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.

Diagnostic Concordance and Measurement Accuracy of WebCeph and Ceph Assistant AI Cephalometric Platforms: A Retrospective Comparative Study

Ali Rokia

aliandalus2017@gmail.com

Al- andalus University for Medical Sciences

Mouetaz Kheirallah

Al- andalus University for Medical Sciences

Research Article

Keywords: Artificial intelligence, Cephalometry, Orthodontics, WebCeph, Ceph Assistant

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11847913/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)-driven cephalometric platforms have rapidly gained popularity in orthodontic practice; however, minor angular discrepancies near diagnostic thresholds may alter clinical treatment decisions. This study aimed to compare the angular accuracy and categorical diagnostic agreement of WebCeph and Ceph Assistant against manual digital tracing.

Materials And Methods

Ninety pretreatment lateral cephalograms were traced manually in MicroDicom viewer (reference standard) and processed automatically on WebCeph and Ceph Assistant without operator correction. Ten angular variables were measured: SNA, SNB, ANB, SN-GoGn, FMA, Gonial angle, U1-SN, IMPA, Interincisal angle, and Nasolabial angle. Patients were grouped by sagittal skeletal class, vertical facial pattern, and lower incisor inclination. Statistical evaluation included intraclass correlation coefficients (ICC), repeated-measures ANOVA, and Cohen's kappa (κ).

Results

Sagittal angles (SNA, SNB, ANB) and SN-GoGn showed strong consistency on both platforms (ICC > 0.86). Lower incisor inclination (IMPA; ICC = 0.776–0.779) demonstrated good consistency, whereas the interincisal angle (ICC = 0.733–0.742) showed moderate reliability. For categorical diagnosis, Ceph Assistant agreed best with manual tracing on vertical growth patterns (84.4%, κ = 0.724 vs. 78.9%, κ = 0.618 for WebCeph). WebCeph performed slightly better on sagittal jaw classification (78.9%, κ = 0.621 vs. 76.7%, κ = 0.575) and IMPA categories (80.0%, κ = 0.625 vs. 78.9%, κ = 0.606). Most diagnostic mismatches occurred in patients whose angles lay close to classification cut-off points.

Conclusions

Both AI tools yield dependable angular values for routine skeletal screening, but orthodontists should verify lower incisor apices and mandibular landmarks manually in borderline cases before finalizing treatment plans.

Trial registration:

not applicable.

1. Introduction

Lateral cephalometric radiography (LCR) is an essential tool in orthodontic practice, routinely used for diagnosis and treatment planning. It provides valuable information regarding anteroposterior jaw relationships, vertical facial proportions, incisor inclination, and soft-tissue profile balance[1]. Although manual digital tracing of lateral cephalograms serves as the clinical reference standard, this procedure is time-consuming and prone to both inter-and intra-examiner variability in landmark identification[2]. Recently, advancements in convolutional neural networks (CNNs) have led to the development of cloud-based cephalometric platforms capable of identifying these landmarks within seconds[3, 4]. Software such as WebCeph and Ceph Assistant can automatically locate cranial, dental, and soft-tissue landmarks from uploaded radiographs without user intervention[5, 6]. Earlier validation studies evaluating AI programs[7, 8] relied on linear or angular differences and reported mean errors ranging from 1° to 2°[9], which are often assumed to be clinically negligible. However, even a minor shift of 1° can alter the diagnosis of critical cephalometric parameters, particularly near borderline values such as the ANB angle shifting from 3.8° to 4.5°, or the IMPA crossing the 95° threshold. Such variations can change a patient's skeletal classification, alter the assessment of incisor inclination, and ultimately impact clinical decisions regarding tooth extraction or orthognathic surgery. To the best of the authors' knowledge, no previous study has directly compared the performance of WebCeph and Ceph Assistant in terms of diagnostic categorical agreement against manual tracing, which serves as the primary motivation for this study.

The aim of this study was to evaluate the angular measurement reliability and categorical diagnostic concordance of WebCeph and Ceph Assistant compared with manual digital tracing across ten standard cephalometric parameters.

2. Materials and Methods

2.1 Ethical Approval and Study Design

The study protocol was reviewed and formally approved by the Institutional Review Board (IRB) of the Faculty of Dentistry at Al-andalus University for Medical Sciences. (Approval No. 12/20, date:3 October, 2026) Clinical trial number not applicable. Given the retrospective design of this investigation, which utilized archived, fully de-identified radiographic records, the requirement for individual patient informed consent was waived by the committee. All procedures were conducted strictly in accordance with the ethical principles outlined in the Declaration of Helsinki. Patient confidentiality was rigorously protected throughout the study by replacing all personal identifiers with randomized alphanumeric codes.

Ninety pretreatment digital lateral cephalograms were selected from the clinical archives of the Department of Orthodontics at the host institution. Sample size calculation was performed in G*Power (*version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf, Germany*) using an intraclass correlation coefficient (ICC = 0.879) and effect size derived from Mahto et al.[7]. With an alpha of 0.05, 90% power, and an effect size (f) of 0.15 across three measurement methods, at least 82 radiographs were needed.

We included 90 radiographs to ensure sufficient representation across skeletal classes (Class I, II, and III) and exceed prior sample sizes (n = 30)[7, 10].

2.2 Inclusion and Exclusion Criteria

Radiographs were included if they met four conditions: (1) high-resolution digital quality with clear hard- and soft-tissue contours, (2) teeth in maximum intercuspation, (3) lips relaxed at rest, and (4) natural head posture [6]. Exclusion criteria were: (1) craniofacial syndromes or cleft lip and palate, (2) marked mandibular asymmetry producing double lower borders, (3) impacted teeth overlapping points A or B or incisor root apices, and (4) fixed orthodontic appliances or motion blur.

2.3 Cephalometric Measurements and Diagnostic Cut-offs

All 90 radiographs were stripped of patient identifiers and coded (Case-01 through Case-90) prior to analysis. Ten angular measurements were recorded across four categories:

- • **Sagittal Skeletal**: SNA, SNB, and calculated ANB (SNA - SNB).
- • **Vertical Skeletal**: SN-GoGn, FMA (FH-MP), and Gonial angle (Ar-Go-Me).
- • **Dental Inclinations**: Upper incisor to Sella-Nasion plane (U1-SN), Lower incisor to Mandibular Plane angle (IMPA), and Interincisal angle (U1-L1).
- • **Soft-Tissue Profile**: Nasolabial angle (Columella-Subnasale-Labrale superius).

Patients were then grouped into clinical diagnostic categories using established cut-offs:

- • **Sagittal Skeletal Classification (ANB)**: Class I ($0° \leq ANB \leq 4°$), Class II ($ANB > 4°$), and Class III ($ANB < 0°$).
- • **Vertical Growth Pattern (SN-GoGn)**: Normodivergent ($28° \leq$ SN-GoGn $\leq 36°$), Hypodivergent (SN-GoGn $< 28°$), and Hyperdivergent (SN-GoGn $> 36°$).
- • **Lower Incisor Inclination (IMPA)**: Normal ($87° \leq IMPA \leq 95°$), Retroclined ($IMPA < 87°$), and Proclined ($IMPA > 95°$).

2.4 Gold Standard vs. AI Platforms and Blinding Protocol

Reference measurements were traced manually on raw DICOM images exported to the MicroDicom viewer (*MicroDicom Ltd., Sofia, Bulgaria*) by a calibrated orthodontist with over 5 years of clinical experience. All manual tracings were performed prior to AI processing, ensuring that the examiner remained blinded to the outputs of WebCeph and Ceph Assistant. Intra-examiner reliability was assessed by randomly re-tracing 20 radiographs two weeks later. Subsequently, the anonymized radiographs were uploaded to WebCeph (*AssembleCircle Corp., Korea*) and Ceph Assistant (*Ceph Assistant Ltd., Hungary*), and the automated angular outputs were recorded without manual landmark repositioning.

2.5 Statistical Analysis

Statistical testing was conducted in SPSS (*version 26.0; IBM Corp., Armonk, NY, USA*). Data distribution was checked with the Shapiro-Wilk test. Repeated-measures ANOVA compared mean angles across the three methods. Intraclass correlation coefficients (ICC) and Bland-Altman 95% limits of agreement (LoA, calculated as mean difference ± 1.96 SD) evaluated measurement consistency against a ± 2° clinical acceptability threshold[9], while Cohen's kappa (κ) measured categorical diagnostic agreement between each AI program and manual tracing.

3. Results

Intra-examiner reliability for MicroDicom tracing was high across all ten parameters (ICC range: 0.95–0.99). The sample consisted of 90 patients (55 females [61.1%] and 35 males [38.9%]) aged 13.1 to 31.8 years (mean: 22.06 ± 5.74 years). Manual reference classifications distributed the sample as follows: 54.4% Class I (n = 49), 38.9% Class II (n = 35), and 6.7% Class III (n = 6) sagittally; 61.1% Normodivergent (n = 55), 25.6% Hyperdivergent (n = 23), and 13.3% Hypodivergent (n = 12) vertically; and 60.0% Normal (n = 54), 23.3% Proclined (n = 21), and 16.7% Retroclined (n = 15) for IMPA.

Table 1 lists the mean ± standard deviation, ANOVA P-values, and ICC values for each platform relative to the gold standard. Mean differences across methods reached statistical significance ($P < 0.05$) for every parameter except ANB ($P = 0.638$). Sagittal variables (SNA, SNB, ANB) and SN-GoGn demonstrated strong agreement (ICC > 0.86), while Interincisal angles showed moderate reliability (ICC = 0.73–0.74) and IMPA demonstrated good reliability (ICC = 0.78).

Table 1

Descriptive statistics, repeated-measures ANOVA, and intraclass correlation coefficients (ICC) with 95% confidence intervals across all ten cephalometric parameters (N = 90).

Parameter	Gold Standard (Mean ± SD)	WebCeph (Mean ± SD)	Ceph Assistant (Mean ± SD)	*P*-value (ANOVA)	ICC (95% CI) WebCeph	ICC (95% CI) Ceph Assistant
SNA (°)	81.90 ± 3.18	82.23 ± 3.13	82.30 ± 3.48	< 0.001*	0.968 (0.938–0.982)	0.952 (0.914–0.971)
SNB (°)	78.63 ± 4.00	78.86 ± 3.96	78.94 ± 4.05	0.001*	0.980 (0.968–0.987)	0.977 (0.961–0.986)
ANB (°)	3.27 ± 2.46	3.36 ± 2.57	3.38 ± 2.52	0.638	0.919 (0.880–0.946)	0.865 (0.802–0.909)
SN-GoGn (°)	33.23 ± 5.00	32.39 ± 5.23	33.00 ± 5.17	< 0.001*	0.917 (0.851–0.951)	0.937 (0.905–0.958)
FMA (°)	26.46 ± 5.09	27.87 ± 5.80	28.13 ± 5.76	< 0.001*	0.880 (0.721–0.938)	0.804 (0.623–0.889)
Gonial Angle (°)	124.73 ± 5.69	126.25 ± 6.19	123.19 ± 6.52	< 0.001*	0.885 (0.730–0.941)	0.839 (0.708–0.906)
U1-SN (°)	105.33 ± 5.93	104.04 ± 6.42	104.47 ± 6.32	< 0.001*	0.916 (0.840–0.952)	0.925 (0.878–0.953)
IMPA (°)	92.20 ± 5.65	90.30 ± 6.23	91.19 ± 6.22	< 0.001*	0.779 (0.619–0.867)	0.776 (0.677–0.847)
Interincisal (°)	130.11 ± 5.79	132.18 ± 6.08	131.93 ± 6.13	< 0.001*	0.742 (0.556–0.845)	0.733 (0.594–0.825)
Nasolabial (°)	101.67 ± 9.21	98.78 ± 9.95	99.64 ± 9.75	< 0.001*	0.851 (0.692–0.919)	0.869 (0.785–0.918)

*Statistically significant difference across methods at P < 0.05. ICC values were calculated using a two-way mixed-effects model with absolute agreement (single measurement). Abbreviations: SD, standard deviation; ANOVA, analysis of variance; ICC, intraclass correlation coefficient; CI, confidence interval

Table 2 shows the categorical agreement percentages and Cohen's kappa coefficients for sagittal classes, vertical patterns, and lower incisor inclination. Overall agreement ranged from 76.7% to 84.4%. Ceph Assistant recorded the highest agreement in vertical facial pattern classification (84.4%, κ = 0.724),

while WebCeph had slightly higher agreement in sagittal skeletal class (78.9%, κ = 0.621) and lower incisor inclination (80.0%, κ = 0.625).(Fig. 1)

Table 2
Categorical Diagnostic Concordance Rates and Cohen's Kappa (κ) Values Compared with the Gold Standard (N = 90).

Diagnostic Dimension	AI Platform	Agreement (%)	Misclassified (n/90)	Kappa (κ)	Agreement Level
Sagittal Class (Class I, II, III)	WebCeph	78.9%	19/90	0.621	Substantial
Sagittal Class (Class I, II, III)	Ceph Assistant	76.7%	21/90	0.575	Moderate
Vertical Pattern (Hypo/Normo/Hyper)	WebCeph	78.9%	19/90	0.618	Substantial
Vertical Pattern (Hypo/Normo/Hyper)	Ceph Assistant	84.4%	14/90	0.724	Substantial
Lower Incisor (IMPA) (Retro/Norm/Pro)	WebCeph	80.0%	18/90	0.625	Substantial
Lower Incisor (IMPA) (Retro/Norm/Pro)	Ceph Assistant	78.9%	19/90	0.606	Substantial

AI, artificial intelligence; Hypo, hypodivergent; Normo, normodivergent; Hyper, hyperdivergent; Retro, retroclined; Norm, normal; Pro, proclined. Agreement levels follow the Landis and Koch scale (0.41–0.60, moderate; 0.61–0.80, substantial).

Bland-Altman agreement analysis (calculated as AI minus manual reference) revealed the narrowest 95% limits of agreement (LoA) for sagittal skeletal variables, particularly on WebCeph for SNA (-1.09° to + 1.76°), SNB (-1.29° to + 1.74°), and ANB (-1.90° to + 2.07°), compared with Ceph Assistant (SNA: -1.50° to + 2.29°; SNB: -1.27° to + 1.88°; ANB: -2.43° to + 2.66°). For vertical skeletal angles, SN-GoGn showed tighter LoA on both WebCeph (-4.63° to + 2.94°) and Ceph Assistant (-3.77° to + 3.31°) than FMA (WebCeph: -3.15° to + 5.98°; Ceph Assistant: -4.35° to + 7.69°) and the gonial angle (WebCeph: -3.36° to + 6.39°; Ceph Assistant:-7.79° to + 4.71°). Conversely, dental inclinations and soft-tissue measurements exhibited substantially wider individual LoA on both platforms, including U1-SN (WebCeph: -5.53° to + 2.95°; Ceph Assistant: -5.12° to + 3.39°), IMPA (WebCeph: -8.77° to + 4.96°; Ceph Assistant:-8.21° to + 6.19°), interincisal angle (WebCeph: -4.86° to + 9.00°; Ceph Assistant: -5.60° to + 9.24°), and nasolabial angle (WebCeph: -11.09° to + 5.32°; Ceph Assistant: -10.52° to + 6.46°).

4. Discussion

Rather than focusing solely on raw degree differences, this study evaluated how automated measurement errors on WebCeph and Ceph Assistant translate into categorical diagnostic shifts. While

most artificial intelligence-based cephalometric studies have focused on evaluating raw linear errors in millimeters[2, 4], only a few have investigated how automated landmark detection impacts the categorical diagnosis of different malocclusion types[5, 8]. In the present study, cephalometric landmarks defining sagittal jaw relationships such as Sella, Nasion, Point A, and Point B are situated along the midsagittal plane and exhibit sharp radiographic contrast. Consequently, both cephalometric platforms, WebCeph and Ceph Assistant, demonstrated high reliability for sagittal parameters, consistent with the findings of Mahto et al.[7] and recent multi-platform evaluations[8, 11]. Sagittal misclassifications were confined to borderline cases with ANB angles ranging from 3.5° to 4.5°. Conversely, vertical skeletal angles (FMA and Gonial angle), dental inclinations (IMPA and Interincisal angle), and the nasolabial angle exhibited greater measurement dispersion, with mean discrepancies exceeding the ± 2.0° clinical acceptability threshold for the nasolabial angle (-2.89° in WebCeph; -2.03° in Ceph Assistant) and the interincisal angle (+ 2.07° in WebCeph). Additionally, the gonial angle displayed opposite directional biases between the two platforms (+ 1.52° for WebCeph and − 1.54° for Ceph Assistant), likely due to algorithmic differences in identifying superimposed bilateral mandibular borders (Fig. 2).

Moreover, the wide 95% Bland-Altman limits of agreement observed for dental inclinations and the nasolabial angle (exceeding ± 8° to ± 11° in individual cases) underscore that low mean group biases can mask substantial case-by-case discrepancies. Overlapping incisor roots and surrounding alveolar bone often obscure dental apices during automated detection, whereas bilateral structures such as Gonion and Porion are affected by anatomical superimposition[5, 6]. Bor et al.[5]and recent reports on WebCeph[11, 12]confirmed clinically relevant errors in IMPA and the nasolabial angle due to difficulties in pinpointing lower incisor apices and the soft-tissue subnasale landmark[12]. Similarly, recent evaluations of Ceph Assistant noted high processing speed alongside notable variability in localizing mandibular plane and soft-tissue landmarks[6].

In contrast, Katyal and Balakrishnan[13] found no significant differences in angular measurements between WebCeph and manual tracing, which may be attributed to their smaller sample size (n = 25) that carried lower statistical power than our cohort of 90 patients. Furthermore, Yassir et al.[14] reported poor agreement for automated IMPA measurements, whereas our results demonstrated moderate-to-good consistency across dental inclinations (IMPA and interincisal angle; ICC = 0.73–0.78). This improvement is likely due to ongoing advancements in deep-learning landmark detection models implemented across successive software updates. Clinically, while fully automated AI tracing saves considerable time during initial patient screening[3], orthodontists should manually verify lower incisor apices, Gonion, and subnasale landmarks before finalizing borderline extraction or surgical treatment plans[6, 10, 11]

Several limitations of this study warrant consideration. First, we evaluated angular measurements and categorical diagnoses rather than linear millimeter coordinates (X-Y errors) of individual landmarks; theoretically, opposite positional errors at two landmarks could offset one another and yield an apparently accurate angle despite inaccurate point localization. Second, the use of conventional two-

dimensional lateral cephalograms introduced bilateral anatomical superimposition, particularly near the gonial angle and the inferior mandibular border, compared with three-dimensional cone-beam computed tomography (CBCT). Finally, the data were derived from a single clinical center, limiting sample variability, and the results may not be generalizable to more complex craniofacial cases due to the exclusion of severe facial asymmetries.

5. Conclusion

Both WebCeph and Ceph Assistant offered reliable angular measurements for the sagittal skeletal parameters (SNA, SNB, and ANB) and the SN-GoGn angle, whereas measurements for lower incisor inclination (IMPA) and the interincisal angle exhibited greater variability. Categorical diagnostic agreement with manual tracing in MicroDicom was substantial in most domains, with Ceph Assistant performing best in vertical growth pattern classification and WebCeph demonstrating slightly higher concordance in sagittal class assignment and IMPA status. While these platforms are useful for routine orthodontic screening, clinicians should manually verify incisor apices and mandibular landmarks in borderline cases to avoid diagnostic misclassification

Abbreviations

AI
Artificial intelligence
ANB
A point-asion-B point angle
ANOVA
Analysis of variance
CBCT
Cone-beam computed tomography
CI
Confidence interval
CNNs
Convolutional neural networks
DICOM
Digital Imaging and Communications in Medicine
FMA
Frankfort mandibular plane angle
ICC
Intraclass correlation coefficient
IMPA
Incisor mandibular plane angle
IRB
Institutional Review Board

LCR
Lateral cephalometric radiography
LoA
Limits of agreement
SD
Standard deviation
SNA
Sella-Nasion-A point angle
SNB
Sella-Nasion-B point angle
SN-GoGn
Sella-Nasion to Gonion-Gnathion angle
U1-L1
Interincisal angle (Upper incisor to lower incisor angle)
U1-SN
Upper incisor to Sella-Nasion plane angle

Declarations

Ethical approval

The study protocol was reviewed and officially approved by the Institutional Review Board (IRB) of the Faculty of Dentistry at Al-Andalus University for Medical Sciences (Approval No. 12/20 , date:3 October, 2026).

Declaration of patient consent

Due to the retrospective nature of this study utilizing archived, completely de-identified radiographic records, the requirement for individual written informed consent was officially waived by the Institutional Review Board (IRB). Patient confidentiality was strictly maintained throughout the investigation in accordance with the Declaration of Helsinki.

Consent for publication

not applicable

Availability of data and materials

The data used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare no competing interests.

Funding

The authors declare no financial support.

Authors' contributions

Ali Rokia performed methodology, data statistics, and conceptualization, and wrote the original draft. Mouetaz Kheirallah performed project administration, supervision, validation, visualization, tables and figures, interpretation of statistical data, and review of it.

Acknowledgments

The authors extend their sincere thanks and appreciation to the clinical staff at Al-Andalus University for Medical Sciences for their technical support of the project.

References

1. Kunz F, Stellzig-Eisenhauer A, Zeman F, Boldt J. Artificial intelligence in orthodontics: Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network. J Orofac Orthop. 2020;81(1):52–68. 10.1007/s00056-019-00203-8. English. doi:.
2. Hwang HW, Park JH, Moon JH, Yu Y, Kim H, Her SB, Srinivasan G, Aljanabi MNA, Donatelli RE, Lee SJ. Automated identification of cephalometric landmarks: Part 2-Might it be better than human? Angle Orthod. 2020;90(1):69–76. Epub 2019 Jul 22. PMID: 31335162; PMCID: PMC8087057.
3. Meriç P, Naoumova J. Web-based Fully Automated Cephalometric Analysis: Comparisons between App-aided, Computerized, and Manual Tracings. Turk J Orthod. 2020;33(3):142–9. PMID: 32974059; PMCID: PMC7491969.
4. Schwendicke F, Chaurasia A, Arsiwala L, Lee JH, Elhennawy K, Jost-Brinkmann PG, Demarco F, Krois J. Deep learning for cephalometric landmark detection: systematic review and meta-analysis. Clin Oral Investig. 2021;25(7):4299–309. 10.1007/s00784-021-03990-w. . Epub 2021 May 27. PMID: 34046742; PMCID: PMC8310492.
5. Bor S, Ciğerim SÇ, Kotan S. Comparison of AI-assisted cephalometric analysis and orthodontist-performed digital tracing analysis. Prog Orthod. 2024;25(1):41. 10.1186/s40510-024-00539-x. . PMID: 39428414; PMCID: PMC11491421.
6. Polizzi A, Lo Giudice A, Conforte C, Isola G, Leonardi R. Influence of head positioning errors on the accuracy of fully automated artificial intelligence-based cephalometric software. Angle Orthod. 2025;95(6):619–28. PMID: 41015410; PMCID: PMC12544317.
7. Mahto RK, Kafle D, Giri A, Luintel S, Karki A. Evaluation of fully automated cephalometric measurements obtained from web-based artificial intelligence driven platform. BMC Oral Health. 2022;22(1):132. 10.1186/s12903-022-02170-w. . PMID: 35440037; PMCID: PMC9020017.

8. Kazimierczak W, Gawin G, Janiszewska-Olszowska J, Dyszkiewicz-Konwińska M, Nowicki P, Kazimierczak N, Serafin Z, Orhan K. Comparison of Three Commercially Available, AI-Driven Cephalometric Analysis Tools in Orthodontics. J Clin Med. 2024;13(13):3733. 10.3390/jcm13133733. . PMID: 38999299; PMCID: PMC11242750.
9. Bulatova G, Kusnoto B, Grace V, Tsay TP, Avenetti DM, Sanchez FJC. Assessment of automatic cephalometric landmark identification using artificial intelligence. Orthod Craniofac Res. 2021;24(Suppl 2):37–42. 10.1111/ocr.12542. . Epub 2021 Nov 29. PMID: 34842346.
10. Waheed A, Abid M. Accuracy and reliability of the artificial intelligence-assisted WebCeph application for lateral cephalometric analysis in comparison with manual tracing. J Coll Physicians Surg Pak. 2025;35(5):612–7.
11. Alassfar A, El Merouani IA, Harti I, Zaoui F, Benyahia H. Agreement and reliability of AI-based cephalometric analysis (WebCeph) compared with digital manual tracing for skeletal and dentoalveolar measurements (A retrospective comparative study). J Bagh Coll Dent. 2026;38(2):20–31. 10.26477/jbcd.v38i2.4215.
12. Gürler G, Toroglu MS, Cam OY. Agreement of WebCeph-Based Automated and Expert-Adjusted Cephalometric Analyses with Manual and Dolphin Tracings. Diagnostics (Basel). 2026;16(12):1836. 10.3390/diagnostics16121836. . PMID: 42351494; PMCID: PMC13298020.
13. Katyal D, Balakrishnan N. Evaluation of the accuracy and reliability of WebCeph – An artificial intelligence-based online software. APOS Trends Orthod. 2022;12:271–6.
14. Yassir YA, Salman AR, Nabbat SA. The accuracy and reliability of WebCeph for cephalometric analysis. J Taibah Univ Med Sci. 2021;17(1):57–66. 10.1016/j.jtumed.2021.08.010. . PMID: 35140566; PMCID: PMC8801471.

Figures

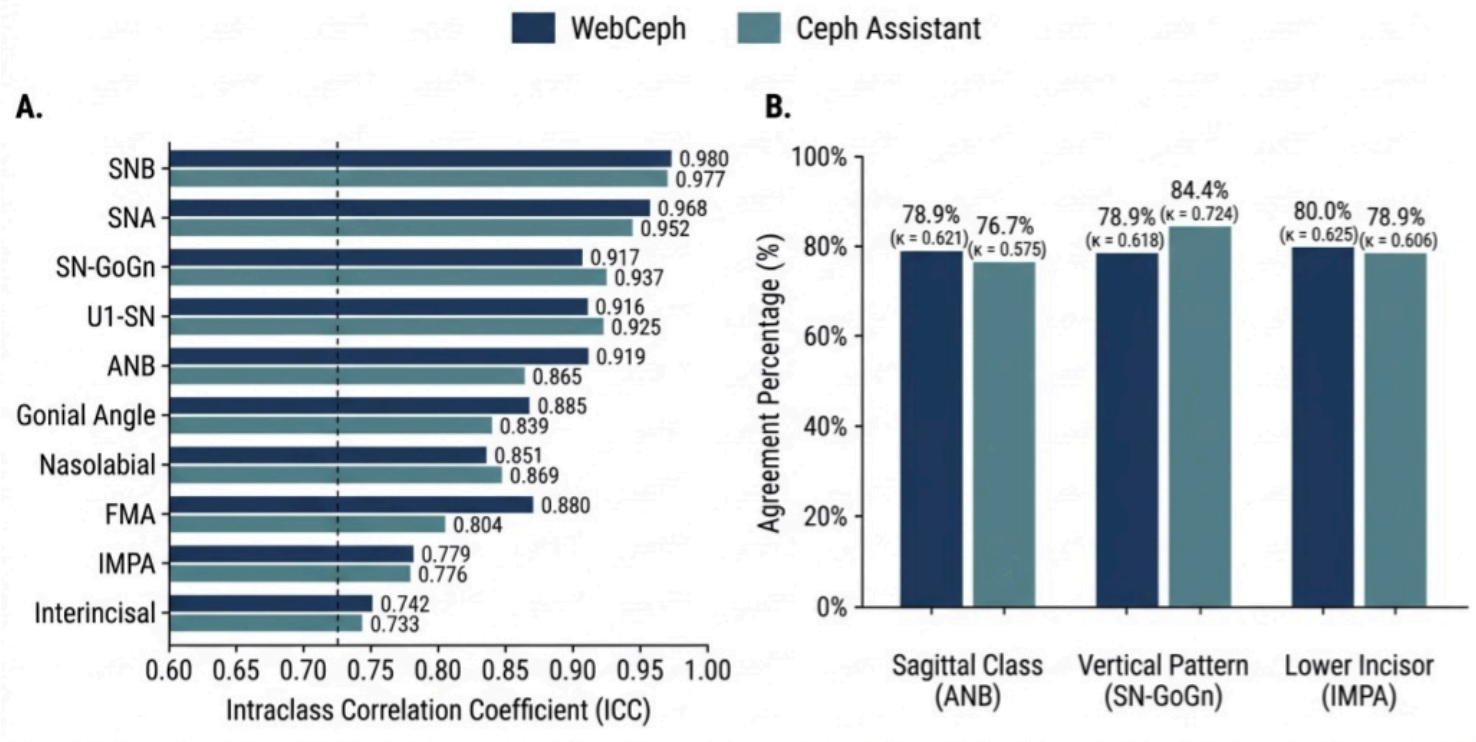


Figure 1

Comparison of measurement reliability and categorical diagnostic agreement between WebCeph and Ceph Assistant against manual digital tracing (N = 90). (A) Intraclass correlation coefficients (ICC) across ten angular cephalometric parameters. (B) Percentage of categorical diagnostic agreement and Cohen's kappa (κ) values for sagittal skeletal classification (ANB), vertical facial growth pattern (SN-GoGn), and lower incisor inclination (IMPA).

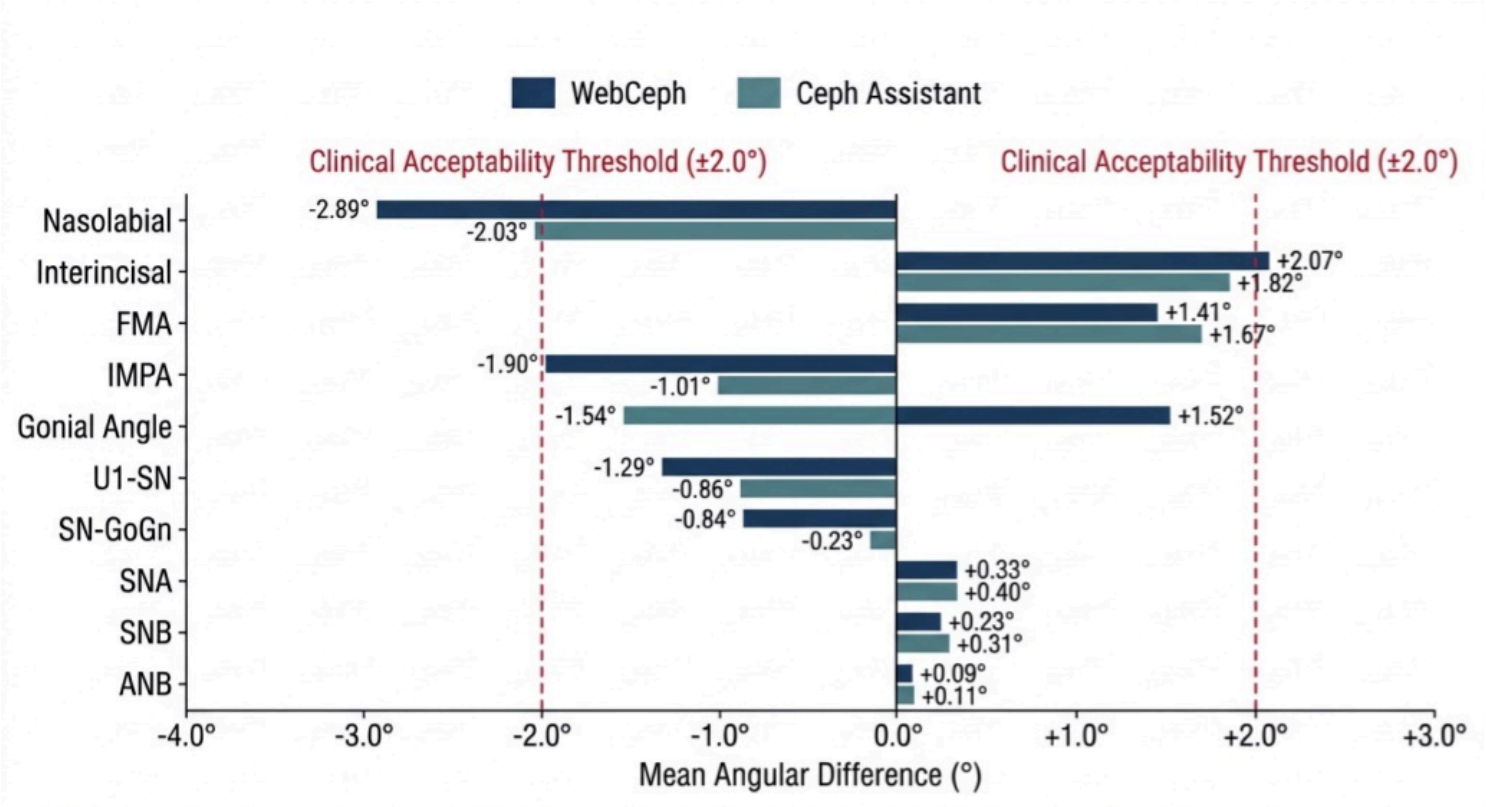


Figure 2

Mean angular discrepancies (in degrees) of WebCeph and Ceph Assistant relative to the manual digital tracing gold standard across ten cephalometric parameters (N = 90). Vertical dashed red lines indicate the ±2.0° clinical acceptability threshold.