A.S., L.T. and C.A.R. drafted the manuscript, which was reviewed and finalised by all authors after discussion with the Coordinators and Participants of Working Group 3, 20th European Workshop on Periodontology.

Acknowledgements

The Authors wish to thank Mrs. Alessandra Viola, University of Ferrara, Italy, for her support in the preparation of the figures.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

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

References

Alalharith, D. M., H. M. Alharthi, W. M. Alghamdi, et al. 2020. "A Deep Learning-Based Approach for the Detection of Early Signs of Gingivitis in Orthodontic Patients Using Faster Region-Based Convolutional Neural Networks." *International Journal of Environmental Research and Public Health* 17: 1–10. https://doi.org/10.3390/ijerph17228447.

Alotaibi, G., M. Awawdeh, F. F. Farook, M. Aljohani, R. M. Aldhafiri, and M. Aldhoayan. 2022. "Artificial Intelligence (AI) Diagnostic Tools: Utilizing a Convolutional Neural Network (CNN) to Assess Periodontal Bone Level Radiographically—A Retrospective Study." *BMC Oral Health* 22: 399. https://doi.org/10.1186/s12903-022-02436-3.

Alqahtani, H. M., S. M. Koroukian, K. Stange, N. K. Schiltz, and N. F. Bissada. 2022. "Identifying Factors Associated With Periodontal Disease Using Machine Learning." *Journal of International Society of Preventive and Community Dentistry* 12: 612–620. https://doi.org/10.4103/jispcd.JISPCD_188_22.

Arksey, H., and L. O'Malley. 2005. "Scoping Studies: Towards a Methodological Framework." *International Journal of Social Research Methodology* 8: 19–32. https://doi.org/10.1080/1364557032000119616.

Asimakopoulou, K., J. T. Newton, B. Daly, Y. Kutzer, and M. Ide. 2015. "The Effects of Providing Periodontal Disease Risk Information on Psychological Outcomes – A Randomized Controlled Trial." *Journal of Clinical Periodontology* 42: 350–355. https://doi.org/10.1111/jcpe.12377.

Bashir, N. Z., Z. Rahman, and S. L. S. Chen. 2022. "Systematic Comparison of Machine Learning Algorithms to Develop and Validate Predictive Models for Periodontitis." *Journal of Clinical Periodontology* 49: 958–969. https://doi.org/10.1111/jcpe.13692.

Bayrakdar, S. K., Ö. çelik, I. S. Bayrakdar, et al. 2020. "Success of Artificial Intelligence System in Determining Alveolar Bone Loss From Dental Panoramic Radiography Images." *Cumhuriyet Dental Journal* 23: 318–324. https://doi.org/10.7126/cumudj.777057.

Chang, H. J., S. J. Lee, T. H. Yong, et al. 2020. "Deep Learning Hybrid Method to Automatically Diagnose Periodontal Bone Loss and Stage Periodontitis." *Scientific Reports* 10: 7531. https://doi.org/10.1038/s41598-020-64509-z.

Chang, J., M. F. Chang, N. Angelov, et al. 2022. "Application of Deep Machine Learning for the Radiographic Diagnosis of Periodontitis." *Clinical Oral Investigations* 26: 6629–6637. https://doi.org/10.1007/s00784-022-04617-4.

Chapple, I. L., F. Van der Weijden, C. Doerfer, et al. 2015. "Primary Prevention of Periodontitis: Managing Gingivitis." *Journal of Clinical Periodontology* 42, no. Suppl 16: S71–S76. https://doi.org/10.1111/jcpe.12366.

Chau, R. C. W., G. H. Li, I. M. Tew, et al. 2023. "Accuracy of Artificial Intelligence-Based Photographic Detection of Gingivitis." *International Dental Journal* 73: 724–730. https://doi.org/10.1016/j.identj.2023.03.007.

Chen, Y., and X. Chen. 2020. "Gingivitis Identification via GLCM and Artificial Neural Network." In *Lecture Notes in Electrical Engineering*, edited by R. Su and H. Liu, 95–106. Springer.

Chiarito, M., L. Luceri, A. Oliva, G. Stefanini, and G. Condorelli. 2022. "Artificial Intelligence and Cardiovascular Risk Prediction: All That Glitters Is Not Gold." *European Cardiology Review* 17: e29. https://doi.org/10.15420/ecr.2022.11.

Chow, D. Y., J. R. H. Tay, and G. G. Nascimento. 2024. "Systematic Review of Prognosis Models in Predicting Tooth Loss in Periodontitis." *Journal of Dental Research* 103: 596–604. https://doi.org/10.1177/00220345241237448.

Collins, G. S., J. B. Reitsma, D. G. Altman, and K. G. M. Moons. 2015. "Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement." *BMJ* 350: g7594. https://doi.org/10.1136/bmj.g7594.

da Silva, S., C. L. Ferreira, J. M. B. Rizzato, et al. 2024. "Infrared Spectroscopy for Fast Screening of Diabetes and Periodontitis." *Photodiagnosis and Photodynamic Therapy* 46: 104106. https://doi.org/10.1016/j.pdpdt.2024.104106.

Dai, F., Q. Liu, Y. Guo, et al. 2024. "Convolutional Neural Networks Combined With Classification Algorithms for the Diagnosis of Periodontitis." *Oral Radiology* 40: 357–366. https://doi.org/10.1007/s11282-024-00739-5.

Deng, K., F. Zonta, H. Yang, G. Pelekos, and M. S. Tonetti. 2023. "Development of a Machine Learning Multiclass Screening Tool for Periodontal Health Status Based on Non-Clinical Parameters and Salivary Biomarkers." *Journal of Clinical Periodontology* 51, no. 12: 1547–1560. https://doi.org/10.1111/jcpe.13856.

Du, M., T. Bo, K. Kapellas, and M. A. Peres. 2018. "Prediction Models for the Incidence and Progression of Periodontitis: A Systematic Review." *Journal of Clinical Periodontology* 45: 1408–1420. https://doi.org/10.1111/jcpe.13037.

von Elm, E., D. G. Altman, M. Egger, et al. 2008. "The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies." *Journal of Clinical Epidemiology* 61: 344–349. https://doi.org/10.1016/j.jclinepi.2007.11.008.

Ertaş, K., I. Pence, M. S. Cesmeli, and Z. Y. Ay. 2022. "Determination of the Stage and Grade of Periodontitis According to the Current Classification of Periodontal and Peri-Implant Diseases and Conditions (2018) Using Machine Learning Algorithms." *Journal of Periodontal & Implant Science* 53, no. 1: 38–53. https://doi.org/10.5051/JPIS.2201060053.

Farhadian, M., P. Shokouhi, and P. Torkzaban. 2020. "A Decision Support System Based on Support Vector Machine for Diagnosis of Periodontal Disease." *BMC Research Notes* 13: 337. https://doi.org/10.1186/s13104-020-05180-5.

Farina, R., R. Lopez, A. Simonelli, and L. Trombelli. 2023. "Accuracy and Applicability of Periodontitis Risk Assessment Tools: A Critical Appraisal." *Periodontology 2000*: 1–18. https://doi.org/10.1111/prd.12498.

Farina, R., A. Simonelli, A. Baraldi, et al. 2021. "Tooth Loss in Complying and Non-Complying Periodontitis Patients With Different Periodontal Risk Levels During Supportive Periodontal Care." *Clinical Oral Investigations* 25: 5897–5906. https://doi.org/10.1007/s00784-021-03895-8.

Farina, R., A. Simonelli, M. E. Guarnelli, et al. 2024. "Efficacy of Communicating Periodontal Risk on Psychological Outcomes and Supragingival Plaque Control in Patients Undergoing First Periodontal Consultation: A Parallel-Arm, Randomized Trial." *Journal of Clinical Periodontology* 51, no. 10: 1289–1301. https://doi.org/10.1111/jcpe.14032.

Guler Ayyildiz, B., R. Karakis, B. Terzioglu, and D. Ozdemir. 2024. "Comparison of Deep Learning Methods for the Radiographic Detection of Patients With Different Periodontitis Stages." *Dento Maxillofacial Radiology* 53: 32–42. https://doi.org/10.1093/dmfr/twad003.

Harrell, F. E., K. L. Lee, and D. B. Mark. 2004. "Prognostic/Clinical Prediction Models: Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors." In *Tutorials in Biostatistics*, edited by R. B. D'Agostino, 1st ed., 223–249. Wiley.

Hong, H. H., C. C. Mei, H. L. Liu, et al. 2020. "The Correspondence of 3D Supporting Bone Loss and Crown-To-Root Ratio to Periodontitis Classification." *Journal of Clinical Periodontology* 47: 825–833. https://doi.org/10.1111/jcpe.13296.

Hoss, P., O. Meyer, U. C. Wölfle, et al. 2023. "Detection of Periodontal Bone Loss on Periapical Radiographs—A Diagnostic Study Using Different Convolutional Neural Networks." *Journal of Clinical Medicine* 12, no. 22: 7189. https://doi.org/10.3390/jcm12227189.

Huang, W., J. Wu, Y. Mao, et al. 2020. "Developing a Periodontal Disease Antibody Array for the Prediction of Severe Periodontal Disease Using Machine Learning Classifiers." *Journal of Periodontology* 91: 232–243. https://doi.org/10.1002/JPER.19-0173.

Jepsen, S., J. Blanco, W. Buchalla, et al. 2017. "Prevention and Control of Dental Caries and Periodontal Diseases at Individual and Population Level: Consensus Report of Group 3 of Joint EFP/ORCA Workshop on the Boundaries Between Caries and Periodontal Diseases." *Journal of Clinical Periodontology* 44: S85–S93. https://doi.org/10.1111/jcpe.12687.

Kabir, T., C. T. Lee, L. Chen, X. Jiang, and S. Shams. 2022. "A Comprehensive Artificial Intelligence Framework for Dental Diagnosis and Charting." *BMC Oral Health* 22: 480. https://doi.org/10.1186/s12903-022-02514-6.

Kazimierczak, W., R. Wajer, A. Wajer, et al. 2024. "Periapical Lesions in Panoramic Radiography and CBCT Imaging—Assessment of AI's Diagnostic Accuracy." *Journal of Clinical Medicine* 13, no. 9: 2709. https://doi.org/10.3390/jcm13092709.

Kim, E. H., S. Kim, H. J. Kim, et al. 2020. "Prediction of Chronic Periodontitis Severity Using Machine Learning Models Based on Salivary Bacterial Copy Number." *Frontiers in Cellular and Infection Microbiology* 10: 571515. https://doi.org/10.3389/fcimb.2020.571515.

Krois, J., C. Graetz, B. Holtfreter, P. Brinkmann, T. Kocher, and F. Schwendicke. 2019. "Evaluating Modeling and Validation Strategies for Tooth Loss." *Journal of Dental Research* 98: 1088–1095. https://doi.org/10.1177/0022034519864889.

Kurt-Bayrakdar, S., İ. Ş. Bayrakdar, M. B. Yavuz, et al. 2024. "Detection of Periodontal Bone Loss Patterns and Furcation Defects From Panoramic Radiographs Using Deep Learning Algorithm: A Retrospective Study." *BMC Oral Health* 24: 155. https://doi.org/10.1186/s12903-024-03896-5.

Kurt-Bayrakdar, S., M. Uğurlu, M. B. Yavuz, et al. 2023. "Detection of Tooth Numbering, Frenulum Attachment, Gingival Overgrowth, and Gingival Inflammation Signs on Dental Photographs Using Convolutional Neural Network Algorithms: A Retrospective Study." *Quintessence International* 54: 680–693. https://doi.org/10.3290/j.qi.b4157183.

Lakshmi, T. K., and J. Dheeba. 2023. "Predictive Analysis of Periodontal Disease Progression Using Machine Learning: Enhancing Oral Health Assessment and Treatment Planning." *International Journal of Intelligent Systems and Applications in Engineering* 11: 660–671.

Lang, N. P., J. E. Suvan, and M. S. Tonetti. 2015. "Risk Factor Assessment Tools for the Prevention of Periodontitis Progression a Systematic Review." *Journal of Clinical Periodontology* 42, no. Suppl 16: S59–S70. https://doi.org/10.1111/jcpe.12350.

Lang, N. P., and M. S. Tonetti. 2003. "Periodontal Risk Assessment (PRA) for Patients in Supportive Periodontal Therapy (SPT)." *Oral Health & Preventive Dentistry* 1: 7–16.

Lee, C. T., T. Kabir, J. Nelson, et al. 2022. "Use of the Deep Learning Approach to Measure Alveolar Bone Level." *Journal of Clinical Periodontology* 49: 260–269. https://doi.org/10.1111/jcpe.13574.

Lee, C.-T., K. Zhang, W. Li, et al. 2023. "Identifying Predictors of Tooth Loss Using a Rule-Based Machine Learning Approach: A Retrospective Study at University-Setting Clinics." *Journal of Periodontology* 94: 1231–1242. https://doi.org/10.1002/JPER.23-0030.

Lee, C.-T., K. Zhang, W. Li, et al. 2024. "Identifying Predictors of the Tooth Loss Phenotype in a Large Periodontitis Patient Cohort Using a Machine Learning Approach." *Journal of Dentistry* 144: 104921. https://doi.org/10.1016/j.jdent.2024.104921.

Li, H., J. Zhou, Y. Zhou, et al. 2020. "Automatic and Interpretable Model for Periodontitis Diagnosis in Panoramic Radiographs." In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, edited by A. L. Martel, P. Abolmaesumi, D. Stoyanov, et al., 454–463. Springer Science and Business Media Deutschland GmbH.

Li, W., Y. Chen, W. Su n, et al. 2019. "A Gingivitis Identification Method Based on Contrast-Limited Adaptive Histogram Equalization, Gray-Level Co-Occurrence Matrix, and Extreme Learning Machine." *International Journal of Imaging Systems and Technology* 29: 77–82. https://doi.org/10.1002/ima.22298.

Li, W., Y. Liang, X. Zhang, et al. 2021. "A Deep Learning Approach to Automatic Gingivitis Screening Based on Classification and Localization in RGB Photos." *Scientific Reports* 11: 16831. https://doi.org/10.1038/s41598-021-96091-3.

Liberati, A., D. G. Altman, J. Tetzlaff, et al. 2009. "The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Healthcare Interventions: Explanation and Elaboration." *BMJ (Clinical Research Ed.)* 339: b2700. https://doi.org/10.1136/bmj.b2700.

Liu, Q., F. Dai, H. Zhu, et al. 2023. "Deep Learning for the Early Identification of Periodontitis: A Retrospective, Multicentre Study." *Clinical Radiology* 78: e985–e992. https://doi.org/10.1016/j.crad.2023.08.017.

Manavella, V., F. Romano, F. Garrone, M. Terzini, C. Bignardi, and M. Aimetti. 2017. "A Novel Image Processing Technique for 3D Volumetric Analysis of Severely Resorbed Alveolar Sockets With CBCT." *Minerva Stomatologica* 66: 81–90. https://doi.org/10.23736/s0026-4970.17.04029-8.

Mohan, R., R. Mark, I. Sing, and A. Jain. 2014. "Diagnostic Accuracy of CBCT for Aggressive Periodontitis." *Journal of Clinical Imaging Science* 4: 2. https://doi.org/10.4103/2156-7514.133258.

Moher, D., A. Liberati, J. Tetzlaff, D. G. Altman, and P. Group. 2009. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement." *Annals of Internal Medicine* 151: 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.

Moons, K. G. M., R. F. Wolff, R. D. Riley, et al. 2019. "PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration." *Annals of Internal Medicine* 170: W1–W33. https://doi.org/10.7326/M18-1377.

Nagarajan, R., C. S. Miller, D. Dawson, and J. L. Ebersole. 2019. "Biologic Modelling of Periodontal Disease Progression." *Journal of Clinical Periodontology* 46: 160–169. https://doi.org/10.1111/jcpe.13064.

Nibali, L., C. Sun, A. Akcalı, X. Meng, Y. K. Tu, and N. Donos. 2017. "A Retrospective Study on Periodontal Disease Progression in Private Practice." *Journal of Clinical Periodontology* 44: 290–297. https://doi.org/10.1111/jcpe.12653.

Padmanabhan, S., A. Dommy, S. R. Guru, and A. Joseph. 2017. "Comparative Evaluation of Cone-Beam Computed Tomography Versus Direct Surgical Measurements in the Diagnosis of Mandibular Molar Furcation Involvement." *Contemporary Clinical Dentistry* 8: 439–445. https://doi.org/10.4103/ccd.ccd_515_17.

Page, R. C., E. A. Krall, J. Martin, L. Mancl, and R. I. Garcia. 2002. "Validity and Accuracy of a Risk Calculator in Predicting Periodontal Disease." *Journal of the American Dental Association (1939)* 133: 569–576. https://doi.org/10.14219/jada.archive.2002.0232.

Page, R. C., J. Martin, E. A. Krall, L. Mancl, and R. Garcia. 2003. "Longitudinal Validation of a Risk Calculator for Periodontal Disease." *Journal of Clinical Periodontology* 30: 819–827. https://doi.org/10.1034/j.1600-051X.2003.00370.x.

Papantonopoulos, G., K. Takahashi, T. Bountis, and B. G. Loos. 2014. "Artificial Neural Networks for the Diagnosis of Aggressive Periodontitis Trained by Immunologic Parameters." *PLoS One* 9: e89757. https://doi.org/10.1371/journal.pone.0089757.

Patel, J. S., K. Patel, H. Vo, et al. 2023. "Enhancing an AI-Empowered Periodontal CDSS and Comparing With Traditional Perio-Risk Assessment Tools." *AMIA Annual Symposium Proceedings* 29: 846–855. eCollection 2022.

Pitchika, V., M. Buttner, and F. Schwendicke. 2024. "Artificial Intelligence and Personalized Diagnostics in Periodontology: A Narrative Review." *Periodontology 2000* 95, no. 1: 220–231. https://doi.org/10.1111/prd.12586.

Qiao, J., S. Wang, J. Duan, et al. 2014. "The Accuracy of Cone-Beam Computed Tomography in Assessing Maxillary Molar Furcation Involvement." *Journal of Clinical Periodontology* 41: 269–274. https://doi.org/10.1111/jcpe.12150.

Ramseier, C. A. 2024. "Diagnostic Measures for Monitoring and Follow-Up in Periodontology and Implant Dentistry." *Periodontology 2000* 95, no. 1: 129–155. https://doi.org/10.1111/prd.12588.

Santamaria, P., G. Troiano, M. Serroni, T. G. Araùjo, A. Ravidà, and L. Nibali. 2024. "Exploring the Accuracy of Tooth Loss Prediction Between a Clinical Periodontal Prognostic System and a Machine Learning Prognostic Model." *Journal of Clinical Periodontology* 51, no. 10: 1333–1341. https://doi.org/10.1111/jcpe.14023.

Sanz, M., D. Herrera, M. Kebschull, et al. 2020. "Treatment of Stage I-III Periodontitis-The EFP S3 Level Clinical Practice Guideline." *Journal of Clinical Periodontology* 47, no. 22: 4–60. https://doi.org/10.1111/jcpe.13290.

Schwendicke, F., L. T. Arsiwala, J. Krois, et al. 2021. "Association, Prediction, Generalizability: Cross-Center Validity of Predicting Tooth Loss in Periodontitis Patients." *Journal of Dentistry* 109: 103662. https://doi.org/10.1016/j.jdent.2021.103662.

Shameer, K., K. W. Johnson, B. S. Glicksberg, J. T. Dudley, and P. P. Sengupta. 2018. "Machine Learning in Cardiovascular Medicine: Are We There Yet?" *Heart (British Cardiac Society)* 104: 1156–1164. https://doi.org/10.1136/heartjnl-2017-311198.

Shen, K. L., C. L. Huang, Y. C. Lin, et al. 2022. "Effects of Artificial Intelligence-Assisted Dental Monitoring Intervention in Patients With Periodontitis: A Randomized Controlled Trial." *Journal of Clinical Periodontology* 49: 988–998. https://doi.org/10.1111/jcpe.13675.

Snider, V., K. Homsi, B. Kusnoto, et al. 2024. "Clinical Evaluation of Artificial Intelligence Driven Remote Monitoring Technology for Assessment of Patient Oral Hygiene During Orthodontic Treatment." *American Journal of Orthodontics and Dentofacial Orthopedics* 165: 586–592. https://doi.org/10.1016/j.ajodo.2023.12.008.

Tastan Eroglu, Z., O. Babayigit, D. Ozkan Sen, and F. Ucan Yarkac. 2024. "Performance of ChatGPT in Classifying Periodontitis According to the 2018 Classification of Periodontal Diseases." *Clinical Oral Investigations* 28: 407. https://doi.org/10.1007/s00784-024-05799-9.

Tonetti, M. S., S. Jepsen, L. Jin, and J. Otomo-Corgel. 2017. "Impact of the Global Burden of Periodontal Diseases on Health, Nutrition and Wellbeing of Mankind: A Call for Global Action." *Journal of Clinical Periodontology* 44: 456–462. https://doi.org/10.1111/jcpe.12732.

Troiano, G., L. Nibali, H. Petsos, et al. 2023. "Development and International Validation of Logistic Regression and Machine-Learning Models for the Prediction of 10-Year Molar Loss." *Journal of Clinical Periodontology* 50: 348–357. https://doi.org/10.1111/jcpe.13739.

Trombelli, L., and R. Farina. 2020. "Implementation of Patient-Based Risk Assessment in Practice." In *Risk Assessment in Oral Health*, edited by I. L. C. Chapple and P. N. Papapanou, 203–223. Springer International Publishing.

Tsoromokos, N., S. Parinussa, F. Claessen, D. A. Moin, and B. G. Loos. 2022. "Estimation of Alveolar Bone Loss in Periodontitis Using Machine Learning." *International Dental Journal* 72: 621–627. https://doi.org/10.1016/j.identj.2022.02.009.

Yusof, N. A. M., E. Noor, N. H. Reduwan, and M. Yusof. 2021. "Diagnostic Accuracy of Periapical Radiograph, Cone Beam Computed Tomography, and Intrasurgical Linear Measurement Techniques for Assessing Furcation Defects: A Longitudinal Randomised Controlled Trial." *Clinical Oral Investigations* 25: 923–932. https://doi.org/10.1007/s00784-020-03380-8.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.