International Platform of Registered Systematic Review and Meta-analysis Protocols

INPLASY

INPLASY202660101

doi: 10.37766/inplasy2026.6.0101

Received: 20 June 2026

Published: 21 June 2026

Corresponding author:
Hassan Assiri

dinho.1010@hotmail.com

Author Affiliation:
King Khalid University, College of Dentistry, Abha, Aseer Region, Saudi Arabia.

Clinical Performance of Artificial Intelligence-Assisted Cephalometric Analysis Compared with Expert or Manual Digital Tracing: A Systematic Review and Meta-Analysis

Assiri, HA; Alnami, HU; Alsaeed, AN; Almojthel, HN.

ADMINISTRATIVE INFORMATION

Support - The Deanship of Research and Graduate Studies at King Khalid University, KSA, for funding this work through the General Research Project (grant number RGP1/494/47).

Review Stage at time of this submission - Data extraction.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202660101

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 21 June 2026 and was last updated on 21 June 2026.

INTRODUCTION

Review question / Objective In patients indicated for lateral cephalometric analysis, how does the clinical performance of artificial intelligence (AI)-assisted cephalometric tracing compare with conventional manual or expert digital tracing, in terms of measurement agreement and accuracy of clinically relevant cephalometric parameters?

Rationale Cephalometric lateral analysis is an essential tool for the diagnosis and treatment planning process in both Orthodontics and Orthognathic Surgery. Traditionally, this has been done through a series of steps including manually identifying landmarks on radiographs (x-rays) followed by manually drawing lines that connect those identified landmarks; this is a very time consuming and variable process due to inter and intra operator differences. A proliferation of Artificial Intelligence (AI) assisted and completely automated cephalometric systems has developed over the past few years that claim to be as accurate as their manual counterparts but require significantly less time to perform. However, there are vast discrepancies in how each platform performs when compared to one another, depending on what parameter(s), etc. are being analyzed. There is also no single quantitative compilation of studies that resolve these discrepancies. This systematic review and meta-analysis will evaluate the accuracy of AI assisted cephalometric tracing when compared to manual or "expert" digitally traced x-rays for all clinically relevant skeletal, dental and soft tissue measurements.

Condition being studied The clinical performance of AI-assisted lateral cephalometric analysis compared with manual or expert digital tracing in patients indicated for cephalometry (a diagnostic-method/measurement-accuracy review, not a single disease).

METHODS

Participant or population Patients whose lateral cephalometric radiographs are indicated for orthodontic or craniofacial assessment. Studies using patient-derived lateral cephalograms are eligible regardless of malocclusion type, gender, or ethnicity.

Intervention AI-assisted or fully automated cephalometric tracing/landmark identification (any commercial or open-source AI platform), producing clinically relevant cephalometric measurements.

Comparator Conventional manual tracing or expert/human digital tracing of the same radiographs, serving as the reference standard.

Study designs to be included Cross-sectional, retrospective, and comparative method-agreement studies that directly compare AI-assisted tracing with manual/expert tracing at the patient level.

Eligibility criteria Inclusion criteria
• Studies using lateral cephalometric radiographs of patients.
• Direct comparison of AI-assisted tracing versus manual/expert digital tracing of the same images.
• Reporting of at least one clinically relevant cephalometric output (angle or linear measurement).
. Studies reported only in English language
Exclusion criteria
• Not a direct comparison of AI-assisted analysis versus manual/expert tracing.
• Education or student-training focus without patient-level diagnostic comparison.
• Single-image reproducibility studies without a patient-level sample.
• Landmark detection only (e.g., mean radial error) without clinically relevant cephalometric output, or AI prediction of values from non-radiographic inputs (reduced-field prediction).

Information sources Three electronic databases were searched: PubMed/MEDLINE, Scopus, and Web of Science Core Collection. No search restrictions.

Main outcome(s) The primary outcome is the agreement and accuracy of AI-assisted cephalometric tracing compared with manual or expert digital tracing for clinically relevant cephalometric measurements. Agreement is expressed as the intraclass correlation coefficient (ICC) between AI-assisted and reference (manual/expert) measurements, and accuracy as the mean difference between methods (in degrees or millimetres), with the proportion of parameters falling within accepted clinical thresholds (±2° or ±2 mm) where reported. Primary parameters of interest are the sagittal skeletal measurements SNA, SNB, and ANB.

Additional outcome(s) Secondary outcomes include analysis/tracing time, intra- and inter-examiner reliability, and per-parameter agreement for additional skeletal, dental, and soft-tissue measurements where reported.

Quality assessment / Risk of bias analysis Methodological quality and risk of bias were assessed using a modified QUADAS-2 tool adapted for cephalometric method-agreement studies, covering patient selection, index test (AI tracing), reference standard (manual/expert tracing), and flow and timing. Each study was independently appraised by two reviewers, with disagreements resolved by discussion.

Strategy of data synthesis A narrative synthesis will describe all included studies (design, AI platform, reference standard, parameters, and reported agreement/accuracy). Where studies report a compatible effect measure for the same cephalometric parameter, results will be pooled using a random-effects model (DerSimonian–Laird), chosen a priori in anticipation of clinical and methodological heterogeneity across AI platforms, reference standards, and populations. Pooled ICCs (and, where data permit, pooled mean differences) with 95% confidence intervals will be calculated for parameters reported by a sufficient number of studies. Where a single study contributes more than one AI arm against a shared comparator, one arm will be selected a priori (the fully automatic arm as primary) to preserve independence and avoid double-counting of the reference cohort. Heterogeneity will be quantified using the I^2 statistic and the χ^2 test. Analyses will be performed in Stata, with significance set at $P < 0.05$.

Subgroup analysis Where the number of contributing studies permits, subgroup analyses are planned by: (1) level of automation (fully automatic AI versus AI with manual landmark correction / semi-automatic); (2) AI platform/software; (3) parameter type (skeletal versus dental versus soft-tissue); and (4) reference standard used (e.g., manual tracing versus validated digital software). Subgroup differences will be interpreted cautiously given expected between-study heterogeneity.

Sensitivity analysis Sensitivity analyses are planned to assess the robustness of pooled

estimates by: (1) excluding studies at high risk of bias on the modified QUADAS-2 assessment; (2) excluding small pilot studies (e.g., very small samples); (3) excluding studies for which a confidence interval was not reported and the standard error had to be derived; and (4) repeating the primary pool using the alternative AI arm where multiple arms were available. Where ten or more studies contribute to a single analysis, small-study/reporting bias will be examined using funnel plots and Egger's test.

Country(ies) involved Saudi arabia.

Keywords Artificial intelligence; cephalometric analysis; lateral cephalogram; automated landmark detection; manual tracing.

Contributions of each author
Author 1 - Hassan Assiri.
Author 2 - Huda Alnami.
Author 3 - Anwar Alsaeed.
Author 4 - Hanaa almojthel.