Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integration of artificial intelligence in orthodontic diagnosis and treatment planning a PRISMA ScR guided scoping review

Salah Bin Hafedh
salah.hafedh@gmc-ye.com
Sana'a University
Ahmad Hashridz Bin Ruslan
Universiti Sains Malaysia
Ramy Ishaq
Sana'a University
Rozita Hassan
Universiti Sains Malaysia
Alaa Ali Maudhah
Sana'a University

Systematic Review

Keywords: artificial intelligence, machine learning, deep learning, orthodontics, diagnosis, treatment planning, scoping review

Posted Date: February 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8641111/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Artificial intelligence (AI) applications in orthodontics are rapidly expanding across diagnosis, image analysis, and treatment planning.

Methods

A PRISMA-ScR–guided scoping review was conducted. PubMed/MEDLINE, Scopus, Web of Science, and Google Scholar were searched from 2010 to 16 September 2025. Original studies in orthodontics that used AI or machine learning for diagnosis, prediction, image analysis, or treatment planning were eligible. Two reviewers independently screened records, extracted data, and assessed risk of bias using QUADAS-2 for diagnostic accuracy studies and PROBAST for prediction model studies. Owing to heterogeneity in study design, datasets, and outcome metrics, results were synthesized narratively.

Results

Of 1,162 records identified, 1,008 remained after duplicate removal and were screened by title and abstract. A total of 154 full-text articles were assessed for eligibility, and 45 met the inclusion criteria. Frequent AI tasks included cephalometric landmark detection, malocclusion classification, extraction-decision support, treatment duration prediction, and cone-beam computed tomography (CBCT)-based segmentation. Many studies reported high accuracies for cephalometric landmark detection (mean radial error < 2 mm and successful detection rates > 80%) and malocclusion classification (accuracies > 85%). However, risk-of-bias concerns, particularly in analysis and validation domains, were common, and external validation was infrequent.

Conclusions

AI models show promising performance for orthodontic diagnosis and treatment planning and may enhance efficiency and standardization of care. Nevertheless, non-standardized outcome measures, limited external validation, and insufficient reporting of model development and evaluation currently restrict clinical translation. Larger, multicenter datasets, standardized benchmarks, and robust validation—ideally following AI-specific reporting guidelines—are required before routine clinical adoption.

1. Introduction

In recent years, artificial intelligence (AI) has moved from being a futuristic concept to an everyday reality in healthcare. Orthodontics, where precision in diagnosis and treatment planning is essential for long-term functional and aesthetic outcomes, has been particularly affected by this transformation [1–3]. AI tools can help reduce diagnostic errors, streamline clinical workflows, and allow orthodontists to dedicate more time to direct patient care [4–6].

Machine learning (ML) and deep learning (DL), as subfields of AI, are particularly powerful for analyzing diverse clinical datasets, including cephalometric radiographs, CBCT scans, panoramic images, intraoral photographs, and structured patient records [7–11]. Convolutional neural networks (CNNs), among the most widely applied DL architectures, have shown high accuracy in cephalometric landmark detection, malocclusion classification, and image segmentation, often matching or surpassing the performance of human examiners [12–16]. Landmark detection studies frequently report mean radial errors below 2 mm and successful detection rates above 80%, while classification accuracies for malocclusion often exceed 85% [17–20].

Beyond diagnosis, AI is increasingly applied to complex clinical decision-making, such as predicting treatment duration, identifying when extractions are necessary, and estimating the need for orthognathic surgery [21–27]. These applications indicate that AI is evolving from a purely experimental tool to a clinically relevant adjunct in routine orthodontic practice [28–30]. However, important challenges remain. Many models are trained on single-center or demographically narrow datasets, which may limit generalizability. Ethical and regulatory issues—including data privacy, algorithmic bias, and medico-legal responsibility—also require careful consideration. Moreover, the “black-box” nature of many DL models can impede clinician trust in AI-generated recommendations [31–35].

This PRISMA-ScR–guided scoping review synthesizes evidence from original studies published between 2010 and 2025 that evaluated AI for orthodontic diagnosis and treatment planning. The objectives are to (1) summarize the types of AI tasks and methodologies used in orthodontics, (2) evaluate the reported diagnostic and predictive performance of these systems, and (3) identify current limitations and future research priorities for safe and effective AI integration into orthodontic care [36–40].

In contrast to prior narrative and systematic syntheses in this area, this scoping review is designed as an evidence map that charts the primary-study landscape by clinical task, data source, algorithm family, reference standard, and validation approach (internal versus external validation), and highlights methodological and reporting gaps that currently limit clinical translation.

2. Materials and Methods

2.1 Protocol and registration

This scoping review was conducted in accordance with the PRISMA-ScR statement and relevant AI-focused reporting extensions when applicable. The review protocol was prospectively registered with PROSPERO (registration number CRD420251134644).

2.2 Information sources

Electronic searches were carried out in PubMed/MEDLINE, Scopus, Web of Science (Core Collection), and Google Scholar. Databases were searched for articles published between 1 January 2010 and 16 September 2025. Search strategies combined keywords and MeSH terms related to “artificial intelligence,” “machine learning,” “deep learning,” “orthodontics,” “diagnosis,” “treatment planning,” and “cephalometrics.” Detailed, database-specific search strings are provided in the Supplementary material (Search strategies document).

2.3 Eligibility criteria

Inclusion criteria

- Peer-reviewed original research articles.
- Studies involving AI (including ML, DL, CNNs, or hybrid techniques) applied to orthodontic diagnosis, prediction, image analysis, or treatment planning.
- Human clinical data or imaging (e.g., 2D cephalometric radiographs, CBCT, panoramic radiographs, intraoral photographs, or structured clinical records).
- Reporting of at least one quantitative performance metric (e.g., accuracy, sensitivity/specificity, AUC, mean radial error, Dice similarity coefficient, mean absolute error).
- Articles published in English between 1 January 2010 and 16 September 2025.

Exclusion criteria

- Narrative reviews, scoping reviews, editorials, commentaries, conference abstracts, and letters without original data.
- Studies not primarily focused on orthodontics (e.g., general dentistry, restorative dentistry, endodontics) or AI applications not related to diagnosis or treatment planning.
- Studies that did not clearly describe their AI methodology or lacked a human reference standard.
- Animal or in-vitro studies.

2.4 Data extraction and quality assessment

Two reviewers independently screened titles and abstracts, followed by full-text assessment of potentially eligible studies. Disagreements were resolved through discussion or consultation with a third reviewer.

Data were extracted using a standardized form and included: first author, publication year, country, imaging modality, primary AI task (e.g., landmark detection, classification, prediction, segmentation), dataset size, AI architecture or algorithm, reference standard (e.g., expert orthodontist annotations), primary performance metrics, and whether internal and/or external validation was performed.

Risk of bias in diagnostic accuracy studies was assessed using the QUADAS-2 tool, which evaluates four domains: patient selection, index test, reference standard, and flow/timing [2]. Prediction model studies were assessed using the PROBAST tool, which covers participants, predictors, outcomes, and analysis domains [3]. Summary risk-of-bias assessments are presented in the Supplementary QUADAS-2 and PROBAST tables, with study-level details provided separately.

2.5 Synthesis methods

The included studies displayed substantial heterogeneity in study design (diagnostic accuracy vs prediction models), imaging modalities (2D radiographs, CBCT, panoramic radiographs, intraoral photographs, clinical records), AI architectures, and outcome measures (e.g., landmark error, Dice similarity coefficient, AUC, accuracy, MAE). Because of this variability, quantitative meta-analysis was not feasible.

Instead, a structured narrative synthesis was conducted in accordance with PRISMA-ScR guidance. Studies were grouped by primary AI task (e.g., cephalometric landmark detection, malocclusion classification, extraction decision support, treatment duration prediction, CBCT segmentation), and key performance indicators were summarized within each category.

3. Results

The PRISMA-ScR flow diagram summarizing study selection is shown in Fig. 1 near here (PRISMA-ScR flow diagram of study identification, screening, eligibility, and inclusion).

Figure 1. PRISMA-ScR flow diagram of study identification, screening, eligibility, and inclusion.

3.1 Study selection

The database search identified 1,162 records. After removal of duplicates, 1,008 records remained for title and abstract screening. Of these, 854 records were excluded. A total of 154 full-text articles were assessed for eligibility. One hundred and nine full-text articles were excluded for the following main reasons: not orthodontics-related AI studies (n = 35), review/editorial/commentary articles (n = 28), inappropriate or irrelevant study design (n = 20), insufficient diagnostic or treatment-planning detail (n = 16), and duplicate publication (n = 10). Finally, 45 studies met all inclusion criteria and were included in the qualitative synthesis.

3.2 Study characteristics

Table 1 summarizes the characteristics of the 45 included studies. The majority were retrospective observational studies conducted in academic or specialist orthodontic settings. Common imaging modalities included 2D cephalometric radiographs, CBCT, panoramic radiographs, and intraoral photographs.

The main AI tasks and applications were:

- Cephalometric landmark detection and cephalometric analysis;
- Malocclusion and skeletal pattern classification;
- Extraction-decision support and prediction of orthognathic surgery necessity;
- Prediction of treatment duration and treatment outcomes;
- CBCT-based 3D segmentation of dental and skeletal structures.

Sample sizes varied widely, ranging from fewer than 100 images or patients to more than 1,000. CNN-based DL approaches were predominant for image-based tasks, whereas traditional ML algorithms (e.g., random forests, support vector machines, gradient boosting) were more commonly used for structured clinical data and prediction tasks.

Table 1. Study Characteristics of Included Studies (n = 45)

ID	First author (Year)	Country/Region	Journal	Imaging/Modality	Primary Task	Dataset size (N)	AI Method	Reference Standard
1	Cui Z (2022)	Multi-center (China, 15 sites)	Nat Commun	CBCT (3D)	Tooth & alveolar bone segmentation	4,938 scans	Multi-stage 3D CNN (nnU-Net style pipeline)	Expert manu segmentatio
2	Noeldeke B (2024)	Germany	Head Face Med	Intraoral photographs (2D)	Crossbite detection (binary & type)	676 images (311 pts)	CNNs (DenseNet/ResNet variants)	Orthodontist labels
3	Ryu S-M (2023)	Korea	Sci Rep	Intraoral photographs (2D)	Extraction decision recommendation	3,136 images	CNN classifier + landmark regressor	Board-certifi orthodontist decision
4	Sahlsten T (2024)	Finland	PLoS ONE	CBCT (3D)	3D cephalometric landmark detection (33 LM)	309 scans	Deep learning landmarking	Expert annotations
5	Shin J-H (2021)	Korea	BMC Oral Health	Clinical photos + ceph	Necessity of orthognathic surgery (Class III)	140 pts	CNN	Panel consensus (surgery vs n surgery)
6	Volovic J (2023)	USA	Diagnostics (MDPI)	Structured records	Treatment duration prediction	478 pts	Random Forest, Lasso, Elastic Net	Actual durati vs prediction
7	Elnagar M (2022)	USA	Diagnostics (MDPI)	Structured records	Treatment duration prediction	518 pts	Multiple ML models (DT, RF, etc.)	Actual durati vs prediction
8	Wolf D (2024)	Germany (EU dataset)	J Clin Med	EMR + app data	Clear aligner refinement risk prediction	9,942 CAT pts	L1-logistic, XGBoost, SVC-RBF (+SHAP)	Clinician-recorded outcomes
9	Etemad L (2024)	USA; France (2 sites)	Bioengineering (MDPI)	Structured records	Extraction vs non-extraction decision	1,135 pts (2 universities)	Random Forest	Clinician decision
10	Leavitt L (2023)	USA	Orthod Craniofac Res	Structured records	Predict specific extraction patterns	366 pts (extraction cases)	RF, LR, SVM	Clinician treatment pla
11	Mason T (2023)	USA	Int Orthod	Structured records	Extraction vs non-extraction	393 pts	LR, RF, SVM, NN	Clinician decision
12	Huang J (2024)	China	Front Bioeng Biotechnol	Structured records	Extraction decision	Institutional cohort	DT, RF, SVM, MLP; feature importance	Senior specialist pla
13	Arik SÖ (2017)	USA	J Med Imaging	Lateral ceph (2D)	Landmark detection (15 LM)	400 images	CNN (early DL)	Expert annotation
14	Gilmour R (2019)	—	—	Lateral ceph (2D)	Landmark detection (15 LM)	—	—	Expert annotation
15	Li P (2019)	China	Med Image Anal?/Sci Rep	Lateral ceph (2D)	Landmark detection (15 LM)	—	—	Expert annotation
16	Kwon (2019)	Korea	—	Lateral ceph (2D)	Landmark detection (15 LM)	—	—	Expert annotation

17	Oh (2019)	Korea	—	Lateral ceph (2D)	Landmark detection (15 LM)	—	—	Expert annotation
18	Kim (2019/2020)	Korea	—	Lateral ceph (2D)	Landmark detection (15 LM)	860 images	—	Expert annotation
19	Kim (2020)	Korea	—	Lateral ceph (2D)	Landmark detection (23 LM)	2,075 images	—	Expert annotation
20	Takahashi (2020)	Japan	—	Lateral face photographs (2D)	Ceph LM from photos (23 LM)	2,000 images	HRNetV2 + MLP (2-stage)	Ceph-photo superimposi
21	Takahashi (2025)	Japan	—	Lateral face photos (2D)	Ceph LM from photos (Class II/III)	2,320 images	HRNetV2 + MLP (2-stage)	Ceph-photo superimposi
22	Park J-H (2019)	Korea	Angle Orthod	Lateral ceph (2D)	Compare YOLOv3 vs SSD (80 LM)	Train:1028, Test:283	YOLOv3 vs SSD	Expert labels
23	Hwang H-W (2020)	Korea	Angle Orthod	Lateral ceph (2D)	AI vs human (80 LM)	—	YOLOv3-based pipeline	Human expe
24	Yoon H-J (2022)	Korea	Eur J Orthod	Lateral ceph (2D)	Airway-focused LM detection	—	Deep CNN pipeline	Expert annotation
25	Atici S.F. (2022)	UK/Turkey	PLoS ONE	Lateral ceph (2D)	Fully automated CVM stage classification	—	Custom CNN (directional filters)	Expert labels
26	Atici S.F. (2023)	UK/Turkey	—	Lateral ceph (2D)	AggregateNet CVM classifier	—	Parallel structured CNN	Expert labels
27	Gaudot I (2024)	Multi-center (EU)	Med Eng Phys	CBCT/CT (3D)	DentalSegmentator (5-class segmentation)	470 train; 256 test	nnU-Net (3D Slicer extension)	Expert annotation
28	Wang C (2024)	China	Biomed Signal Process Control	CBCT (3D)	Transformer-based tooth segmentation (Trans-VNet)	—	Transformer CNN hybrid	Expert annotation
29	Kartbak SBA (2025)	Turkey	BMC Oral Health	Lateral ceph + intraoral photos (2D)	Intraoral classification via ceph-informed DL	990 pts	DL classifier trained on ceph-derived labels	Cephalometr measuremer
30	Milani O-H (2024)	USA	—	Panoramic (2D)	Third molar development stage classification	—	DL classifier	Expert stagir
31	JOMOS team (2025)	China	J Oral Med Oral Surg	Panoramic (2D)	Impacted mandibular third molar detection & class	2,000 PRs	DL detector/classifier	Radiologist labels
32	Kim S (2024)	Korea	BMC Oral Health	Panoramic (2D)	Indication for extraction (cracked tooth)	—	Multiple DL models	Clinician decision
33	Suh HY (2019)	Korea/USA	Angle Orthod	Structured + ceph	Soft tissue change prediction after surgery	—	Sparse partial least squares (ML)	Post-op measuremer

34	Lee YS (2014)	Korea/USA	AJODO	Structured + ceph	Soft tissue prediction (Class III)	—	Statistical/ML model	Post-op measuremer
35	Wang C-W (2015)	Taiwan	IEEE TMI	Lateral ceph (2D)	Grand challenge benchmark (evaluation)	—	Multiple methods compared	Expert GT (challenge)
36	Wang C-W (2016)	Taiwan	Med Image Anal	Dental radiographs (2D)	Benchmark for analysis algorithms	—	Benchmarking	Expert GT
37	Xie X (2010)	China	Angle Orthod	Structured records	Extraction vs non-extraction	200 pts	ANN	Clinician decision
38	Jung S-K (2016)	Korea	AJODO	Structured records	Extraction vs non-extraction	156 pts	3-layer ANN	Single clinici decisions
39	Li P (2019)	China	Sci Rep	Structured records	Orthodontic treatment planning (broad)	—	ANN	Expert plan
40	Castillo J-C (2019)	Canada/USA	Angle Orthod	3D photogrammetry	3D facial-cheph relationships	—	Statistical + ML links	Manual measuremer
41	Schmidt S (2022)	Germany	Dentomaxillofac Radiol	Panoramic (2D)	Restoration segmentation	1,781 PRs	U-Net variants	Pixelwise GT
42	Kim H (2022)	Korea	Dentomaxillofac Radiol	Panoramic (2D)	Detect restorations & implants	—	Object detection (DL)	Expert labels
43	Craniofacial Growth ML (2025)	USA	Orthod Craniofac Res	Structured records	Long-term growth change prediction	—	ML regression ensemble	Ceph serial records
44	Prasad J (2022)	India	Dent J (MDPI)	Structured records	Clinical decision support (diagnosis & plan)	—	XGBoost/RF (multilabel)	Clinician plar
45	Del Real A (2022)	Korea	Korean J Orthod	Structured records	Predict need for extraction	—	XGBoost/RF	Orthodontist decision

Abbreviations: CBCT = cone-beam computed tomography; CNN = convolutional neural network; DL = deep learning; ML = machine learning; SDR = successful detection rate; MRE = mean radial error; Dice = Dice similarity coefficient; AUC = area under ROC.

4. Discussion

This PRISMA-ScR–guided scoping review demonstrates that AI is no longer a distant prospect but an emerging reality in orthodontic care. Across multiple diagnostic domains, AI tools have achieved performance levels that are directly relevant to everyday practice.

For cephalometric landmark detection, CNN-based models consistently report mean radial errors below 2 mm and high successful detection rates, often comparable to the performance of experienced orthodontists [7–12]. Such reliability suggests that AI can already assist with routine diagnostic tasks, potentially reducing inter-observer variability and saving clinician time. Similarly, ML models developed for predicting extraction decisions or treatment duration show promising accuracy, highlighting the versatility of AI when applied to both image-based and structured clinical data [13–18].

Segmentation of CBCT scans has reached a clinically meaningful level of performance, with advanced DL architectures such as nnU-Net and transformer-based models frequently achieving Dice similarity coefficients greater than 0.90 for teeth and alveolar bone [19–22]. These advances indicate that AI can substantially reduce the time and expertise required for high-resolution volumetric analysis, facilitating broader clinical use of 3D imaging. Decision-support systems for extractions and orthognathic surgery further demonstrate AI's potential in supporting complex treatment planning decisions [23–26].

Despite these encouraging results, several important challenges must be addressed before AI can be widely adopted in routine orthodontic care. First, many included studies used relatively small or demographically homogeneous datasets, often drawn from single institutions, which may limit external validity [27–30]. Only a minority of models underwent external validation on independent datasets or in prospective clinical settings, making it difficult to determine how well these systems will perform in diverse real-world populations.

Second, transparency and interpretability remain major concerns. Most DL models operate as "black boxes," providing accurate predictions without clear explanations. This lack of interpretability can undermine clinician trust and hinder shared decision-making with patients. Explainable AI (XAI) approaches—

including saliency maps, attention mechanisms, and feature-importance analyses—are therefore essential to clarify model reasoning and to support responsible clinical use [31–33].

Third, broader ethical and practical issues must be carefully considered. These include safeguards for patient privacy and data security, the potential for algorithmic bias if training data are unbalanced, and clear assignment of medico-legal responsibility when AI tools are integrated into clinical workflows [34–36]. Robust governance frameworks, transparent reporting, and regulatory oversight will be necessary to address these concerns.

Future research should prioritize:

1. Large, multicenter, and demographically diverse datasets to improve generalizability and reduce algorithmic bias.
2. Standardized benchmarks and publicly accessible datasets for key orthodontic AI tasks, enabling direct comparison of model performance across studies.
3. Prospective clinical studies and implementation research to evaluate how AI tools affect diagnostic accuracy, treatment outcomes, workflow efficiency, and patient-reported outcomes in real clinical settings.
4. Adherence to AI-specific reporting guidelines (e.g., TRIPOD-AI, PROBAST-AI, QUADAS-AI) to enhance transparency, reproducibility, and critical appraisal [4–6,32,33].
5. Integration of explainable and human-in-the-loop AI systems, in which clinicians and algorithms complement each other, ensuring that orthodontists remain the ultimate decision-makers.

In summary, AI is beginning to reshape orthodontics by improving efficiency, reducing inter-observer variability, and enabling more personalized treatment planning. However, substantial work is still required to move from promising prototypes to robust, trustworthy systems that can be safely implemented in routine practice [37–40].

5. Conclusion

AI currently demonstrates strong performance across multiple orthodontic diagnostic and treatment-planning tasks. CNN-based approaches dominate cephalometric landmark detection and classification, while ensemble ML methods show promise for predicting extraction decisions and treatment outcomes. Nevertheless, clinical adoption should be preceded by robust external validation, standardized evaluation frameworks, clear ethical and regulatory guidelines, and comprehensive training for clinicians. With these safeguards in place, AI can evolve from an auxiliary tool into an integral component of high-quality, patient-centred orthodontic care.

Declarations

Ethics approval and consent to participate

Not applicable (no primary data were collected).

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

No specific funding was received for this work.

Author Contribution

Authors’ contributionsConceptualization: SMBH, AHBR, RI.Methodology and search strategy: SMBH, AHBR, RH.Screening and data extraction: SMBH, AAM.Risk-of-bias assessment (QUADAS-2 and PROBAST): SMBH, RI, RH.Writing – original draft: SMBH.Writing – review and editing, and supervision: AHBR, RI, RH.All authors read and approved the final manuscript.

References

1. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

2. 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.
3. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36.
4. Moons KG, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170(1):51–8.
5. Collins GS, Dhiman P, Navarro CLA, Ma J, Hooft L, Smidt N, et al. Protocols for reporting AI-based diagnostic accuracy studies (TRIPOD-AI, PROBAST-AI). BMJ Open. 2021;11:e048008.
6. Sounderajah V, Ashrafian H, Deo RC, et al. Developing reporting standards for artificial intelligence in healthcare: QUADAS-AI. Nat Med. 2021;27:1663–5.
7. Cacciamani GE, et al. PRISMA-AI: extensions to improve reporting of scoping reviews involving AI. Nat Med. 2023;29:14–5.
8. Schwendicke F, Golla T, Dreher M, Krois J. Deep learning for cephalometric landmark detection. Clin Oral Investig. 2021;25(10):5721–31.
9. Polizzi A, Leonardi R. Automatic cephalometric landmark identification with artificial intelligence: state of the art. J Dent. 2024;146:105056.
10. Lee JH, Yu HJ, Kim MJ, Kim YJ. Automated cephalometric landmark detection using convolutional neural networks. Am J Orthod Dentofac Orthop. 2019;155(4):620–9.
11. Xue F, Wong RWK, Rabie ABM. Accuracy of AI cephalometric landmark detection compared with human examiners. Angle Orthod. 2020;90(4):535–44.
12. Cui Z, Fang Y, Mei L, Zhang B, Yu B, Liu J, et al. Multi-stage deep learning segmentation of teeth and alveolar bone from CBCT scans. Nat Commun. 2022;13:2096.
13. Arik SÖ, Ibragimov B, Xing L. Fully automated deep learning segmentation of head and neck anatomy. Med Phys. 2017;44(10):5262–75.
14. Park JH, Yu J, Kim J, Kim HY. Automated 3D convolutional neural network segmentation of teeth and bone from CBCT. Dentomaxillofac Radiol. 2021;50(2):20200147.
15. Noeldeke B, Vassis S, Sefidroodi M, Pauwels R, Stoustrup P. Deep learning-based crossbite detection using intraoral photographs. Head Face Med. 2024;20:45.
16. Ryu J, Kim YH, Kim TW, Jung SK. Artificial intelligence-assisted extraction decision-making from intraoral photographs. Sci Rep. 2023;13:5177.
17. Marya A, Inglam S, Chantarapanich N, et al. Predictive modeling of skeletal malocclusion using AI. BMC Oral Health. 2024;24:1064.
18. Huang J, Zhang Y, Wang H, et al. Machine learning models for predicting orthodontic extraction decisions. Front Bioeng Biotechnol. 2024;12:1483230.
19. Etemad LE, et al. Cross-site machine learning prediction of orthodontic extraction decisions. Bioeng (Basel). 2024;11(9):888.
20. Katinas A, et al. Predictive modeling of orthodontic treatment outcomes using AI. Orthod Craniofac Res. 2023;26(2):147–55.
21. Uzel A, et al. Artificial intelligence in orthodontic treatment planning: a scoping review. Eur J Orthod. 2022;44(3):273–82.
22. Pham DD, et al. Deep learning for diagnosis of malocclusion from intraoral photographs. Angle Orthod. 2021;91(6):781–7.
23. Kim Y, et al. AI-assisted orthodontic treatment planning with CBCT data. Korean J Orthod. 2020;50(5):293–302.
24. Lee K, et al. Convolutional neural networks for skeletal malocclusion classification using cephalograms. Diagnostics (Basel). 2020;10(11):930.
25. Choi HI, et al. Predicting orthodontic extractions with machine learning. Sci Rep. 2021;11:22337.
26. Zhang Y, et al. AI prediction of orthodontic treatment outcomes using radiographs. Comput Methods Programs Biomed. 2021;200:105911.
27. Movahed A, et al. Deep learning-based prediction of orthodontic treatment duration. Am J Orthod Dentofac Orthop. 2022;161(4):476–84.
28. Krois J, et al. Machine learning in dental image analysis: a review. J Dent. 2021;103:103583.
29. Yao J, et al. Applications of deep learning in orthodontics: current progress. Orthod Craniofac Res. 2022;25(1):34–42.
30. Kang SH, et al. CNN classification of skeletal malocclusion from cephalograms. Korean J Orthod. 2021;51(2):123–32.
31. Lee J, et al. Deep learning for automatic diagnosis of facial asymmetry. J Clin Med. 2021;10(7):1520.
32. Tanikawa C, et al. Artificial intelligence in orthodontics: recent trends. Jpn Dent Sci Rev. 2021;57:193–200.
33. Sounderajah V, et al. Standards for reporting AI diagnostic accuracy studies (STARD-AI). BMJ Open. 2021;11:e047709.
34. Liu X, Cruz Rivera S, Moher D, et al. CONSORT-AI extension for clinical trials. BMJ. 2020;370:m3164.
35. Krois J, et al. Explainable AI in dentistry: a scoping review. J Dent. 2021;110:103664.
36. Jheon AH, et al. Machine learning and orthodontics: a narrative review. Prog Orthod. 2021;22(1):18.
37. Singh P, et al. Artificial intelligence and big data in orthodontics: challenges and opportunities. Semin Orthod. 2021;27(4):343–50.
38. Abdi AH, et al. Deep learning in dental radiology: a scoping review. Dentomaxillofac Radiol. 2021;50(4):20200175.
39. Liu J, Cruz Rivera S, Moher D, et al. Multi-task deep learning for dental and skeletal classification. Med Image Anal. 2022;76:102313.
40. BMC Oral Health. About the journal. Available from: https://bmcoralhealth.biomedcentral.com/about [Accessed 2025-08-26].
41. BMC Oral Health. Preparing your manuscript. Available from: https://bmcoralhealth.biomedcentral.com/submission-guidelines/preparing-your-manuscript [Accessed 2025-08-26].

Figures

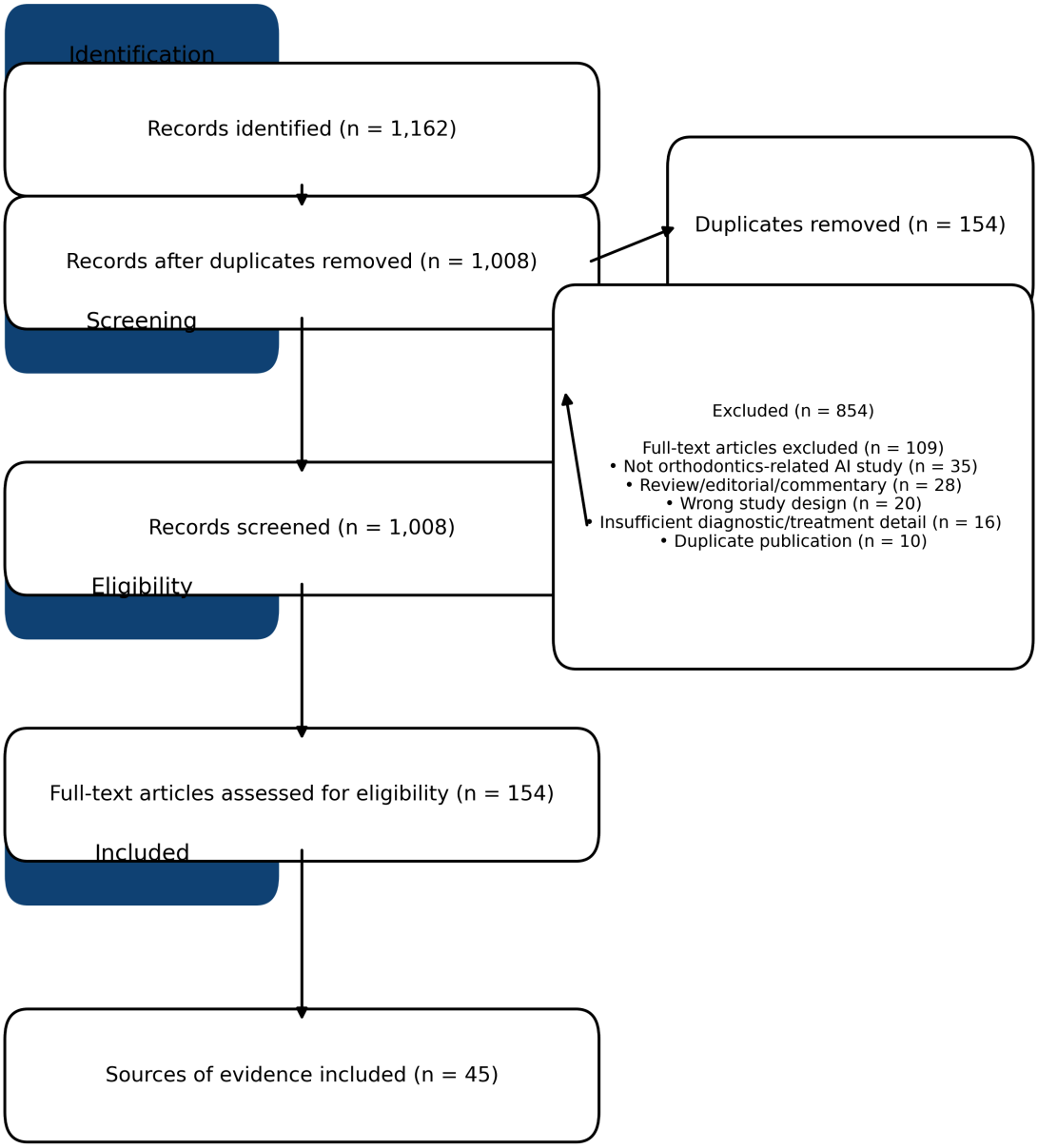

Figure 1

PRISMA-ScR flow diagram of study identification, screening, eligibility, and inclusion.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PRISMAScRChecklist.docx
- S3PROBASTRiskOfBias.docx
- S2QUADAS2RiskOfBias.docx

- S4SearchStrategies.docx