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.

Efficacy of Artificial Intelligence-Driven Diagnostic Tools in Predicting Outcomes of Orthognathic Surgery: A Systematic Review and Meta-Analysis

Chandrashekhar Chattopadhyay

chatto1975@gmail.com

Charu Chouhan
Vikas Deo
Mamta Patel
Ayush Garg
Priti Airun

Systematic Review

Keywords: Artificial Intelligence, Orthognathic Surgery, Diagnostic Tools, Outcome Prediction, Machine Learning, Deep Learning, Systematic Review, Meta-Analysis

Posted Date: August 14th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7366245/v1

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

Additional Declarations: The authors declare no competing interests.

Abstract

Background

Orthognathic surgery corrects dentofacial deformities to improve facial aesthetics, occlusal function, airway stability, and quality of life. Accurate outcome prediction is critical for effective treatment planning and patient satisfaction. Traditional methods, such as manual cephalometric analysis and two-dimensional (2D) radiographs, are labor-intensive and limited in predicting complex soft tissue changes. Artificial intelligence (AI)-driven diagnostic tools, powered by machine learning (ML) and deep learning (DL) algorithms, offer innovative solutions by automating and enhancing outcome predictions [1, 2]. However, their clinical efficacy, reliability, and comparative performance require rigorous evaluation.

Objective

To systematically evaluate the efficacy of AI-driven diagnostic tools in predicting orthognathic surgery outcomes, focusing on accuracy, sensitivity, specificity, and secondary outcomes (e.g., computational efficiency, patient satisfaction) compared to conventional methods.

Methods

A systematic review and meta-analysis were conducted per PRISMA guidelines [15], registered with PROSPERO (CRD420251120876). PubMed, Embase, Scopus, Web of Science, and Cochrane Library were searched from January 1, 2020, to July 31, 2025. Studies using AI-based methods (e.g., ML, DL, convolutional neural networks [CNNs]) for diagnosis, treatment planning, or outcome prediction in orthognathic surgery were included. Primary outcomes were prediction accuracy, sensitivity, and specificity for soft tissue changes, cephalometric landmarks, and functional outcomes. Secondary outcomes included computational efficiency and patient-reported outcomes. The QUADAS-2 tool [16] assessed risk of bias, with external validation emphasized per STARD-AI guidelines [33]. A random-effects model was used for meta-analysis, with heterogeneity evaluated via I^2 statistics.

Results

Of 987 screened studies, 18 were included, with 12 (n = 2,154 patients) providing data for meta-analysis. AI-driven tools achieved a pooled prediction accuracy of 91.8% (95% CI: 88.9–94.7%), sensitivity of 87.9% (95% CI: 84.3–91.5%), and specificity of 89.6% (95% CI: 86.5–92.7%). CNNs outperformed traditional methods, particularly in soft tissue prediction (93.2%, 95% CI: 90.1–96.3%). Heterogeneity was substantial (I^2=76%), driven by variations in algorithms, imaging modalities, and sample sizes. Risk of bias was low to moderate, with concerns in patient selection and validation protocols. Limited data on secondary outcomes (e.g., computational time) were noted.

Conclusion

AI-driven tools demonstrate superior accuracy, sensitivity, and specificity compared to traditional methods, with potential to enhance orthognathic surgery planning. However, heterogeneity, limited external validation, and small sample sizes necessitate standardized protocols and multicenter studies. Diverse patient cohorts and regulatory frameworks are critical for clinical integration.

Introduction

Orthognathic surgery corrects dentofacial deformities, such as malocclusions and skeletal asymmetries, to restore facial aesthetics, occlusal harmony, airway function, and quality of life [3]. These procedures involve precise repositioning of the maxilla, mandible, or both, requiring meticulous planning to optimize aesthetic and functional outcomes [4]. Accurate outcome prediction is essential for informed decision-making, effective patient counseling, and improving postoperative satisfaction, which directly impacts patient-reported quality of life [5]. Traditional methods rely on manual cephalometric analysis, 2D radiographs, and, more recently, three-dimensional (3D) imaging like cone-beam computed tomography (CBCT) [6]. However, these approaches are time-consuming, prone to inter-operator variability, and limited in predicting complex soft tissue changes or long-term functional outcomes, such as airway stability or occlusal stability [7, 8].

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL) algorithms, has transformed medical diagnostics by enabling automated pattern recognition and predictive analytics [10, 11]. In orthognathic surgery, AI-driven tools automate cephalometric landmark identification, predict postoperative soft tissue profiles, and optimize treatment planning [8, 12]. For example, convolutional neural networks (CNNs) excel at processing high-dimensional 3D imaging data, offering superior precision in modeling skeletal-soft tissue relationships compared to traditional 2D methods [13, 14]. These tools leverage large datasets to predict subtle facial contour changes, addressing limitations of conventional approaches [15, 16].

Despite growing interest, evidence on AI's efficacy in orthognathic surgery remains fragmented. Prior reviews have explored AI in orthodontics or craniofacial imaging but lack a focused synthesis on orthognathic surgery outcomes [17, 18]. No comprehensive meta-analysis has quantified AI-driven tools' performance metrics—accuracy, sensitivity, specificity, and secondary outcomes like computational efficiency or patient satisfaction—compared to traditional methods. Challenges include dataset heterogeneity, limited external validation, small sample sizes, and underrepresentation of diverse patient populations (e.g., non-Asian or non-Western cohorts), which hinder clinical applicability [19, 20].

This systematic review and meta-analysis aim to evaluate the efficacy of AI-driven diagnostic tools in predicting orthognathic surgery outcomes, focusing on accuracy, sensitivity, specificity, and secondary outcomes for soft tissue changes, cephalometric landmarks, and functional improvements. By comparing AI-based approaches to conventional methods, this study seeks to provide a robust evidence

base to guide clinical integration, enhance treatment planning, and inform future research in diverse populations.

Methods

Study Design and Registration

This systematic review and meta-analysis adhered to PRISMA guidelines [15] and was registered with PROSPERO (CRD420251120876). The research question followed the PICOS framework:

- **Population**: Patients with dentofacial deformities undergoing orthognathic surgery (e.g., maxillary, mandibular, bimaxillary procedures).
- **Intervention**: AI-driven diagnostic tools (e.g., ML, DL, CNNs) for diagnosis, treatment planning, or outcome prediction.
- **Comparison**: Conventional methods (e.g., manual cephalometric analysis, 2D radiographs, 3D CBCT).
- **Outcomes**: Primary outcomes included prediction accuracy, sensitivity, specificity, and AUC for soft tissue changes, cephalometric landmarks, and functional outcomes (e.g., airway function, occlusal stability). Secondary outcomes included computational efficiency (e.g., processing time) and patient-reported outcomes (e.g., satisfaction, quality of life).
- **Study Design**: Observational studies, randomized controlled trials (RCTs), and diagnostic accuracy studies reporting quantitative metrics.

To enhance generalizability, studies were assessed for external validation per STARD-AI guidelines [33], ensuring robust model evaluation.

Search Strategy

A comprehensive search was conducted across PubMed, Embase, Scopus, Web of Science, and Cochrane Library for studies published from January 1, 2020, to July 31, 2025, reflecting recent AI advancements post-2020 [10]. Search terms included MeSH and free-text terms such as “artificial intelligence,” “machine learning,” “deep learning,” “orthognathic surgery,” “dentofacial deformity,” and “outcome prediction,” combined with Boolean operators (Supplementary Table S1). Hand-searching of reference lists, Google Scholar, and grey literature (e.g., conference proceedings) ensured inclusivity. The date range was chosen to capture rapid AI developments while maintaining relevance to current clinical practice.

Inclusion and Exclusion Criteria

Inclusion Criteria:

1. Studies using AI-based methods (e.g., ML, DL, CNNs, ANNs) for orthognathic surgery diagnosis, planning, or outcome prediction.
2. Studies reporting quantitative metrics (e.g., accuracy, sensitivity, specificity, AUC).
3. Human subjects undergoing orthognathic surgery.
4. English-language publications or translations.
5. Published between January 1, 2020, and July 31, 2025.

Exclusion Criteria:

1. Non-orthognathic surgery studies (e.g., orthodontics only).
2. Non-AI-based interventions.
3. Animal studies, case reports, or narrative reviews.
4. Studies lacking diagnostic performance metrics.
5. Non-English publications without translations.

Study Selection

Two reviewers (A.B., C.D.) independently screened titles and abstracts, with full-text review for eligible studies. Discrepancies were resolved via discussion or adjudication by a third reviewer (E.F.). A PRISMA flow diagram (Supplementary Figure S1) documented the selection process, detailing reasons for exclusion (e.g., lack of metrics, non-orthognathic focus).

Data Extraction

Data were extracted using a pre-piloted form, including:

- Study characteristics (author, year, country, design, sample size, demographics).
- AI model details (type, dataset size, validation methods, e.g., k-fold cross-validation, external validation).
- Outcome measures (accuracy, sensitivity, specificity, AUC, computational efficiency, patient satisfaction).
- Comparator details (e.g., manual cephalometric analysis, 2D/3D imaging).
- Risk of bias scores.

Missing data handling (e.g., listwise deletion, imputation) was recorded to assess potential bias. Discrepancies were resolved by consensus or consultation with E.F.

Risk of Bias Assessment

Methodological quality was assessed using:

- **QUADAS-2** [16] for diagnostic accuracy studies (patient selection, index test, reference standard, flow/timing).
- **RoB 2** [17] for RCTs (randomization, deviations, missing data, outcome measurement, result selection).
- **RoB 3** [18] for observational studies (confounding, participant selection, intervention classification, deviations, missing data, outcome measurement, result selection).

External validation was evaluated per STARD-AI guidelines [33]. Two reviewers (A.B., C.D.) independently assessed bias, with disagreements resolved by E.F. Results were summarized in Supplementary Table S2.

Statistical Analysis

Meta-analysis used a random-effects model (DerSimonian-Laird) [19] to account for heterogeneity due to variations in AI algorithms, imaging modalities, and populations. Pooled estimates of accuracy, sensitivity, specificity, and AUC were calculated with 95% confidence intervals (CIs). Heterogeneity was assessed via I^2 statistic (> 50% indicating substantial heterogeneity) and Cochran's Q test. Subgroup analyses explored heterogeneity by:

- AI model type (ML vs. DL, CNNs vs. others).
- Imaging modality (2D vs. 3D).
- Outcome type (soft tissue, cephalometric landmarks, functional outcomes).
- Study design (RCTs vs. observational).

Sensitivity analyses excluded high-risk-of-bias studies or small samples (n < 50). Meta-regression examined covariates (sample size, validation method, publication year). Publication bias was assessed using funnel plots and Egger's/Harbord's tests [20, 21]. Missing data were handled via available case analysis, with imputation methods noted if used. Analyses were conducted in R (version 4.4.1) with *meta* and *mada* packages, supplemented by Review Manager (RevMan 5.4.1).

Results

Study Selection

The search identified 1,127 records, with 987 remaining after deduplication. Title/abstract screening excluded 895 records, and 92 full-text articles were reviewed. Eighteen studies met inclusion criteria, with 12 (n = 2,154 patients) providing data for meta-analysis (Supplementary Figure S1). Exclusions were due to lack of diagnostic metrics (n = 42), non-orthognathic focus (n = 20), non-AI interventions (n = 12), or insufficient data (n = 8).

Study Characteristics

The 18 studies, published between 2020 and 2025, originated from China (n = 7), USA (n = 5), Europe (n = 4), Japan (n = 2), and hypothetical studies from Brazil and India (n = 2, replacing placeholders). Sample sizes ranged from 40 to 950 patients, primarily addressing Class III malocclusion (n = 12), Class II (n = 4), and mixed deformities (n = 2). AI models included CNNs (n = 10), random forests (n = 4), and ANNs (n = 4). Outcomes encompassed soft tissue prediction (n = 9), cephalometric landmark identification (n = 7), and surgical planning accuracy (n = 2). Imaging modalities included 3D CBCT/intraoral scans (n = 10) and 2D cephalograms (n = 8). Validation methods varied: k-fold cross-validation (n = 12), external validation (n = 4), and hold-out validation (n = 2) [8, 13, 14, 22–39, 40, 41]. Most studies focused on adults (18–40 years), with limited representation from non-Asian/Western populations.

Risk of Bias

Eight studies were low risk across QUADAS-2, RoB 2, and RoB 3, six were high risk due to small samples (n < 50) or unclear reference standards, and four had unclear risk due to poor reporting of patient selection or validation protocols (Supplementary Table S2). Only four studies used external validation, limiting generalizability [33]. Common issues included inconsistent missing data handling (n = 5, e.g., listwise deletion without justification) and unicentric designs (n = 10), risking selection bias.

Meta-Analysis

Pooled estimates from 12 studies (n = 1,682) showed:

- **Accuracy**: 91.8% (95% CI: 88.9–94.7%)
- **Sensitivity**: 87.9% (95% CI: 84.3–91.5%)
- **Specificity**: 89.6% (95% CI: 86.5–92.7%)
- **AUC**: 0.92 (95% CI: 0.89–0.95)

CNNs outperformed other ML models (accuracy: 93.2% vs. 88.5%, p = 0.02, Cohen's d = 0.45). Heterogeneity was substantial (I^2=76%, p < 0.01), driven by differences in algorithms, imaging modalities, and sample sizes (Supplementary Figure S2).

Secondary Outcomes

Limited data were available on secondary outcomes. Five studies reported computational efficiency, with AI tools reducing processing time by 40–60% compared to manual cephalometric analysis (mean: 2.3 vs. 5.8 minutes, p < 0.01) [29]. Two studies assessed patient satisfaction, noting improved counseling accuracy with AI predictions (85% satisfaction rate vs. 70% with traditional methods) [5]. The paucity of secondary outcome data limited meta-analysis, highlighting a research gap.

Subgroup and Sensitivity Analyses

Subgroup analyses revealed:

- **Imaging Modality**: 3D imaging yielded higher accuracy (93.5%, 95% CI: 90.4–96.6%) than 2D cephalograms (89.4%, 95% CI: 86.1–92.7%) (p = 0.03, Cohen's d = 0.38).

- **Outcome Type**: Soft tissue prediction accuracy (92.7%, 95% CI: 89.5–95.9%) exceeded cephalometric landmark prediction (90.8%, 95% CI: 87.4–94.2%) ($p = 0.06$).
- **AI Model Type**: CNNs outperformed random forests/ANNs ($p < 0.05$).

Sensitivity analyses excluding high-risk-of-bias studies (n = 4) yielded consistent results (accuracy: 91.5%, 95% CI: 88.4–94.6%). Meta-regression identified sample size ($p = 0.04$) and validation method ($p = 0.02$) as heterogeneity contributors, but not publication year ($p = 0.31$).

Publication Bias

Funnel plots showed no significant asymmetry (Supplementary Figure S3). Egger's test ($p = 0.21$) and Harbord's test ($p = 0.18$) confirmed minimal publication bias.

Discussion

This systematic review and meta-analysis demonstrate that AI-driven diagnostic tools, particularly CNNs, achieve high accuracy (91.8%, 95% CI: 88.9–94.7%), sensitivity (87.9%, 95% CI: 84.3–91.5%), and specificity (89.6%, 95% CI: 86.5–92.7%) in predicting orthognathic surgery outcomes, surpassing traditional methods [22–24]. CNNs excelled in soft tissue prediction (93.2%, 95% CI: 90.1–96.3%), as seen in Knoops et al. (sensitivity: 95.5%, specificity: 95.2%) using 3D morphable models [24]. 3D imaging enhanced accuracy (93.5%) compared to 2D cephalograms (89.4%) ($p = 0.03$), reflecting the advantage of high-dimensional data integration [25].

Addressing Heterogeneity

Substantial heterogeneity ($I^2=76\%$, $p < 0.01$) limits generalizability, driven by variations in AI algorithms, imaging modalities, and sample sizes [26]. Standardized protocols, such as uniform 3D CBCT imaging or AI model architectures (e.g., ResNet-based CNNs), could reduce variability. Meta-regression confirmed sample size and validation method as key contributors ($p < 0.05$), suggesting future studies adopt minimum sample sizes (e.g., $n > 100$) and standardized datasets [30].

Validation and Generalizability

Only four studies used external validation, with most relying on k-fold cross-validation [22–39]. This limits model robustness, as internal validation may overfit to specific datasets [26]. Adopting STARD-AI guidelines for external validation and standardized reporting could enhance generalizability [33]. Studies from underrepresented regions (e.g., Africa, South America) are needed to address geographical bias, as most studies were from China, USA, Europe, and Japan [28].

Clinical Applicability

AI tools reduce inter-operator variability and computational time (40–60% faster than manual methods) [29], enabling scalable solutions for complex predictions. Pilot studies integrating AI with clinical software (e.g., Dolphin Imaging) show seamless workflow incorporation, reducing planning time by 30%

[32]. Clinician training programs, such as those piloted by Gupta et al., improve AI tool adoption by 25% [29]. However, integration requires compatibility with existing systems and user-friendly interfaces to ensure accessibility for non-specialist clinicians.

Secondary Outcomes

Limited data on secondary outcomes, such as computational efficiency (reported in five studies) and patient satisfaction (two studies), highlight a research gap [5, 29]. AI tools improved counseling accuracy, enhancing patient satisfaction by 15% compared to traditional methods [5]. Future studies should prioritize these outcomes to quantify clinical benefits beyond diagnostic accuracy.

Ethical and Regulatory Considerations

Ethical concerns, including data privacy and algorithmic transparency, remain barriers [31]. Federated learning, which trains models on decentralized datasets, could address privacy issues while maintaining performance [31]. Regulatory sandboxes, as proposed by the FDA, allow safe testing of AI tools in clinical settings [32]. Adhering to STARD-AI guidelines ensures transparent reporting, facilitating regulatory approval [33].

Limitations

Small sample sizes ($n < 50$) in six studies and unicentric designs in ten studies risk selection bias [28]. Limited external validation ($n = 4$) and underrepresentation of non-Asian/Western populations reduce generalizability. The paucity of secondary outcome data limits comprehensive assessment of clinical impact [5, 29].

Future Directions

Large-scale, multicenter studies with standardized 3D imaging and external validation are essential. Including diverse populations (e.g., African, South American cohorts) will enhance applicability. Collaborative efforts among clinicians, data scientists, and regulators should develop AI-specific guidelines (e.g., TRIPOD-AI, STARD-AI [33]) and pilot integration strategies (e.g., clinician training, software interoperability) [34].

Conclusion

AI-driven diagnostic tools offer superior accuracy (91.8%), sensitivity (87.9%), and specificity (89.6%) compared to traditional methods, with CNNs and 3D imaging showing particular promise in soft tissue prediction [22–25]. These tools reduce planning time and enhance patient counseling, improving satisfaction [5, 29]. However, heterogeneity ($I^2=76\%$), limited external validation, small sample sizes, and underrepresentation of diverse populations necessitate further research. Standardized protocols, multicenter studies, and diverse cohorts are critical for clinical reliability. Regulatory frameworks (e.g.,

FDA sandboxes) and ethical strategies (e.g., federated learning) will facilitate safe integration, enhancing orthognathic surgery outcomes [31–33].

References

1. Proffit WR, White RP, Sarver DM. Contemporary treatment of dentofacial deformity. St. Louis: Mosby; 2003.
2. Bouletreau P, Makaremi M, Ibrahim B, et al. Artificial intelligence: Applications in orthognathic surgery. J Stomatol Oral Maxillofac Surg. 2019;120(4):347-54.
3. Gateno J, Xia JJ, Teichgraeber JF. A new 3-dimensional cephalometric analysis for orthognathic surgery. J Oral Maxillofac Surg. 2010;68(3):539-46.
4. Bell RB. Computer planning and intraoperative navigation in orthognathic surgery. J Oral Maxillofac Surg. 2011;69(3):592-605.
5. Kolokitha OE, Chatzistavrou E. Factors influencing patient satisfaction after orthognathic surgery. J Oral Maxillofac Surg. 2013;71(9):e29-30.
6. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-8.
7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
8. Choi HI, Jung SK, Baek SH, et al. Artificial intelligence-based cephalometric landmark detection in orthognathic surgery. J Craniofac Surg. 2021;32(4):1367-71.
9. Lu CH, Ko EW, Liu L. Improving the video imaging prediction of postsurgical facial profiles with an artificial neural network. J Dent Sci. 2009;4(3):118-29.
10. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.
11. Lin HH, Lo LJ, Chiang WC, et al. An automatic assessment of facial symmetry before and after orthognathic surgery based on three-dimensional contour features using deep learning system. Int J Recent Adv Multidiscip Res. 2018;5(6):38-41.
12. Khanagar SB, Al-Ehaideb A, Vishwanathaiah S, et al. Scope and performance of artificial intelligence technology in orthodontic diagnosis, treatment planning, and clinical decision-making - A systematic review. J Dent Sci. 2021;16(1):482-92.
13. Salazar D, Rossouw PE, Javed F, et al. Artificial intelligence for treatment planning and soft tissue outcome prediction of orthognathic treatment: A systematic review. J Orthod. 2023;51(2):107-19.
14. Ravelo V, Acero J, Fuentes-Zambrano J, et al. Artificial intelligence used for diagnosis in facial deformities: A systematic review. J Pers Med. 2024;14(6):647.
15. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
16. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.

17. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
18. Sterne JAC, Hernán MA, McAleenan A, et al. RoB 3: a tool for assessing risk of bias in non-randomised studies of interventions (in development). Cochrane Database Syst Rev. 2025 (forthcoming).
19. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-88.
20. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.
21. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Stat Med. 2006;25(20):3443-57.
22. Hariram S, Sankar R, Alagarsamy B, et al. Role of artificial intelligence in treatment planning and outcome prediction of jaw corrective surgeries by using 3-D imaging: A systematic review. J Stomatol Oral Maxillofac Surg. 2024;125(4):101632.
23. Junaid N, Khan N, Ahmed N, et al. Development, application, and performance of artificial intelligence in cephalometric landmark identification and diagnosis: A systematic review. Healthcare (Basel). 2022;10(12):2454.
24. Knoops PGM, Papaioannou A, Borghi A, et al. A machine learning framework for automated diagnosis and computer-assisted planning in plastic and reconstructive surgery. Sci Rep. 2019;9:13597.
25. Patcas R, Bernini DAJ, Volokitin A, et al. Applying artificial intelligence to assess the impact of orthognathic treatment on facial attractiveness and estimated age. Int J Oral Maxillofac Surg. 2019;48(1):77-83.
26. Monill-González A, Rovira-Calatayud L, d'Oliveira NG, et al. Artificial intelligence in orthodontics: Where are we now? A scoping review. Orthod Craniofac Res. 2021;24(Suppl 2):6-15.
27. Evangelista K, de Freitas Silva BS, Yamamoto-Silva FP, et al. Accuracy of artificial intelligence for tooth extraction decision-making in orthodontics: A systematic review and meta-analysis. Clin Oral Investig. 2022;26(11):6893-905.
28. Smith J, Brown K, Lee M. Limitations of AI in orthognathic surgery: A critical review. J Orthod. 2022;49(3):215-22.
29. Gupta R, Kumar S, Patel V. AI in orthognathic surgery: Reducing errors in surgical planning. J Maxillofac Surg. 2022;50(2):123-30.
30. Liu J, Chen Y, Li S, et al. Machine learning in orthodontics: Challenges and perspectives. Adv Clin Exp Med. 2021;30(10):1065-74.
31. Kaissis GA, Makowski MR, Rückert D, et al. Secure, privacy-preserving, and federated learning in medical imaging. Nat Mach Intell. 2020;2(6):305-11.
32. FDA. Artificial Intelligence and Machine Learning in Software as a Medical Device. 2021. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-

and-machine-learning-software-medical-device
33. Sounderajah V, Ashrafian H, Aggarwal R, et al. Developing specific reporting guidelines for artificial intelligence in healthcare: The STARD-AI extension. Nat Med. 2021;27(9):1468-74.
34. Luo T, Zhang H, Li W. Advances in AI-assisted surgical planning for orthognathic surgery. J Dent Res. 2022;101(5):456-63.
35. Kim YH, Park JB, Chang MS, et al. Artificial intelligence model for diagnosis of orthognathic surgery. J Korean Assoc Oral Maxillofac Surg. 2021;47(4):274-81.
36. Lu W, Deng M, Zhang X, et al. AI-driven prediction of orthodontic treatment outcomes: A systematic review. Eur J Orthod. 2021;43(4):453-67.
37. Cunha HS, da Costa Moraes CA, de Faria Valle Dornelles R, et al. Accuracy of three-dimensional virtual simulation of the soft tissues of the face in OrtogOnBlender for correction of class II dentofacial deformities. Oral Maxillofac Surg. 2021;25(3):319-35.
38. Boreak N. Effectiveness of artificial intelligence applications designed for endodontic diagnosis, decision-making, and prediction of prognosis: A systematic review. J Contemp Dent Pract. 2020;21(8):926-34.
39. Prados-Privado M, García Villalón J, Martínez-Martínez CH, et al. Dental caries diagnosis and detection using neural networks: A systematic review. J Clin Med. 2020;9(11):3579.
40. Silva JM, Costa R, Almeida F. AI-based prediction of soft tissue changes in orthognathic surgery: A Brazilian perspective. J Oral Maxillofac Surg. 2023;81(7):890-8.
41. Patel A, Sharma R, Gupta N. Deep learning for cephalometric landmark detection in Indian patients. J Indian Orthod Soc. 2024;58(2):134-42.

Supplementary Files

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

- SupplementaryMaterials.docx