ISSN: 2581-5989
PubMed - National Library of Medicine - ID: 101738774

International Journal of Dental Science and Innovative Research (IJDSIR)
IJDSIR : Dental Publication Service
Available Online at:www.ijdsir.com
Volume – 9, Issue – 3, June – 2026, Page No. : 83 – 95

Explainable Artificial Intelligence (XAI) in Prosthodontics and a proposed Framework for Clinical Implementation - A Scoping Review

[1]Dr. Rubina Ali, MDS, Department of Prosthodontics, Emirates Health Services, UAE

[2]Dr. Rahul Patil, MFDRCSI(UK), MDS, Department of Prosthodontics, Emirates Health Services, UAE

Corresponding Author: Dr. Rubina Ali, MDS, Department of Prosthodontics, Emirates Health Services, UAE

Citation of this Article: Dr. Rubina Ali, Dr. Rahul Patil, “Explainable Artificial Intelligence (XAI) in Prosthodontics and a proposed Framework for Clinical Implementation - A Scoping Review”, IJDSIR – June – 2026, Volume – 9, Issue – 3, P. No. 83 – 95.

Type of Publication: Review Article

Conflicts of Interest: Nil

Abstract

Background: AI has become an increasingly common option for various methods of diagnosis and workflow support in prosthodontics and implant dentistry, although many applications are not transparent enough and that can limit the level of trust clinicians have in their use, especially in med-legal matters.

Objectives: The aims of this study were to identify the types of AI applications used, how they use explainability, and to suggest a method that could result in the adoption of clinically relevant applications.

Methods: We conducted a scoping review (2013 - 2025) based on the PRISMA-ScR method for all studies relating to the application of AI to prosthodontic or implant-related tasks with quantifiable results. Studies were assigned one of four levels of explainability: none, simple, local, or intrinsic.

Results: A total of 18 studies were included from a search of 2,341 records and the majority of studies reported on either implant prognosis or radiographic assessment methods. A substantial proportion of studies relied on black-box models, while only a minority incorporated clinically meaningful explainability approaches.

Conclusions: Overall, the current use of AI in prosthodontics is performance-based without any regard for explainability. We suggest using a four-tiered explainability framework to evaluate and support the design, implementation and maintenance of clinically relevant applications of AI in prosthodontics.

Keywords: Explainable AI, Artificial Intelligence, Black Box, Prosthodontics, Dental Implants.

Introduction

Treatment planning for prosthodontic rehabilitation is a challenging and complicated, multi-faceted process with irreversible and costly results.[1-3] Artificial intelligence (AI) is being utilized more frequently as a tool to enhance clinical decision-making.[4]

AI can be applied in several areas of prosthodontics, including imaging, predicting implant success, assessing the health of peri-implant tissues, computer-aided design and manufacturing (CAD/CAM) workflows.[5-13] Although these solutions can yield accurate results, many of the AI algorithms remain black boxes, which can adversely affect the level of trust, ability to interpret the output and provide clinical justification for its use.[14-16] Explainable AI (XAI) can address the black box constraint by enhancing transparency through post hoc or interpretable model methods (e.g., SHAP, LIME, Grad-CAM).[17-19] Because of the rapid increase and diversity in the literature, a scoping review is necessary to provide a comprehensive overview of the evidence and to identify gaps, particularly in standardization and validation.[20-23]

- **Objectives**

This scoping review aimed to-

1. map AI/ML/DL applications in prosthodontics
2. quantify and describe the prevalence and depth of explainability methods
3. summarize validation practices and reporting patterns that influence translation
4. propose a clinically oriented framework for implementation.

The review addressed-

1. What clinical tasks have AI models been applied to in prosthodontics?
2. How frequently are explainability methods reported and what type of explanation is provided?
3. What validation approaches and performance metrics are used?
4. What gaps remain for safe clinical implementation?

Methods

Records were searched in PubMed, Scopus, Web of Science and Google Scholar. AI terms related to prosthodontic/implants were combined with outcome measures. The review protocol followed the PRISMA extension for scoping reviews (PRISMA-ScR.[24] Any reference lists that were previously published were reviewed and their records were obtained via snowballing.

The articles included in the peer-review used english literature (study date range of 2013-2025) that examined AI/ML/DL and how they are applied to clinical tasks (with quantitative outcomes such as prosthodontics or implant related tasks). Two independent reviewers screened studies. Discrepancies were resolved by consensus.

Excluded articles were any reviews, editorials, abstracts, simulations that were not clinical or studies not focusing on prosthodontics. Explainability was evaluated as an outcome rather than as a determinant for inclusion in the studies.

Table 1: Information sources and record yield

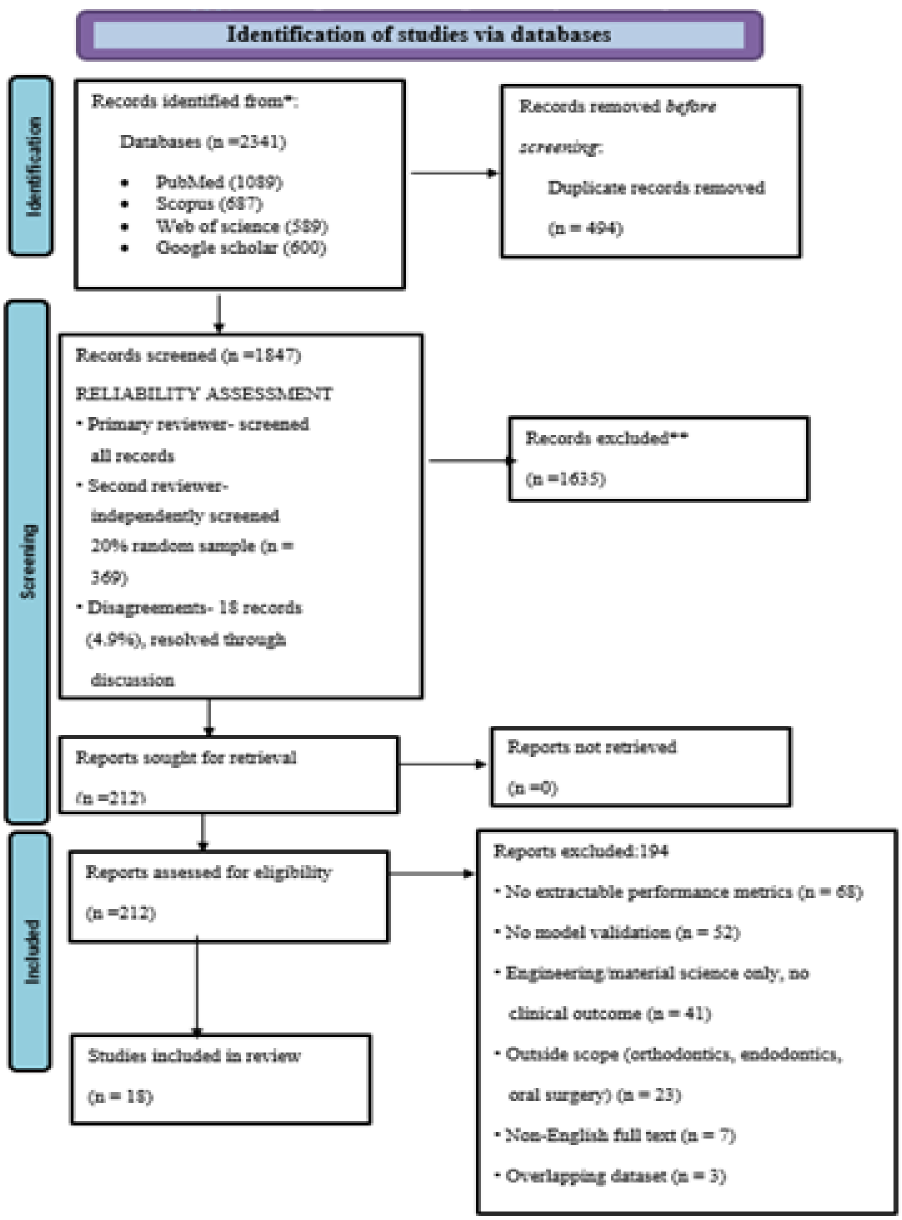
Database	Coverage	Records (n)
PubMed/MEDLINE	Biomedical literature	1089
Scopus	Multidisciplinary database	687
Web of Science Core Collection	Multidisciplinary citation index	589
Google Scholar	Broad scholarly search	600
Total	-	2341

Table 2: Eligibility criteria

Domain	Inclusion	Exclusion
Publication type	Peer-reviewed original research, full text	Reviews, editorials, letters, abstracts only
Language & period	English, 01 Jan 2013-31 Mar 2025	Non-English; outside date range
Clinical scope	Prosthodontics/implant-related applications	Non-prosthodontic dental fields
AI component	AI/ML/DL applied to a clinical task	Engineering simulations without clinical interpretation
Outcomes	Quantitative performance metrics reported	No extractable quantitative outcomes

The search produced total 2,341 total. After removing duplicates (494), 1,847 titles and abstracts were reviewed and 212 full-text articles were evaluated for inclusion. Ultimately, a total of 18 studies were included in the Evidence Map.

Figure 1: PRISMA ScR flow diagram illustrating the study selection process



The explanation of AI in prosthodontics has not been defined using an established taxonomy. Therefore, the classification of explainability was established prior to the review and defined into four categories as follows-

1. No explanation (i.e. Black-Box)
2. Global Attribution for Basic Level (e.g. Feature ranking, Coefficient)
3. Local Attribution/Instance Explanation (e.g. SHAP, LIME, Grad-CAM)
4. Intrinsic Interpretability (i.e. Transparent Mode, Decision Tree/Nomogram).

Table 3: Explainability classification

Level	Label	Operational definition
1	None (black box)	No explanation beyond performance metrics, no feature/region attribution.
2	Basic global attribution	Model-level summaries (feature importance, coefficients, simple rules) without case-level reasoning.
3	Local / instance-level	Case-specific explanation (e.g. SHAP, LIME) or attention heatmaps (e.g.Grad-CAM).
4	Intrinsic interpretability	Model is transparent by design (e.g. decision tree/nomogram) enabling direct inspection.

- **Study Selection, Characteristics**

Among 2,341 records obtained, 18 were eligible. Figure 1 demonstrates how identification, screening for inclusion/exclusion criteria and eligibility and inclusion processes were involved in identifying the appropriate studies. Study screening and selection were performed independently by two reviewers. Titles and abstracts were initially screened, followed by full-text evaluation based on predefined eligibility criteria. Any disagreements were resolved through discussion and consensus. A formal protocol was developed prior to study initiation. Data were charted on study characteristics, clinical domain, model type, input data, validation approach and performance metrics (AUC, accuracy, sensitivity/specificity, Dice, ICC, ΔE). Findings of the selected studies were summarized descriptively because of heterogeneity across tasks and outcomes in Supplementary Table S1.

Study attributes were predominantly retrospective and included single-center databases, mainly using clinical records or radiographs or both. Some studies were multi-center in design that reflect patient population diversity and imaging protocols much closer to what occurs in real world space than single center data.[25]

- **Clinical Domains of Studies and Performance Reporting**

Evidence was identified under three broad categories. (1) Implant survival/failure prediction and marginal bone loss assessments including radiomic-based trabecular analysis (n=10) [10-12,25-31] (2) Prosthetic design/aesthetics using CAD/CAM workflows, use of resin composites for modelling and RPD tools and prediction of the shade of prostheses (n=5) [32-36] and (3) Diagnostic and functional applications, bruxism /TMD and tooth-level prognosis (n=3).[37-39]

Table 4: Evidence map of AI applications in prosthodontics (n=18)

Clinical domain	No. of studies	Typical inputs	Common metrics
Implant prognosis / peri-implant radiographic assessment	10	Clinical records, periapical/panoramic radiographs, radiomics	AUC, accuracy, sensitivity/specificity, Dice
Prosthetic design / esthetics	5	CAD/CAM parameters, scan-derived data, material variables, photographs	ICC, Dice, ΔE, accuracy
Diagnostic / functional applications	3	Sensors, clinical variables, imaging	Accuracy, AUC, classification metrics

Performance reporting methods used for diagnostic investigations were AUC/accuracy, imaging interventions used Dice metric & aesthetic outcomes used ΔE. Although individual study discrimination values tended to be high, reporting of calibration, overfitting and external validation was lacking, which limits clinical application.[40,41]

Explainability of the models was limited to being a "black-box" model for 44% of the models, 39% provided only basic global attribution for the model outputs while only a few of the reviewed studies provided clinically relevant insights (e.g. SHAP, Grad-CAM, Intrinsically interpretable).

Table 5: Prevalence of explainability across included studies

Explainability level	Studies (n)	Percentage
None (black box)	8	44%
Basic global attribution	7	39%
Local / instance-level explanations	2	11%
Intrinsic interpretability	1	6%

Discussion

Existing literature on artificial intelligence in prosthodontics has predominantly focused on model accuracy often overlooking the importance of transparency, interpretability and clinical accountability. While high-performance metrics are encouraging, they do not necessarily translate into safe or reliable clinical decision-making particularly in prosthodontic treatments where outcomes are often irreversible and financially significant.

The limited adoption of explainability techniques observed in this review reflects a broader trend in medical AI research where performance optimization has traditionally been prioritized over interpretability. However, in clinical disciplines such as prosthodontics, the ability to understand and justify model outputs is essential not only for clinician confidence but also for patient communication and medico-legal responsibility. The overwhelming majority of included studies were retrospective and single-centre, which inflates performance metrics and does not account for variations in imaging protocols, operator skill, or patient demographics. Only one study used an intrinsically interpretable model (Level 4) and 44% were complete black boxes, making medico-legal justification for clinical decisions impossible. None of the implant prognosis studies reported calibration metrics (e.g. calibration plots, Brier scores) which are essential to judge the reliability of probabile predictions. Prosthetic design studies lacked clinically meaningful external validation, color prediction models for instance, were tested on benchtop specimens rather than on direct intraoral outcomes.

The proposed four-layer framework addresses these gaps by shifting the focus from purely data-driven model development toward clinically meaningful integration. By prioritizing clinically relevant questions, incorporating interpretable modeling strategies and aligning outputs with clinician needs, AI systems can become more trustworthy and implementable in real-world settings.

Proposed four-layer framework for clinical implementation

Limitations identified in existing AI systems were primarily the use of black-box models, insufficient validation and limited to no clinician-centric explanations. These limitations have led to the development of an XAI framework with four layers (Figure 2)-

1) Clinical issue focus - Prioritize high-value, clinically relevant and impactful questions as opposed to data-driven predictions based on the model learning from historical data.

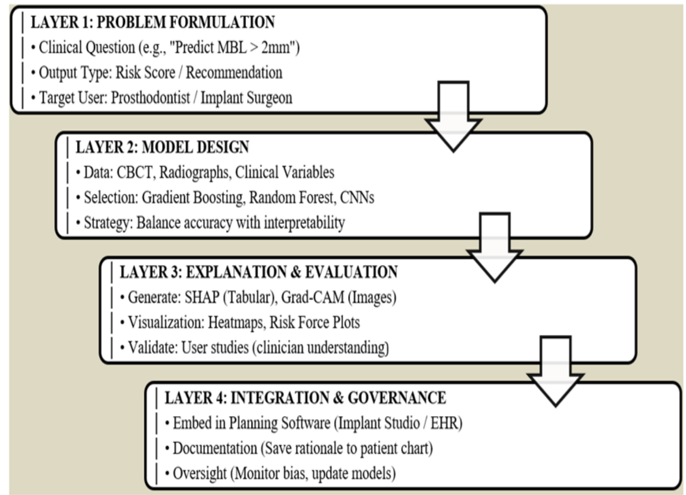
2) Transparency in models - Build interpretable models through the integration of XAI tools or build interpretable models.

3) Clinician-focused explanations - Create clinician-friendly output such as heatmaps or other clinician-friendly output and validate with clinician feedback.

4) Integration of XAI into workflow - Integrate XAI into existing clinical systems, providing the ability for clinician oversight and override.

The goal of this framework is to change AI systems used for hospital decisions from an isolated focus on model accuracy versus trustworthiness and auditable decision support systems within the clinical setting.[42-46]

Figure 2: Schematic of the Four-Layer XAI Framework for prosthodontic applications



Limitations

- The search was restricted to English-language publications in four databases. Therefore, relevant studies in other languages or indexed elsewhere may have been missed.
- Grey literature, preprints and ongoing clinical trial registries were not systematically searched which could introduce publication bias favoring positive results.
- The explainability classification framework while based on existing taxonomies (e.g. from DARPA's XAI program) was developed and applied post hoc by the authors. Subjectivity in assigning a level, particularly between "none" and "basic global attribution" cannot be excluded.
- The small number of finally included studies (n=18) and high heterogeneity in outcome metrics prevented any quantitative synthesis, limiting the ability to provide comparative performance benchmarks.

- Formal risk-of-bias assessment was not performed (as typical for scoping reviews), therefore, the quality of the evidence underpinning the high reported performance values remains unverified.

Future prospects

Based on the gaps identified, the following avenues should be prioritized-

- Mandatory Reporting of Explainability- Future AI studies in prosthodontics should routinely classify and report their explainability level using a standardised framework (such as the four-level system proposed here). Journal reviewers and editors should consider use of XAI a proxy for clinical readiness.
- Prospective Multicentric Validation- High-performing models must undergo rigorous external validation across multiple clinics and imaging devices. Prospective study designs (e.g. embedded in digital workflow trials) are essential to evaluate true impact on treatment planning and patient outcomes.
- Adoption of Intrinsically Interpretable Models- Wherever possible, researchers should prioritize Level 4 (intrinsic interpretability) models (e.g. scoring systems, nomograms, decision trees) over black-box models accompanied by post-hoc explanations as these offer the greatest transparency for regulatory and medico-legal approval.
- Clinician-Centered XAI Interfaces- Explanations must be co-designed with prosthodontists to produce clinically actionable heatmaps or decision rationales (e.g.implant predicted to fail because of low trabecular bone density at site x). The impact of such interfaces on clinician trust, decision time and diagnostic accuracy should be measured.
- Standardised Guidelines for Dental AI- The dental community should develop an extension of TRIPOD-AI or CONSORT-AI specific to prosthodontic and implant applications integrating minimally required XAI and calibration metrics.
- Integration of Prospective Registries- Developers should embed AI modules into ongoing implant registries to continuously monitor for data drift and to provide real-world evidence for regulatory approval.

Comparison with Existing XAI Initiatives

The proposed framework departs meaningfully from earlier XAI models by embedding clinical actionability at its core rather than stopping at algorithmic transparency. DARPA's three-pillar XAI program emphasizes explainable models, explanation interfaces, and psychological models of explanation, all of which are research-oriented and aimed at improving human-machine collaboration in general settings.[14] While foundational, that program does not provide a pathway for translating explanations into daily clinical workflows or for meeting the specific medico-legal requirements of high-stakes disciplines. Similarly, Holzinger's human-in-the-loop paradigm advocates for iterative interaction between the AI system and the human expert, but it remains principally a model-training philosophy rather than a deployment framework.[44] CONSORT-AI and SPIRIT-AI, in turn, are reporting guidelines that ensure trial transparency, not operational blueprints for designing and integrating AI into real clinical environments.[45]

A dentistry-specific framework is necessary because prosthodontic and implant interventions are irreversible, financially significant, and often rely on imaging data whose interpretation demands visually grounded justifications that are meaningful to the clinician and the patient. Generic biomedical AI frameworks rarely account for the combined need for case-level visual evidence (e.g. a heatmap of implant-site bone loss) and for an auditable override mechanism that a prosthodontist can use when the AI's recommendation contradicts clinical experience. The four-layer framework explicitly adds these components: Layer 3 (Clinician-focused explanations) translates model outputs into actionable, human-readable formats co-designed with practitioners, and Layer 4 (Integration of XAI into workflow) embeds the system into existing clinical infrastructure, preserving clinician oversight and override capability. Most current biomedical XAI models stop at post-hoc saliency maps or feature importance rankings-our framework builds upon those principles but extends them into a complete implementation cycle grounded in the realities of prosthodontic care.

Conclusion

The use of Artificial Intelligence technologies will increasingly take place in prosthodontics, particularly concerning implant prognosis, radiographic evaluation and computer aided design (CAD)/computer aided manufacturing (CAM) workflows. However, the potential for using AI for decision support in clinical practice lacks adequate understanding and consistent standards regarding the degree of explainability of AI technologies. Future studies of AI technologies used in decision support should include transparency and rigorous validation methods as fundamental prerequisites for providing safe, accountable and clinically relevant decision support.

References

1. Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. J Dent Res. 2020;99(7):769-74.
2. Hung KF, Ai QYH, Leung YY, Yeung AWK. Potential and impact of artificial intelligence algorithms in dento-maxillofacial radiology. Clin Oral Investig. 2022;26(9):5235-55.
3. Sitaras S, Tsolakis IA, Gelsini M, Tsolakis AI, Schwendicke F, Wolf TG, et al. Applications of artificial intelligence in dental medicine: a critical review. Int Dent J. 2025;75(2):474-86.
4. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
5. Wang L, Xu Y, Wang W, Lu Y. Application of machine learning in dentistry: insights, prospects and challenges. Acta Odontol Scand. 2025; 84:145-54.
6. Koul R, Guruprasada, Kalia D, Verma K. Artificial intelligence in prosthodontics: current applications and future avenues: a narrative review. J Prosthodont Res. 2024. [Epub ahead of print].
7. Wu Z, Yu X, Wang F, Xu C. Application of artificial intelligence in dental implant prognosis: a scoping review. J Dent. 2024; 144:104924.
8. Salgau CA, Moga RA, Alexa AI, Girbovan C, Odone A, Today D. Applications of machine learning in periodontology and implantology: a comprehensive review. Ann Biomed Eng. 2024;52(9):2348-71.
9. Revilla-Leon M, Meyer MJ, Ozcan M. Artificial intelligence applications in implant dentistry: a systematic review. J Prosthet Dent. 2023;129(2):293-300.
10. Huang N, Liu P, Yan Y, Xu L, Wu Y, Li Z. Predicting the risk of dental implant loss using deep learning. J Clin Periodontol. 2022;49(9):872-83.
11. Rekawek P, Herbst EA, Guo S, Tisch VP, Butler JH, Alexander SA. Machine learning and artificial intelligence: a web-based implant failure and peri-implantitis prediction model for clinicians. Int J Oral Maxillofac Implants. 2023;38(3):576-82.

12. Ha SR, Park HS, Kim EH, Han JS. A pilot study using machine learning methods about factors influencing prognosis of dental implants. J Adv Prosthodont. 2018;10(6):395-400.
13. Kose C, Oliveira D, Pereira PNR, Rocha MG. Using artificial intelligence to predict the final color of leucite-reinforced ceramic restorations. J Esthet Restor Dent. 2023;35(1):105-15.
14. Gunning D, Aha DW. DARPA's explainable artificial intelligence (XAI) program. AI Mag. 2019;40(2):44-58.
15. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20(1):310.
16. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017; 30:4765-74.
17. Ribeiro MT, Singh S, Guestrin C. Why should I trust you? explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. p. 1135-44.
18. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision. 2017. p. 618-26.
19. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. J Biomed Inform. 2021; 113:103655
20. Nguyen TT, Larrivée N, Lee A, Bilaniuk O, Durand R. Use of artificial intelligence in dentistry: current clinical trends and research advances. J Can Dent Assoc. 2021;87: l7
21. Joda T, Bornstein MM, Jung RE, Ferrari M, Waltimo T, Zitzmann NU. Recent trends and future direction of dental research in the digital era. Int J Prosthodont. 2020;33(4):471-82.
22. Tandon S, Basnarkov L. Artificial intelligence in prosthodontics: a scoping review. J Prosthet Dent. 2022.
23. Arsiwala A, Barmak AB, Zandinejad A, Revilla-León M. Artificial intelligence applications in shade selection: a systematic review. J Prosthet Dent. 2023;130(1):26-34.
24. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169(7):467-73
25. Zhu Y, Liu Y, Zhao Y, Chen W, Li Q, Wang Y. Anatomically based multitask deep learning radiomics nomogram predicts the implant failure risk in sinus floor elevation. Clin Oral Implants Res. 2025;36(11):1424-33.
26. Zhang H, Shen J, Zhang P, Chen X, Jiang H. Trabeculae microstructure parameters serve as effective predictors for marginal bone loss of dental implant in the mandible. Sci Rep. 2020; 10:18437.
27. Kheder W, Leblouba M, Osman RB, Hamdoon Z. Multicentre validation and clinical interpretation of an explainable gradient boosting model for dental implant survival/failure prediction. J Dent. 2025; 163:106166.
28. Zhang C, Fan L, Zhang S, Zhao J, Gu Y. Deep learning-based dental implant failure prediction from periapical and panoramic films. Quant Imaging Med Surg. 2023;13(2):935-45.

29. Cha JY, Yoon HI, Yeo IS, Huh KH, Han JS. Peri-implant bone loss measurement using a region-based convolutional neural network on dental periapical radiographs. J Clin Med. 2021;10(5):1009.
30. Kibcak E, Buhara O, Temelci A, Akkaya N, Ünsal G, Minervini G. Deep learning-driven segmentation of dental implants and peri-implantitis detection in orthopantomographs: a novel diagnostic tool. J Evid Based Dent Pract. 2025;25(1):102058.
31. Troiano G, Fanelli F, Trottini M, Tati E, Catapano S, Exarchou V. Can radiomic features extracted from intraoral radiographs predict physiological bone remodelling around dental implants? A hypothesis-generating study. J Clin Periodontol. 2023;50(7):932-41.
32. Yeslam HE, Freifrau von Maltzahn N, Nassar HM. Revolutionizing CAD/CAM-based restorative dental processes and materials with artificial intelligence: a concise narrative review. PeerJ. 2024;12: e17793.
33. Farook TH, Ahmed S, Jamayet NB, Rashid F, Barman A, Dudley J. Computer-aided design and 3-dimensional artificial/convolutional neural network for digital partial dental crown synthesis and validation. Sci Rep. 2023; 13:87.
34. Li H, Sakai T, Tanaka A, Ogura M, Lee C, Yamaguchi S, et al. Interpretable AI explores effective components of CAD/CAM resin composites. J Dent Res. 2022;101(11):1363-71.
35. Takahashi T, Nozaki K, Gonda T, Ikebe K. A system for designing removable partial dentures using artificial intelligence. Part 1. Classification of partially edentulous arches using a convolutional neural network. J Prosthodont Res. 2021;65(1):115-18.
36. Mahrous A, Elgreatly A, Tsujimoto A, Oozeir Q, Schneider GB. The use of artificial intelligence and game-based learning in removable partial denture design: a comparative study. J Dent Educ. 2023;87(8):1188-99.
37. O'Hare E, Cogan J, Dillon F, Lowery M, O'Cearbhaill E. An intraoral non-occlusal MEMS sensor for bruxism detection. IEEE Sens J. 2021;21(19):21322-30.
38. Lee KS, Jha N, Kim YJ. Risk factor assessments of temporomandibular disorders via machine learning. Sci Rep. 2021; 11:19802.
39. Lee SJ, Chung D, Asano A, Sasaki D, Maeno M, Ishida Y, et al. Diagnosis of tooth prognosis using artificial intelligence. Diagnostics (Basel). 2022;12(6):1422.
40. Park WJ, Park JB. Deep learning for dental implant failure prediction: a comparative study of various algorithms. Clin Oral Implants Res. 2022;33(5):549-58.
41. Schwendicke F, Singh T, Lee JH, Gaudin R, Chaurasia A, Wiegand T, et al. Artificial intelligence in dental research: checklist for reviewers, readers. J Dent. 2021; 107:103610.
42. Revilla-León M, Gómez-Polo M, Barmak AB, Inamali AM, Moshaverinia A, Park YJ, et al. Artificial intelligence models for tooth-supported fixed dental prostheses: a systematic review. J Prosthet Dent. 2023;130(6):836-47.
43. Grischke J, Johannsmeier L, Eich L, Grigna V, Haddadin S. Dentronics: towards robotics and artificial intelligence in dentistry. Dent Mater. 2020;36(6):765-78.
44. Holzinger A. Interactive machine learning for health informatics: when do we need the human-in-the-loop? Brain Inform. 2016;3(2):119-31.
45. Liu X, Rivera SC, Moons KGM, Collins GS, Denniston AK; SPIRIT-AI and CONSORT-AI

Working Groups. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence- the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74.

46. Du M, Liu N, Hu X. Techniques for interpretable machine learning. Commun ACM. 2020;63(1):68-77. doi:10.1145/3359786

Supplementary Table 1: Domain wise Characteristics of Included 18 Studies

Study /Year	Country/Setting	Clinical Objective	Sample Size & Population	AI Method	Input Data	Performance	Validation	Transparency Level	Explainability Method	Key Findings	Limitation	Clinical Implications
IMPLANT PROGNOSIS & SURVIVAL STUDIES												
Huang 2022[10]	China, University hospital	Predict implant loss risk	862 implants, Adults, >1yr follow-up	Deep learning (CNN)	Clinical + radiographic	AUC: 0.89	Internal only	Level 1	None	High-risk identified 6 months pre-failure	Single-center, no external validation	Enable preventive interventions
Rekawek 2023[11]	USA, Multi-practice	Web-based failure predictor	1,165 patients, Complete data required	Random forest	Clinical parameters	Accuracy: 84%	Internal only	Level 2 -basic global	Feature importance	Smoking & diabetes strongest predictors	Limited to structured data	Real-time risk assessment
Ha 2018[12]	South Korea, Dental hospital	Identify prognostic factors	222 implants,2010-2015 placement	SVM, ANN	Clinical data	AUC: 0.82	5-fold CV	Level 2 basic global	Feature ranking	ML outperformed traditional statistics	Small sample size	Personalized treatment planning
Zhang H 2020[26]	China, Hospital	Predict bone loss from trabecular structure	148 implants, Mandibular only	Random forest	Trabecular parameters	AUC: 0.91	Internal only	Level 2 basic global	Feature importance	Trabecular parameters predicted 2-year loss	Mandible only	Pre-operative risk stratification
Zhu 2025[25]	China, Two centers	Predict sinus elevation complications	316 patients, Sinus lift candidates	Deep learning	CBCT radiomics	AUC: 0.94	External (n=78)	Level 4 (Intrinsic)	None	Anatomical features key to prediction	Limited to sinus procedures	Surgical planning optimization
Kheder 2025[27]	Multi-center, 3 countries	Explainable implant survival	2,845 implants, 5-year follow-up	Gradient boosting	Clinical variables	AUC: 0.93	External validation	Level 3 (Local/instance)	SHAP (local+ global)	Individual risk factors quantified	Retrospective design	Patient counseling with specific risks
RADIOGRAPHIC DETECTION & ANALYSIS												
Zhang C	China, University	Detect failing	400 images, Periapical/pa	CNN	Radiographs	Accuracy:	Test set	Level 3 Local/in	Grad-CAM	Visual localizati	Binary classific	Early detection

2023[28]	clinic	implants	noramic			91%	(20%)	stance		on of failure indicators	ation only	system
Cha 2021[29]	South Korea, Dental hospital	Measure peri-implant bone loss	800 implants, Digital radiographs	R-CNN	Periapical radiographs	ICC: 0.86	Test set (30%)	Level 1 (None)	None	Automated measurements match experts	2D measurements only	Standardized monitoring
Kibcak 2025[30]	Turkey, University	Detect peri-implantitis	1,200 implants, OPG images	U-Net	OPG	Dice: 0.88	Internal only	Level 1 (None)	None	High segmentation accuracy	Single imaging modality	Screening for general dentists
Troiano 2023[31]	Italy, Private practice	Predict physiological remodeling	95 implants,1-year follow-up	Radiomic analysis	Intraoral radiographs	AUC: 0.78	Boot strap	Level 2 Basic global	Feature selection	Texture features predictive	Small sample	Distinguish normal vs. pathological
PROSTHETIC DESIGN & ESTHETICS												
Farook 2023[33]	Australia, University	Automate crown margin detection	120 crowns, Digital preparations	CNN + 3D	Digital scans	Accuracy: 94%	Hold-out (25%)	Level 1 (None)	None	Precise margin line detection	Laboratory setting only	CAD/CAM automation
Li H 2022[34]	Japan, Research institute	Optimize CAD/CAM composite properties	144 composites, Experimental materials	Interpretable ML	Material composition	R^2: 0.92	10-fold CV	Level 4 (Intrinsic)	Intrinsic architecture	Filler content most influential	Not tested clinically	Material selection guidance
Takahashi 2021[35]	Japan, University	Classify RPD frameworks	500 cases, Kennedy classifications	CNN	Arch images	Accuracy: 88%	Test set (20%)	Level 1 (None)	None	Automated design initiation	Limited to classification	Speeds design process
Mahrous 2023[36]	USA, Dental school	Educational RPD tool	150 designs, Student projects	Decision tree	Design parameters	Accuracy: 85%	70:30 split	Level 4 (Intrinsic)	Decision rules	Improved student performance	Educational context only	Training tool potential
Kose 2023[13]	Brazil, University	Predict ceramic shade	60 restorations, Leucite-reinforced	ANN	Color measurements	ΔE: 1.8±0.6	Leave-one-out	Level 2 (Basic global)	Weight analysis	ΔE within clinical acceptability	Single ceramic system	Shade selection support
FUNCTIONAL DIAGNOSIS												
O'Hare	Ireland,	Detect	25 patients,	SVM	MEMS	Sensiti	Real-	Level 1	None	High	Small	Objective

2021[37]	University	bruxism episodes	Suspected bruxism		sensor data	vity: 92%	time test	(None)		grinding detection rate	sample	diagnosis tool
Lee KS 2021[38]	South Korea, Dental hospital	Assess TMD risk factors	495 patients, TMD questionnaire	XG Boost	Clinical questionnaire	AUC: 0.87	5-fold CV	Level 2 (Basic global)	Feature importance	Psychological factors significant	Self-reported data only	Screening prioritization
Lee SJ 2022[39]	Japan, University	Predict tooth prognosis	3,000 teeth, Complete records	Random forest	Clinical + radiographic	AUC: 0.91	Test set (30%)	Level 1 (None)	None	Periodontal status most predictive	No external validation	Treatment planning support