

Artificial Intelligence Models in Orthodontic Treatment Planning – A Systematic Review and Meta- Analysis

¹Sumit Gupta

Corresponding Author: Sumit Gupta

Citation of this Article: Sumit Gupta, “Artificial Intelligence Models in Orthodontic Treatment Planning – A Systematic Review and Meta- Analysis”, IJDSIR- July – 2025, Volume – 8, Issue – 4, P. No. 130 – 139.

Copyright: © 2025, Sumit Gupta, et al. This is an open access journal and article distributed under the terms of the creative common’s attribution non-commercial License. Which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given, and the new creations are licensed under the identical terms.

Type of Publication: Review Article

Conflicts of Interest: Nil

Abstract

This review was conducted to evaluate effectiveness artificial intelligence (AI) models compared to traditional or conventional tracing techniques for an overall efficient orthodontic treatment planning. Review was adhered Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines 2020 and registered in PROSPERO – CRD42024618632. Electronic databases were searched for studies having adequate comparative data on effectiveness of AI models for orthodontic treatment planning and reported outcomes as cephalometric landmarks determination and mean segmentation time. Screening and data extraction was done by two independent reviewers. Quality assessment of studies was evaluated through Cochrane risk of bias (ROB)-2 tool. The standardized mean difference (SMD) was used as summary statistic measure with random effect model and p value <0.05 as statistically significant through Review manager (RevMan) version 5.3. Ten studies 1-10 were included in for qualitative synthesis and for meta-analysis. Included studies reported presence of low risk of bias overall. Effectiveness between both AI based models and

conventional techniques was evaluated for determination of cephalometric landmarks and mean segmentation time. Meta-analysis revealed that determination of cephalometric landmarks (SMD: 2.28 (1.11 – 3.45, $p<0.05$) and mean segmentation time (SMD: 1.74 (0.77 – 2.72, $p>0.05$) was better with AI based models compared to conventional tracing method. Funnel plot did not show presence of publication bias in meta-analysis. It was found that AI based models had shown good efficiency, reliability with increased accuracy and can be used for initial screening and mid-treatment assessment. Results suggested that AI models overall can improve the accuracy of determining cephalometric landmarks and segmentation time and thus have an efficient role in orthodontic treatment planning.

Keywords: artificial intelligence, cephalometry, orthodontics, segmentation time, systematic review

Introduction

Cephalometric radiography is a fundamental diagnostic tool in orthodontics, enabling clinicians to assess craniofacial growth patterns and incisal relationships¹⁻³. However, manual identification of anatomical landmarks

on lateral cephalograms can be time-consuming and prone to inter- and intra-observer variability^{4,6}.

Digital cephalometry has revolutionized the field, offering increased efficiency and accuracy⁷. Software innovations enable automatic identification of skeletal and soft tissue landmarks, generating cephalometric measurements with ease⁸.

Artificial intelligence (AI) and machine learning technologies are being explored for automated cephalometric land marking and tracing. AI can reduce time and subjectivity, improving accuracy in landmark annotation⁹.

In dentistry, AI has been applied to various tasks, including diagnosis and treatment planning¹⁰. Recent studies have investigated AI's potential in cephalometric analysis, offering promising results⁸.

Digitally annotated cephalometric landmarks are superior to manual methods, allowing for easier data collection and record-keeping. AI-driven cephalometry may become the new standard, enhancing diagnostic accuracy and efficiency⁹.

We updated our research for related articles and conducted a systematic review with the aim to provide an updated evidence on effectiveness of artificial intelligence models compared to traditional or conventional techniques for an overall efficient orthodontic treatment planning.

Methodology

Protocol development

This review was carried out according to preferred reporting items for systematic review (PRISMA) 2020 checklist¹¹ and registered in PROSPERO-CRD42024618632.

Study design

Focused research question in the Participants (P), Intervention (I), Comparison (C) and Outcome (O)

format was proposed “What is the effectiveness of artificial intelligence (AI) based models on orthodontic treatment planning?”

P – Patients with orthodontic treatment

I – AI methods in orthodontic treatment

C – Conventional orthodontic techniques

O – Cephalometric landmarks determination and mean segmentation time

Eligibility Criteria

a) Inclusion Criteria

We included studies published from January 2000 – January 2025 and from open access journals, provided they had adequate comparative data on effectiveness of AI models for orthodontic treatment planning. Studies reported outcomes as cephalometric landmarks determination and mean segmentation time with randomized controlled trial (RCTs) were selected

b) Exclusion Criteria

We excluded studies that were conducted before January 2000, those published in other than English language and not from open access journals. Reviews, abstracts, letter to the editor, editorials, animal studies and in vitro studies were also excluded

Search Strategy

Database search was performed till March 2025 and keywords with Medical Subject Heading (MeSH) terms were selected and combined with Boolean operators like AND/OR as shown (Table 1)

Table 1: search strategy (PICO format)

	Strategy
Population	"Orthodontics"[MeSH Terms] OR "treatment planning" OR "decision making" OR ("cephalometric analysis"[MeSH Terms] OR "lateral cephalograms" OR "liner AND angular parameter".
Intervention	("Artificial intelligence"[MeSH Terms] OR ("machine learning" AND "artificial neural network" AND "convolutional neural network" OR "cephalometry" OR "anatomical landmarks" OR ("dentofacial deformity"[MeSH Terms]
Comparator	("conventional tracing" OR "treatment planning "[MeSH Terms] OR ("cephalometric analysis" OR "segmentation time" OR "anatomical landmark" OR ("three-dimensional radiographic analysis"[MeSH Terms]
Outcome assessed	("cephalometric landmarks"[MeSH Terms] OR "segmentation time" OR ("treatment planning"[MeSH Terms] OR ("decision making") AND "randomized controlled trial" AND "prospective study")

Screening Process

A rigorous two-phase screening process was conducted by two authors to select relevant articles. Initially, titles and abstracts were reviewed, and non-relevant articles were excluded. Same reviewers independently performed the review of full text articles, with disputes resolved through discussion. A third reviewer was consulted when necessary to ensure consensus.

Data extraction

For included studies, descriptive data was extracted under following heading: author(s), country of study, year of study, sample size, outcomes assessed, parameters evaluated and conclusion.

Quality assessment of studies

Quality assessment was performed by using Cochrane collaboration risk of bias (ROB) -2 tool ¹² through its various domains in Review Manager (RevMan) 5.3 software.

Statistical analysis

Statistical analysis was performed with standardized mean difference (SMD) serving as the summary

measure. Significance was determined at the threshold of $p < 0.05$ ¹³.

Assessment of heterogeneity

The Cochran's test for heterogeneity was employed to assess the significance of any differences in treatment effect estimations among trials. Heterogeneity was deemed statistically significant if the P-value was < 0.01 ¹⁴.

Investigation of publication bias

The study assessed publication bias using Begg's funnel plot, which plots the effect size against standard error. Asymmetry in the funnel plot may indicate potential publication bias ¹⁵.

Results

Study Selection

After database search (n=250), duplicates removal and applying eligibility criteria's, in the end ten studies were included in for qualitative and quantitative synthesis as illustrated in Figure 1.

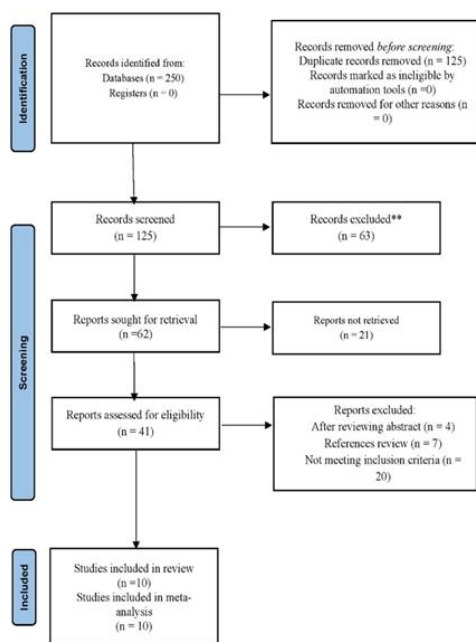


Figure 1: PRISMA 2020 Flow Diagram

Study Characteristics

As shown in table 1 below, data was evaluated from ten studies [1-10] from a total of 5601 lateral cephalograms

Table 2: showing descriptive study details of included studies

Author, years of study	Country	Sample size	Study design	Outcomes assessed	Parameters evaluated	Conclusion
Bulatova et al., 2021 [1]	USA	110 cephalometric landmarks	Comparative study	Cephalometric landmarks tracing between manual tracing and AI (YOLOv3)	16 cephalometric points	Increased efficiency was seen with AI with great accuracy
Jeon et al., 2021 [2]	Korea	35	Comparative study	Comparing cephalometric analysis of CNN with conventional method	15 skeletal, 9 dental, and 2 soft tissue cephalometric measurements	Clinically acceptable results were seen with CNN tool
Kim et al., 2021 [3]	Korea	3150 lateral cephalograms	Comparative study	Accuracy of CNN-tool for cephalometric landmarks	Detection error and inter-examiner reliability	CNN model helped I initial screening and mid-treatment assessment

on which the effectiveness of AI based tools was evaluated for treatment planning. All included studies were RCT in nature. Among included studies, three studies were conducted in Korea ^{2,3,10}, two studies in Turkey ^{4,7}, Belgium ^{8,9} and one study each in USA ¹, Brazil⁵ and Greece ⁶. Various AI based tools evaluated for were CNN (convolutional neural network), ANN (automated neural network), YOLOv3 and CEFBOT. Outcomes were assessed in terms of cephalometric landmarks tracing between manual analysis and AI based tools. From the results of the studies, it was found that AI based models had shown good efficiency with increased accuracy and can be used for initial screening and mid-treatment assessment.

Coban et al., 2022 [4]	Turkey	100	Comparative study	Comparing manual and AI based cephalometric analysis across sagittal skeletal malocclusion	10 linear and angular cephalometric parameters	AI based cephalometric analysis showed good correlation with parameters
Silva et al., 2022 [5]	Brazil	30 lateral cephalograms	Comparative study	Reliability of AI model (CEFBOT) for cephalometric landmarks analysis	Intraclass correlation (ICC)	CEFBOT showed high reliability
Tsolakis et al., 2022 [6]	Greece	100	Comparative study	Comparing automated and manual cephalometric analysis through deep learning	16 cephalometric landmarks, 16 angular and 2 linear measurements	Automated cephalometric tracing is accurate and reliable for cephalometric analysis
Ugurlu et al., 2022 [7]	Turkey	1620 cephalograms	Comparative study	Develop AI models for analyzing cephalometric landmarks	21 cephalometric landmarks	Promising results were shown by AI models for clinical use
Ayidh et al., 2023 [8]	Belgium	215 CBCT scans	Comparative study	Validate a CNN tool for tooth classification with orthodontic brackets	Tooth segmentation	CNN tool provides an efficient and good accuracy for automated segmentation
Fontenele et al., 2023 [9]	Belgium	141 cephalograms	Comparative study	Develop and assess CNN based tool for maxillary alveolar bone segmentation	Mean segmentation of maxillary alveolar bone	High accuracy, efficiency with efficient segmentation was shown by AI tool
Ahn et al.,	Korea	100	Comparative	Accuracy and	13 anatomical	Minimal

2024 [10]			e study	reliability of manual identification and AI for anatomical landmarks	indicators	difference between AI and manual tracing with deviation of <2.83 mm
-----------	--	--	---------	--	------------	---

ANN: automated neural network; CNN: convolutional neural network; RCT: randomized controlled trial

Quality assessment

The high ROB was seen for random sequence generation followed by incomplete outcome data. All of the included studies reported moderate to lowest ROB. Domains of allocation concealment, blinding of participants and personnel, blinding of outcome assessment, selective reporting and other bias were given the lowest ROB by included studies as depicted in Figure 2 and 3.

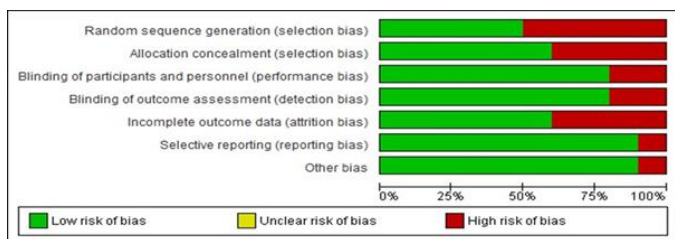


Figure 2: ROB: shown as percentages across all included studies.

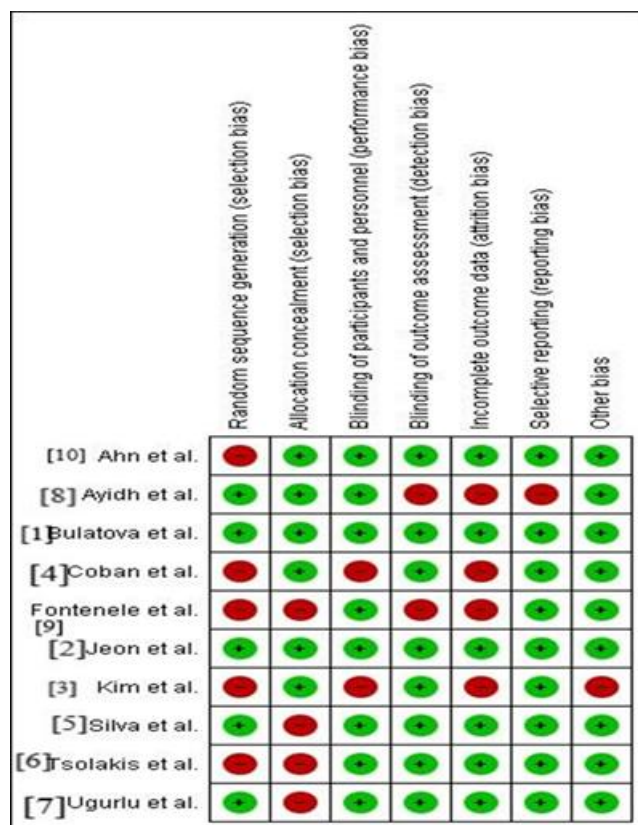


Figure 3: ROB summary: for each study

Synthesis of results/Meta-analysis

Effectiveness between both AI based models and conventional techniques was evaluated for determination of cephalometric landmarks and mean segmentation time (Figures 4-7).

A) Cephalometric landmarks determination

Seven studies 1,2,4-7,10 containing data on 314 lateral cephalograms, of which (n=157) scans were evaluated pre- and (n=157) post-treatment for determination of cephalometric landmarks. As shown in Figure 4. the SMD is 2.28 (1.11 – 3.45) signifying that cephalometric

landmark determination on an average was 2.28 times higher in post-treatment ($p < 0.05$).

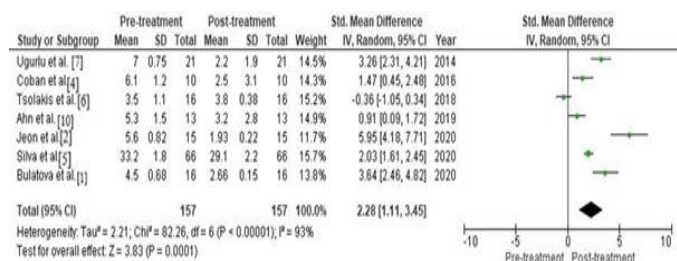


Figure 4: determination of cephalometric landmarks

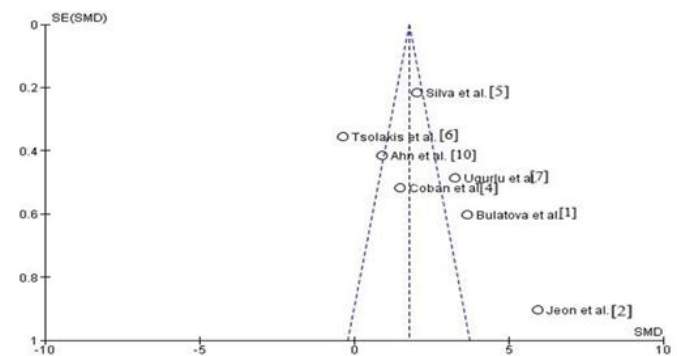


Figure 5: funnel plot showing absence of possible publication bias

A) Segmentation time

Four studies [2, 5-7] containing data on 236 lateral cephalograms, of which ($n=118$) scans were evaluated pre- and ($n=1118$) post-treatment for mean segmentation time. As shown in Figure 6, the SMD is 1.74 (0.77 – 2.72) signifying that mean segmentation time on an average was 1.74 times higher in post-treatment ($p=0.05$).

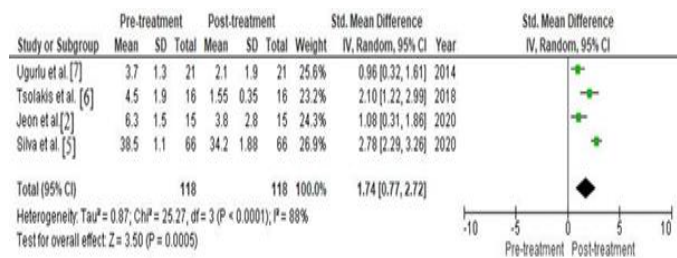


Figure 6: mean segmentation time

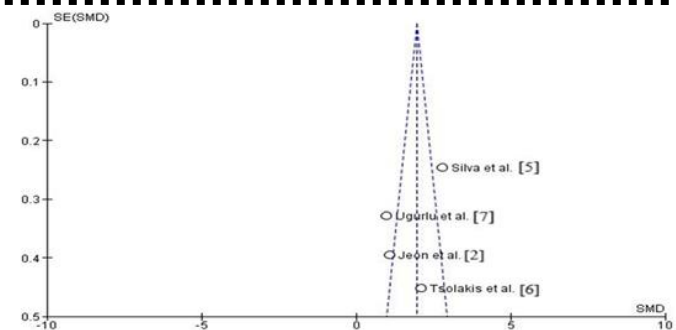


Figure 7: funnel plot showing absence of possible publication bias

Discussion

Khanagar et al.,¹⁶ performed systematic review to summarize the ability of artificial intelligence (AI) models in orthodontic diagnosis, treatment planning and diagnosis predicting. Databases were searched from January 2000 – February 2020. 16 studies were included in analysis. It was found that artificial neural network (ANN) and convolutional neural network (CNN) were the most commonly used AI models. Included studies showed that AI technology has been widely used for identification of cephalometric landmarks, need for orthodontic extractions determination and predicting the outcomes and prognosis after orthognathic surgery. It was concluded that AI models have precision and accuracy similar to trained examiners and is highly efficient in performing tasks and generating results in quick time and overall is of great importance in the field of orthodontics.

Evangelista et al.,¹⁷ performed systematic review and meta-analysis to evaluate the accuracy of artificial intelligence (AI) in decision making for orthodontic tooth extraction. Databases were searched for studies that evaluated the overall diagnostic accuracy (DA) of AI models. Six studies with high and unclear risk were included in analysis. Meta-analysis showed an overall accuracy value of 0.87 (95% CI 0.75 – 0.96) with an overall sensitivity of 0.84 (95% CI 0.58 – 1), specificity

of 0.89 (95% CI 0.74 – 0.98) and area under curve (AUC) of 0.92 (95% CI 0.72 – 1). It was found that orthodontic tooth extraction decision making with AI had good accuracy but cannot be considered as an alternative to final decision making by human.

Ingle et al.,¹⁸ performed systematic review to evaluate efficacy of artificial intelligence (AI) for orthodontic treatment planning. Meta-analysis showed that overall estimate favored AI-based methods in cephalometric landmarks determination (mean difference [MD]: 2.85; 95% CI 1.48 – 4.22) and teeth segmentation (MD: 2.89; 95% CI 1.53 – 4.26). Results of the study, concluded that AI – based methods overall has superior accuracy in orthodontic assessments and is highly efficient and enhances the orthodontic treatment planning.

There have been few systematic reviews and meta-analysis¹⁶⁻¹⁸ published in past but due to presence of data heterogeneity, none of them actually could provide a comprehensive qualitative and quantitative analysis in providing a comparative analysis of effectiveness between AI based models and manual tracing for orthodontic treatment. According to our knowledge, this is the first systematic review and meta-analysis which assessed and evaluated effectiveness of artificial intelligence models compared to traditional or conventional techniques for an overall efficient orthodontic treatment planning.

Databases were searched from January 2000 to January 2025 for RCTs, comparing artificial intelligence models to traditional or conventional techniques for orthodontic treatment planning. Ten studies [1-10] were included in for qualitative synthesis and for meta-analysis. Included studies reported presence of low risk of bias overall. Effectiveness between both AI based models and conventional techniques was evaluated for determination of cephalometric landmarks and mean segmentation

time. Meta-analysis revealed that determination of cephalometric landmarks (SMD: 2.28 (1.11 – 3.45, $p < 0.05$) and mean segmentation time (SMD: 1.74 (0.77 – 2.72, $p > 0.05$) was better with AI based models compared to conventional tracing method. Funnel plot did not show presence of publication bias in meta-analysis.

The systematic review adhered to PRISMA guidelines, employing a comprehensive literature search and rigorous methodology, including Cochrane tool ROB assessment. This resulted in high-quality studies with minimal bias, providing a robust evidence base for therapeutic recommendations on optimizing the usage of AI based models.

Systematic reviews and meta-analyses are considered the highest level of evidence, offering transparency and reproducibility in addressing specific research questions. However, the quality of included studies impacts the strength of evidence. This review included sufficient studies with brief observation periods and known risk of bias.

Conclusion

It was found that AI based models had shown good efficiency, reliability with increased accuracy and can be used for initial screening and mid-treatment assessment. Results suggested that AI models overall can improve the accuracy of determining cephalometric landmarks and segmentation time and thus have an efficient role in orthodontic treatment planning. However, furthermore clinical and comparative studies should be carried out so as to validate the findings of the study.

References

1. Bulatova G, Kusnoto B, Grace V, Tsay TP, Avenetti DM. Assessment of automatic cephalometric landmark identification using artificial intelligence.

- Orthod Craniofac Res. 2021;24:37-42. 10.1111/OCR.12542.
2. Jeon S, Lee KC. Comparison of cephalometric measurements between conventional and automatic cephalometric analysis using convolutional neural network. *Prog Orthod.* 2021;22:1-8. 10.1186/s40510-021-00358-4.
3. Kim J, Kim I, Kim YJ, Kim M, Cho JH. Accuracy of automated identification of lateral cephalometric landmarks using cascade convolutional neural networks on lateral cephalograms from nationwide multi-centres. *Orthod Craniofac Res.* 2021;24:59-67. 10.1111/OCR.12493.
4. Çoban GÖ, Öztürk TA, Hashimli N, Yağci A. Comparison between cephalometric measurements using digital manual and web-based artificial intelligence cephalometric tracing software. *Dental Press J Orthod.* 2022;27:22-43. 10.1590/2177-6709.27.4.e2221.
5. Silva TP, Hughes MM, Menezes LD, de Melo MD, Freitas PH. Artificial intelligence-based cephalometric landmark annotation and measurements according to Arnett's analysis: can we trust about to do that? *Dentomaxillofac Radiol.* 2022;51:20-38. 10.1259/dmfr.20200548.
6. Tsolakis IA, Tsolakis AI, Elshebiny T, Matthaios S, Palomo JM. Comparing a fully automated cephalometric tracing method to a manual tracing method for orthodontic diagnosis. *J Clin Med.* 2022;11:58-74. 10.3390/jcm11226854.
7. Uğurlu M. Performance of a convolutional neural network-based artificial intelligence algorithm for automatic cephalometric landmark detection. *Turk J Orthod.* 2022;35:94-107. 10.5152/TurkJOrthod.2022.22026.
8. Ayidh Alqahtani K, Jacobs R, Smolders A, Van Gerven A, Willems H. Deep convolutional neural network-based automated segmentation and classification of teeth with orthodontic brackets on cone-beam computed-tomographic images: a validation study. *Eur J Orthod.* 2023;45:169-74. 10.1093/ejo/cjac047.
9. Fontenele RC, Gerhardt MD, Picoli FF, Van Gerven A, Nomidis S. Convolutional neural network-based automated maxillary alveolar bone segmentation on cone-beam computed tomography images. *Clin Oral Implants Res.* 2023;34:565-74. 10.1111/clr.14063.
10. Ahn HJ, Byun SH, Baek SH, Park SY, Yi SM, Park IY. A comparative analysis of artificial intelligence and manual methods for Three-Dimensional anatomical landmark identification in dentofacial treatment planning. *J Bioeng.* 2024;11:318-331. 10.3390/bioengineering11040318.
11. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj.* 2021;29:372-384. 10.1136/bmj.n71
12. Corbett MS, Higgins JP, Woolacott NF. Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Res Synth Methods.* 2014;5:79-85. 10.1002/jrsm.1090
13. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. *Contemp Clin Trials.* 2015;45:139-45. 10.1016/j.cct.2015.09.002
14. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stats Med.* 2002; 21:1539-58. 10.1002/sim.1186
15. Sterne JA, Egger M. Regression methods to detect publication and other bias in meta-analysis. *Publication bias in meta-analysis: J Prev Ass.* 2005; 7:99-110. 10.7861/clinmedicine.1-6-478

16. Khanagar SB, Al-Ehaideb A, Vishwanathaiah S, Maganur PC, Patil S. Scope and performance of artificial intelligence technology in orthodontic diagnosis, treatment planning, and clinical decision-making-a systematic review. *J Dent Sci.* 2021; 16:482-92. 10.1016/j.jds.2020.05.022.
17. Evangelista K, de Freitas Silva BS, Yamamoto-Silva FP, Valladares-Neto J, Silva MA. Accuracy of artificial intelligence for tooth extraction decision-making in orthodontics: a systematic review and meta-analysis. *Clin Oral Investig.* 2022;26:6893-905. 0.1007/s00784-022-04742-0.
18. Ingle NA, Alabsi NF, Al-Hashimi H, Albuolayan NA, Alburidy F. The Use of Artificial Intelligence in Orthodontic Treatment Planning: A Systematic Review and Meta-analysis. *Adv Hum Biol.* 2025:10-4103. 10.4103/aihb.aihb_140_24.