

Accuracy of cephalometric landmark identification on artificial intelligence-based software: a comparative study

İpek **ŞAVKAN**¹

 <https://orcid.org/0000-0002-8144-1320>

Sara Nur **ÖZÇANKAYA**¹

 <https://orcid.org/0009-0002-6393-7215>

Öykü Naz **TURAN**² 

 <https://orcid.org/0009-0009-7668-9960>

Submitted: June 23, 2025 • Revised and accepted: July 15, 2025

 oykunazturan@ogr.iu.edu.tr

How to cite: Şavkan İ, Özçankaya SN, Turan ÖN. Accuracy of cephalometric landmark identification on artificial intelligence-based software: a comparative study. Dental Press J Orthod. 2026;31(2):e2625144.

Editor-in-Chief: Matheus Melo Pithon / Associate Editor: Matheus Melo Pithon

(1) İstanbul Kent University, Faculty of Dentistry, Department of Orthodontics (İstanbul, Türkiye).

(2) İstanbul University, Faculty of Dentistry, Department of Orthodontics (İstanbul, Türkiye).

ABSTRACT

Introduction: Cephalometry has long been a cornerstone of orthodontic diagnosis and treatment planning. While traditionally performed manually, cephalometric analysis has increasingly shifted toward semi-automated or fully automated computer-based software, due to advances in technology. **Objective:** This study aims to evaluate and compare the accuracy of anatomical landmarks identified manually by orthodontists using digital tools versus those identified automatically by various artificial intelligence (AI)-based cephalometric software. **Methods:** A total of 25 lateral cephalometric radiographs were selected for analysis. Manual landmark identification was independently performed by two orthodontists using the digital cephalometric software NemoCeph. These were compared to the automatic landmark identifications generated by two AI-based programs: CephNinja and WebCeph™. A total of 17 anatomical landmarks, 15 hard tissue and 2 soft tissue points, were evaluated. The 17 cephalometric landmarks selected in this study are among the most commonly used and reliably identifiable reference points in the literature. To facilitate direct visual comparison, landmark positions derived from all three methods were overlaid using semi-transparent image layers. One-way ANOVA and intraclass correlation coefficient (ICC) tests were conducted to assess inter-method differences. The threshold for statistical significance was set at $p < 0.05$ and $p < 0.01$. **Results:** Statistically significant differences were observed at the porion, basion, gonion, and the apexes of the upper and lower incisors ($p < 0.05$). Compared to the manual tracings in NemoCeph, AI-based software generally placed these landmarks in more distal and superior positions. Among the AI tools, CephNinja showed greater similarity to NemoCeph for the gonion and basion points, while WebCeph™ was closer at the incisor apexes. **Conclusion:** Although AI-based cephalometric software offers rapid and reproducible assessments, certain anatomical landmarks – particularly porion, basion, gonion, and the apexes of the upper and lower incisors – demonstrate statistically significant positional discrepancies, compared to manual tracings. Clinicians should be cautious when relying exclusively on automated analyses for diagnostic or treatment planning purposes, since AI struggles with certain landmarks, due to algorithmic limitations.

Keywords: Cephalometrics. Orthodontic imaging. Artificial intelligence. Diagnostic records. Digitalization.

RESUMO

Introdução: A cefalometria tem sido, por muito tempo, um pilar do diagnóstico e do planejamento do tratamento ortodôntico. Embora tradicionalmente realizada de forma manual, a análise cefalométrica tem migrado cada vez mais para *softwares* computadorizados semiautomatizados ou totalmente automatizados, devido aos avanços tecnológicos. **Objetivo:** Este estudo teve como objetivo avaliar e comparar a acurácia dos marcos anatômicos identificados manualmente por ortodontistas por meio de ferramentas digitais *versus* aqueles identificados automaticamente por diversos *softwares* cefalométricos baseados em inteligência artificial (IA). **Métodos:** Um total de 25 telerradiografias laterais foram selecionadas para análise. A identificação manual dos marcos anatômicos foi realizada de forma independente por dois ortodontistas, utilizando o *software* de cefalometria digital NemoCeph. Essas foram comparadas às identificações automáticas de marcos geradas por dois programas baseados em IA: CephNinja e WebCeph™. Foram avaliados 17 marcos anatômicos no total, sendo 15 em tecido duro e 2 em tecido mole. Os 17 marcos cefalométricos selecionados neste estudo estão entre os pontos de referência mais comumente utilizados e citados na literatura. Para facilitar a comparação visual direta, as posições dos marcos derivadas dos três métodos foram sobrepostas utilizando camadas de imagem semitransparentes. Os testes ANOVA de uma via e o coeficiente de correlação intraclassa (ICC) foram utilizados para avaliar as diferenças entre os métodos. O nível de significância estatística foi definido em $p < 0,05$ e $p < 0,01$. **Resultados:** Diferenças estatisticamente significativas foram observadas no pório, básico, gônio e nos ápices dos incisivos superiores e inferiores ($p < 0,05$). Em comparação aos traçados manuais no NemoCeph, os *softwares* baseados em IA geralmente posicionaram esses marcos em localizações mais distais e superiores. Entre as ferramentas de IA, o CephNinja apresentou maior similaridade com o NemoCeph para os pontos gônio e básico, enquanto o WebCeph™ foi mais próximo nos ápices dos incisivos. **Conclusões:** Embora os *softwares* cefalométricos baseados em IA ofereçam avaliações rápidas e reproduzíveis, certos marcos anatômicos — particularmente o pório, o básico, o gônio e os ápices dos incisivos superiores e inferiores — demonstram discrepâncias posicionais estatisticamente significativas, em comparação aos traçados manuais. Os clínicos devem ser cautelosos ao confiar exclusivamente em análises automatizadas para fins de diagnóstico ou planejamento de tratamento, uma vez que a IA apresenta dificuldades com certos marcos, devido a limitações algorítmicas.

Palavras-chave: Cefalometria. Imaginologia ortodôntica. Inteligência artificial. Registros diagnósticos. Digitalização.

INTRODUCTION

Although radiography was initially designed as a research laboratory tool, it has now become an indispensable diagnostic method in orthodontics, particularly for treatment planning and growth assessment.¹ Lateral cephalometric radiography, which captures a side-view image of the skull, is a vital imaging technique used to evaluate sagittal and vertical craniofacial relationships. Cephalometry has long been a central tool for assessing dental and skeletal development, post-treatment changes, and for supporting clinical and research studies.²

Traditional cephalometric analysis begins with the manual identification of anatomical landmarks on specialized tracing paper, followed by angular and linear measurements using a ruler and compass. Although this method is still considered the gold standard in many studies, it is time-consuming and associated with a higher risk of error in landmark identification.³

With the advent of digital technology, semi-automated and fully automated cephalometric analysis tools have increasingly replaced manual tracing methods.⁴ Software such as NemoCeph, although not artificial intelligence (AI)-based, allows for digital manual landmark identification and measurement, offering increased efficiency and consistency. These tools reduce manual effort while still requiring user input and validation, thus ensuring accuracy through a combination of automation and human oversight.

Today, AI has emerged as a powerful tool across many domains, including Orthodontics. AI is defined as the science and engineering of creating intelligent machines, particularly smart computer programs.⁵ In Orthodontics, one of the most prominent applications of AI is automated cephalometric analysis.

Several commercially available AI-based software programs — such as CephX[®] (ORCA Dental AI, Las Vegas, USA), AudaxCeph (Audax, Ljubljana, Slovenia), WebCeph[™] (Assemblecircle, Seoul, Korea), and CephNinja (Cyncronus LLC, United States) — offer automatic reference point detection.⁶ These AI systems are designed to automatically identify cephalometric landmarks, lines, and angles on radiographs, reducing human error and saving time.

Moreover, AI models can be adapted by clinicians to suit specific patient characteristics or clinical conditions, enabling the development of more individualized and precise treatment plans.⁷

Unlike previous studies that focused primarily on angular or linear measurements, the present study uniquely evaluated positional discrepancies using a coordinate-based approach across multiple commercial AI platforms, offering a more precise understanding of landmarking inconsistencies.⁸

This study aimed to evaluate and compare the accuracy of anatomical landmark identification performed manually by experienced orthodontists and automatically by different AI-based cephalometric software programs.

MATERIAL AND METHODS

Ethical approval was obtained from the Istanbul Kent University (Approval No: 2024-06)

This study was based on a comparison between manual tracings performed by two orthodontists and automatic tracings produced by various AI-based cephalometric software programs. Twenty-five lateral cephalometric radiographs were selected using a simple random sampling method from an anonymized archive of radiographs taken at Kent University's Department of Orthodontics. The radiographs were obtained over a six-month period, between June 2024 and December 2024. Inclusion criteria were: high image quality, clear visibility of anatomical landmarks, and no history of craniofacial anomalies or prior orthodontic treatment.

Exclusion criteria included radiographs with poor image quality, patients with a history of surgical intervention, pregnancy, radiation sensitivity, or craniofacial syndromes. No restrictions were placed regarding age or gender. These criteria were designed to ensure sample homogeneity and to improve data reliability.

The required sample size was calculated using G*Power software. *A priori* power analysis indicated that a minimum of 25 radiographs would be sufficient to detect large effect sizes (effect size = 0.95) with a power of 95% at an alpha level of 0.05. This calculation assumed moderate variability in landmark positioning across methods.

Two orthodontists, each with a minimum of 5 years of clinical experience, independently traced the same radiographs using the NemoCeph software (Nemotec, Madrid, Spain) (Fig. 1). Each tracing was repeated after a 15-day interval, and the intra- and inter-observer reliability was assessed using the Intraclass Correlation Coefficient (ICC) (Table 1).

The same films were then analyzed using two AI-based cephalometric programs: WebCeph™ (Assemblecircle, Seoul, Korea) (Fig. 2) and CephNinja (Cyncronus LLC, USA) (Fig. 3). Seventeen cephalometric landmarks selected in this study are among the most commonly used and reliably identifiable reference points in the literature (Table 2). In the AI software, anatomical landmarks were automatically placed and not altered by the researchers. However, in the manual tracings, orthodontists were allowed to adjust landmark positions to reflect the most anatomically accurate placement.

Table 1: Intraclass Correlation Coefficient (ICC).

	ICC	p
Sella	0.846	0.001**
Nasion	0.823	0.001**
Orbita	0.888	0.001**
Basion	0.947	0.001**
Porion	0.915	0.001**
Gonion	0.904	0.001**
Menton	0.875	0.001**
Gnathion	0.867	0.001**
Pogonion	0.840	0.001**
A Point	0.861	0.001**
B Point	0.836	0.001**
Tip of upper incisor	0.874	0.001**
Apex of upper incisor	0.841	0.001**
Tip of lower incisor	0.898	0.001**
Apex of lower incisor	0.941	0.001**
Subnasale	0.908	0.001**
Nasal tip	0.887	0.001**

The correlation coefficient (ICC) values for the specified class in all measurements were found to be close to 1 ($p < 0.01$). It was observed that the same or similar results were obtained in both measurements.

** statistically significant.

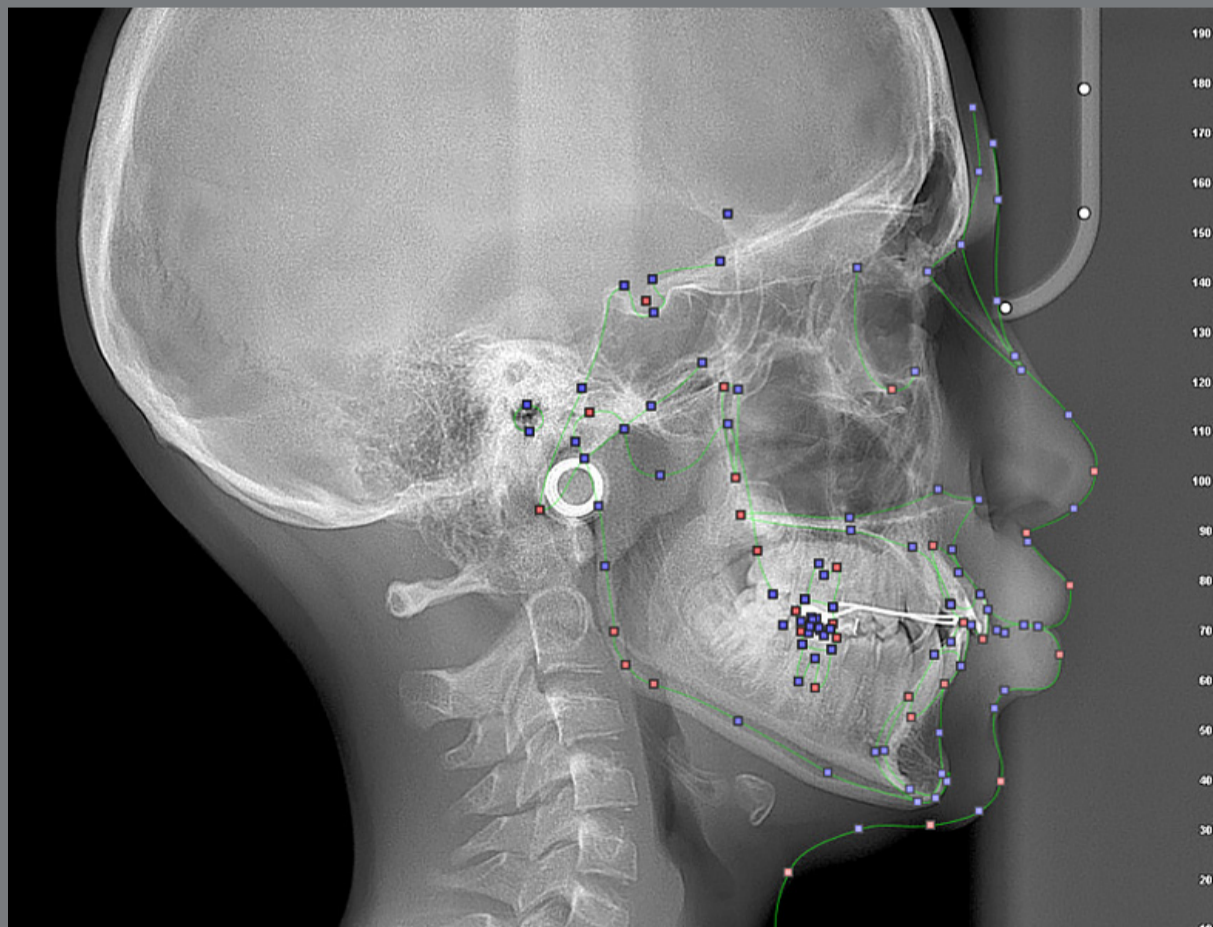


Figure 1: Anatomical reference points traced using the Nemoceph software.

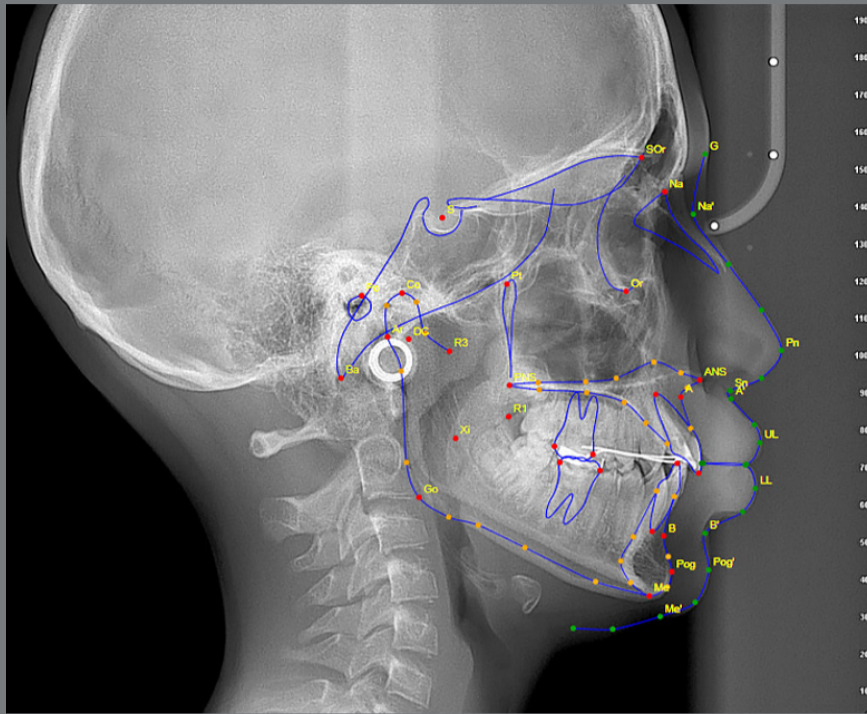


Figure 2: Anatomical reference points traced using the Webceph software.

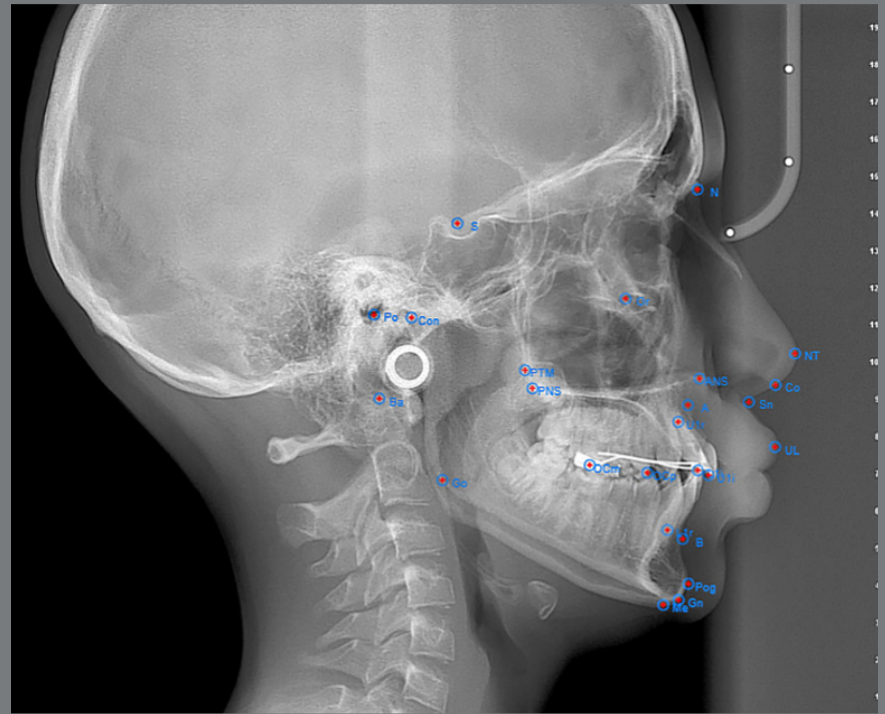


Figure 3: Anatomical reference points traced using the CephNinja software.

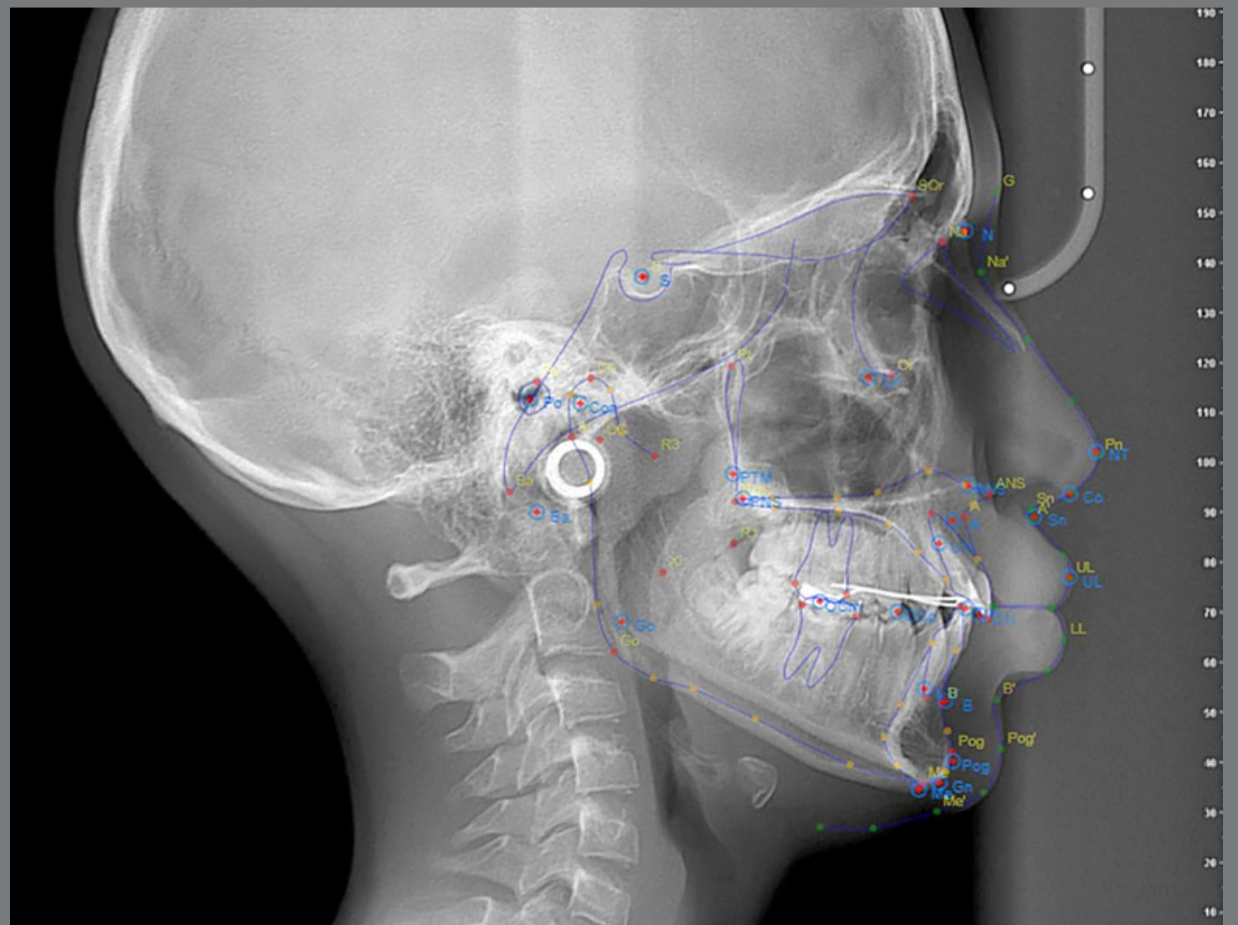
Table 2: Anatomical landmarks.

Landmark	Description
Sella	Center of the sella turcica
Nasion	Intersection of the frontal and nasal bones
Porion	Upper margin of the external auditory meatus
Basion	Most inferior-posterior point of the foramen magnum
Gonion	Most posterior-inferior point on the angle of the mandible
Gnathion	Midpoint between the most anterior and inferior points of the chin
Menton	Most inferior point on the mandibular symphysis
Pogonion	Most anterior point on the chin
A Point	Deepest point on the curvature of the anterior maxilla
B Point	Deepest point on the curvature of the anterior mandible
Orbita	Lowest point on the infraorbital rim
Tip of upper incisor	Incisal edge of the most prominent upper central incisor
Apex of upper incisor	Root apex of the most prominent upper central incisor
Tip of lower incisor	Incisal edge of the most prominent lower central incisor
Apex of lower incisor	Root apex of the most prominent lower central incisor
Subnasale	Junction of the columella base and the upper lip
Nasal tip	Most anterior point of the nose

The traced radiographs were overlaid across the three software platforms to visually compare the positioning of each landmark. The images were made semi-transparent, and alignment was verified using the reference rulers embedded in each software (Fig. 4).

The X and Y coordinates of each anatomical landmark were recorded separately for each software program. The distances were measured in millimeters. The positional differences (error margins) between the coordinates of the same landmark across different programs were calculated and compared.

Figure 4: Overlay of analyses performed using the Nemoceph, Cephinja, and Webceph software programs.



STATISTICAL ANALYSIS

All statistical analyses were conducted using SPSS version 24.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics including mean, standard deviation, median, frequency, ratio, minimum, and maximum were computed.

To compare the coordinates of landmarks across the three software programs, one-way ANOVA was applied, followed by Bonferroni correction for *post-hoc* pairwise comparisons.

To assess intra- and inter-observer reliability, Intraclass Correlation Coefficient (ICC) values were calculated.

Statistical significance was evaluated at $p < 0.05$ and $p < 0.01$ levels.

RESULTS

According to the coordinate system used in this study, X-axis values increased from posterior to anterior, while Y-axis values increased from cranial to caudal (Fig. 5).

Statistically significant differences were found in the anatomical positions of the following landmarks: porion, basion, gonion, apex of the upper incisor, and apex of the lower incisor ($p < 0.05$) (Table 3).

Table 3: Comparison of anatomical landmark coordinates.

		Coordinate values of anatomical points			^a p
		NemoCeph Mean ± SD	CephNinja Mean ± SD	WebCeph Mean ± SD	
Sella	X-axis	1162.36±332.07	1162.16±332.28	1161.96±332.04	0.920
	Y-axis	410.8±83.53	410.32±82.08	410.88±83.13	0.932
Nasion	X-axis	1353.96±384.84	1354.84±384.27	1349.48±382.94	0.945
	Y-axis	381.16±82.33	379.84±79.61	381.28±79.7	0.956
Porion	X-axis	1101±312.46	1012.04±313.69	1050.88±313.37	0.044*
	Y-axis	465.32±95.43	435.6±99.51	426.72±94.97	0.032*
Basion	X-axis	1098.84±314.71	1058.36±314.56	1055.44±309.89	0.034*
	Y-axis	525.8±109.78	507.72±113.14	510.08±112.05	0.037*
Gonion	X-axis	1155.56±323.09	1131.44±326.67	1125.32±323.1	0.035*
	Y-axis	623.32±134.97	611.6±137.95	608.44±137.2	0.039*
Gnathion	X-axis	1335.48±379.35	1335.44±379.71	1335.4±378.27	0.975
	Y-axis	702.96±155.94	702.76±156.27	702.08±157.43	0.912
Menton	X-axis	1324.88±375.08	1323.44±375.63	1324.92±374.55	0.981
	Y-axis	705.52±158.07	705.24±157.98	704.88±156.93	0.989
Pogonion	X-axis	1342.64±381.03	1343.24±382.93	1341.72±380.95	0.991
	Y-axis	689.68±153.26	690.44±152.93	685.44±151.42	0.992
A Point	X-axis	1382.92±408.72	1384±407.2	1387.24±408.72	0.994
	Y-axis	552.16±124.16	551.72±122.53	553.16±122.57	0.931
B Point	X-axis	1334.64±383.37	1336.24±383.83	1334.8±382.84	0.978
	Y-axis	655.64±143.5	654±144.11	652.24±143.47	0.922
Orbita	X-axis	1310.6±375.47	1298.56±371.65	1313±373.09	0.989
	Y-axis	460.88±95.15	466.76±97.91	463.24±96.7	0.977
Tip of upper incisor	X-axis	1362.96±391.25	1361.96±391.32	1362.56±391	0.978
	Y-axis	607.24±132.72	604.4±132.26	606.56±132.85	0.992
Apex of upper incisor	X-axis	1336.52±382.02	1312.68±384.51	1327.12±382.61	0.041*
	Y-axis	542.96±115.21	527.96±121.61	536.24±116.72	0.039*
Tip of lower incisor	X-axis	1350.44±389.04	1350.88±389.53	1349.76±388.17	0.932
	Y-axis	599.4±130.63	602.12±130.74	599.64±129.83	0.997
Apex of lower incisor	X-axis	1319.24±377.26	1300.64±378.88	1305.48±379.92	0.039*
	Y-axis	657.24±143.62	639.76±141.16	647.36±142.8	0.045*
Subnasale	X-axis	1399±395.35	1396.44±396.64	1399.84±396.17	0.939
	Y-axis	542±116.07	550.84±116.08	544.12±119.77	0.962
Nasal tip	X-axis	1434.36±396.88	1433.84±398.43	1435.24±398.55	0.957
	Y-axis	512.12±107.32	510.44±106.87	510.2±106.04	0.998

^a One-way ANOVA test.

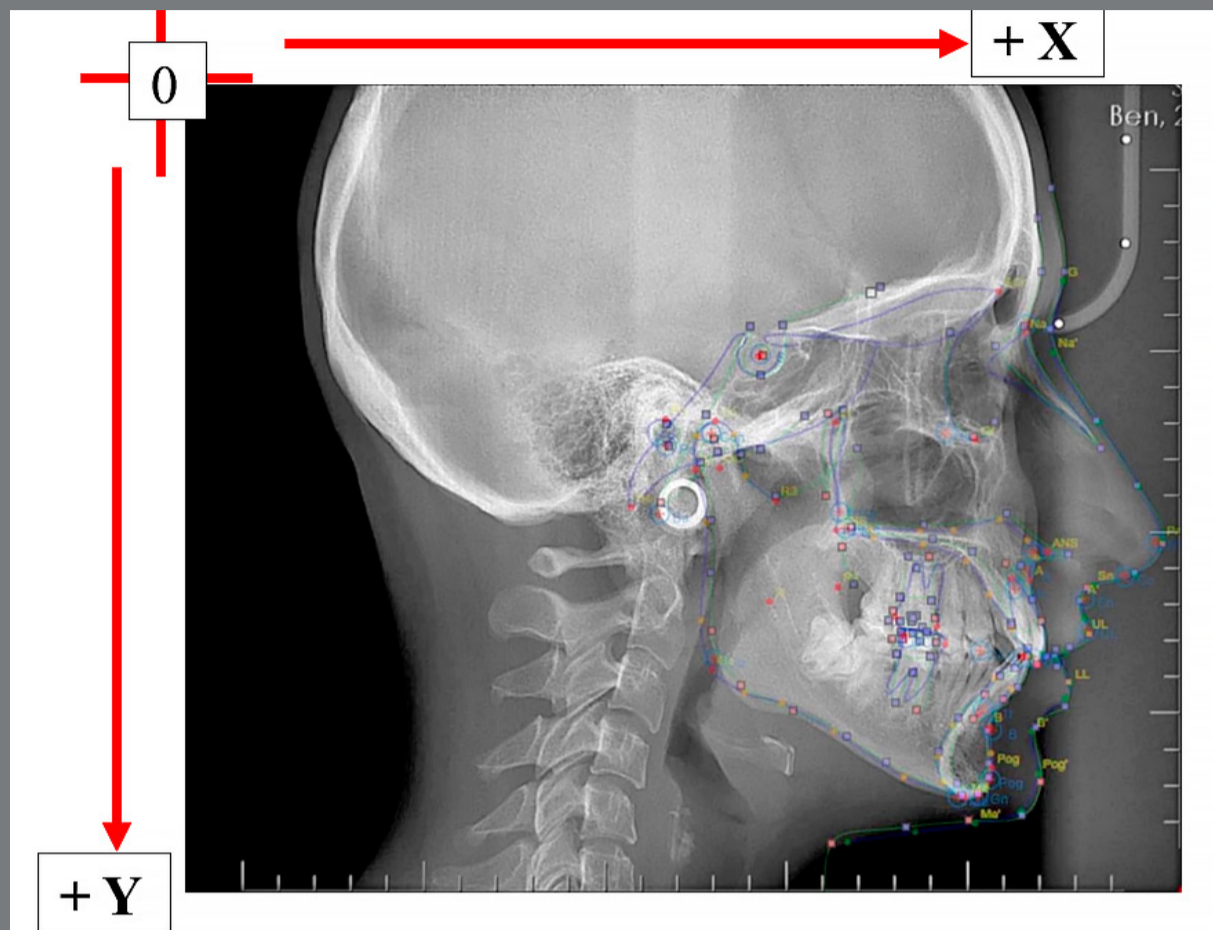


Figure 5: Superimposition of lateral cephalometric films obtained with Ceph-ninja, WebCeph, and Nemo programs, and the use of a coordinate system.

For the porion point, the mean coordinates on both X and Y axes differed significantly across the three software programs ($p = 0.044$ for the X-axis; $p = 0.032$ for the Y-axis). According to pairwise comparisons, AI-based software localized this point more distally and superiorly than the manual software. The coordinates determined by CephNinja were closer to those of NemoCeph in both the X and Y directions.

For the basion point, statistically significant differences in both X and Y coordinates were also observed ($p = 0.034$ for the X-axis; $p = 0.037$ for the Y-axis). AI-based programs again positioned this landmark more distally and superiorly, compared to manual tracings. However, in X direction, CephNinja showed greater

similarity to NemoCeph, while in the Y direction, WebCeph™ was more similar to NemoCeph.

Similarly, the gonion point was identified more distally and superiorly by both AI-based software programs. CephNinja's coordinate values in both X and Y directions were closer to those of NemoCeph ($p = 0.035$ for the X-axis; $p = 0.039$ for the Y-axis).

The apex of the upper incisor was also located in a more distal and superior position by both AI software programs. Among them, WebCeph™ provided closer coordinate values to NemoCeph in both axes ($p = 0.041$ for the X-axis; $p = 0.039$ for the Y-axis).

For the apex of the lower incisor, significant differences were again observed in both axes ($p = 0.039$ for the X-axis; $p = 0.045$ for the Y-axis). According to the pairwise comparison, this point was also more distally and superiorly positioned by the AI software, with WebCeph™ providing closer values to NemoCeph than CephNinja.

Intraclass correlation coefficient (ICC) values for all reference points were found to be close to 1, indicating high intra- and inter-observer agreement ($p < 0.01$).

Post-hoc analyses comparing software outputs for each anatomical landmark are presented in Table 4.

Table 4: Post Hoc Analysis Coordinate Values of Anatomical Points.

		NemoCeph - CephNinja	NemoCeph - WebCeph	CephNinja - WebCeph
Porion	X-axis	0.039	0.041	p>0.05
	Y-axis	0.021	0.035	p>0.05
Basion	X-axis	0.031	0.037	p>0.05
	Y-axis	0.027	0.033	p>0.05
Gonion	X-axis	0.031	0.044	p>0.05
	Y-axis	0.043	0.049	p>0.05
Apex of upper incisor	X-axis	0.035	0.046	p>0.05
	Y-axis	0.027	0.039	p>0.05
Apex of lower incisor	X-axis	0.039	0.041	p>0.05
	Y-axis	0.022	0.037	p>0.05

DISCUSSION

Lateral cephalometric radiographs continue to play a key role in orthodontic diagnosis and treatment planning. However, the reliability of these analyses depends heavily on correctly identifying anatomical landmarks. Several studies have pointed out the limitations of conventional methods, especially when it comes to the consistency of landmark placement and measurement accuracy.^{9,10}

Fortunately, with advances in technology, digital systems have made landmark identification and measurements faster and often more reliable.^{11,12} Many digital software programs have been developed, and a large body of research has confirmed that these tools can be as dependable as traditional manual methods. For example, software such as FACAD[®], Cef-X 2001, and PACS have all demonstrated consistent results, when compared to manual tracing, making them suitable for clinical use.¹³⁻¹⁶

Other digital programs like Dolphin Imaging[®], NemoCeph[®], and AutoCEPH[®] have also been widely tested for their performance. Erkan et al.¹⁷ compared several programs, including Dolphin, NemoCeph, Quick Ceph[®], and Vistadent[®], and reported no significant differences between them. Paixão et al.¹⁸ also confirmed the reliability of Dolphin software, showing its results were consistent with manual methods. Similarly, Albarakati et al.¹⁹ found both conventional and digital techniques to be highly reliable, even though they observed small differences in some measurements like ANB angle, convexity angle, and lower anterior facial height (LAFH), which they considered clinically insignificant. Other researchers have highlighted how computer-assisted systems like CADCAS can save time and reduce manual errors.^{20,21}

Even so, landmark identification remains one of the biggest sources of variability in cephalometric analysis.²²⁻²⁶ While digital tools help, they still often rely on the operator's experience and visual judgment, which can be affected by fatigue or other factors. In addition, the high cost of advanced software has sometimes made them less accessible to all clinicians.

More recently, artificial intelligence (AI) has been introduced as a way to improve efficiency and reduce human error in cephalometric analysis. AI-supported software like WebCeph[™] and CephNinja are now commonly used.²⁷⁻²⁹ These programs automatically locate anatomical landmarks, helping to reduce the orthodontist's workload and providing easy access to

cephalometric analysis through computers or mobile devices. Mahto et al.³⁰ showed promising results, reporting good agreement between AI-generated cephalometric measurements from WebCeph™ and traditional manual tracings.

Accurately identifying landmarks is one of the most important steps in cephalometric analysis, because even small errors in positioning can lead to major mistakes in angular and linear measurements. These mistakes can directly affect diagnosis or growth assessment. For example, if landmarks like the porion or nasion are placed incorrectly, it can distort calculations of cranial base length or mandibular plane angles — both of which are crucial for determining skeletal relationships and planning treatment. While many previous studies mainly focused on how accurate the angular or linear measurements were, the present study took a different approach. We specifically examined how much the landmark positions themselves varied by using a coordinate-based comparison across different commercial AI platforms. This provides a clearer picture of where these systems may be inconsistent when marking key anatomical points.

In the present study, we concentrated on the accuracy of landmark placement itself, rather than just measurement outcomes. We used NemoCeph, marked by experienced clinicians, as our gold standard and compared those results to the automatically identified landmarks from WebCeph™ and CephNinja.

Previous research has explored similar comparisons. Kumar et al.³¹ studied Steiner's analysis using CephNinja and NemoCeph and found no significant differences in skeletal parameters, but minor variations in dental measurements like NA and NB. The present study supports these findings, showing no significant differences at the root tips of upper and lower incisors. In another study, Kumar et al.³² found significant differences in Y-axis orientation, the incisor occlusal plane angle, and upper incisor to A-Pog measurements, which aligns with our finding of discrepancies at the porion and incisor apex points.

Duran et al.³³ compared manual tracing, Dolphin software, OrthoDx™, and WebCeph™, and found acceptable consistency overall, but noted that OrthoDx™ performed slightly better than WebCeph™. In our results, WebCeph™ showed greater consistency than CephNinja, particularly in identifying the porion and incisor apex points.

Çoban et al.³⁴ compared Dolphin Imaging and WebCeph™, noting significant differences across several measurements, which they linked to malocclusion type and the difficulty of identifying certain landmarks, like the condyle. Similarly, in the present study, the incisor apex points marked by AI-based software were positioned more distally and superiorly than those identified with NemoCeph, highlighting areas where AI needs further development.

Prince et al.³⁵ compared WebCeph™, AutoCEPH®, and manual methods, finding strong agreement among them and supporting the use of AI-assisted platforms in clinical practice. The present results are consistent with these findings to a degree. While AI software accurately located most landmarks, we still observed inaccuracies at critical points like porion, basion, gonion, and the incisor apex — points that play an essential role in both angular and linear measurements needed for diagnosis and treatment planning.

Overall, as AI technology advances, it offers great potential to simplify and improve cephalometric analysis. However, clinicians should still carefully review AI-generated landmark placements, especially when dealing with key points that directly influence treatment outcomes. Continued research and algorithm refinement will be necessary to make AI-based cephalometric tools even more reliable and clinically effective.

LIMITATIONS

A larger sample size would enhance the statistical power and generalizability of the findings.

Blinded evaluations by multiple orthodontists could reduce observer bias and strengthen the validity of the manual tracing results.

Although NemoCeph was considered the reference standard in this study, additional ICC analyses with more raters could help further validate its reliability.

In future studies, the inclusion of additional anatomical landmarks, such as the occlusal plane, may enhance the comprehensiveness of the study.

CONCLUSION

In this study, we evaluated the accuracy of anatomical landmark identification across manual digital software (NemoCeph) and two AI-based cephalometric software programs (WebCeph™ and CephNinja). It was found that certain anatomical landmarks —porion, basion, gonion, and the apexes of the upper and lower incisors— were located more distally and superiorly in the AI-based software, compared to NemoCeph.

WebCeph™ showed a closer alignment with NemoCeph at the porion and incisor apex points, while CephNinja demonstrated greater similarity at the basion and gonion landmarks. While artificial intelligence has proven effective in most cases, further refinement is required —particularly in points prone to variability or clinical sensitivity.

Clinicians should remain cautious when relying solely on AI-generated cephalometric analyses, and manual validation of key reference points is recommended to ensure diagnostic accuracy and optimal treatment planning.

AUTHORS' CONTRIBUTIONS

İpek Şavkan (IS)

Sara Nur Özçankaya (SNO)

Öykü Naz Turan (ONT)

Conception or design of the study:

IS

Data acquisition, analysis or interpretation:

IS, SNO, ONT

Writing the article:

IS, SNO, ONT

Critical revision of the article:

IS, SNO, ONT

Final approval of the article:

IS, SNO, ONT

Overall responsibility:

IS, SNO, ONT

» Patients displayed in this article previously approved the use of their facial and intraoral photographs.

» The authors report no commercial, proprietary or financial interest in the products or companies described in this article.

» **Data Availability Statement:** Due to sensitive nature of the data and ethical restrictions datasets are not publicly available

REFERENCES

1. Agarwal D. Cephalometric analysis for diagnosis and treatment of orthodontic patients. *J Oral Health Community Dent.* 2013;7(2):75–9. doi: 10.5005/johcd-7-2-75.
2. Sandler PJ. Reproducibility of cephalometric measurements. *Br J Orthod.* 1988;15(2):105-10. doi: 10.1179/bjo.15.2.105.
3. Uzel I, Enacar A. Ortodontide Sefalometri (Cephalometry in Orthodontics). Taş Kitapevi: Ankara; 2000. 2nd ed, p. 3–210.
4. 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. doi: 10.1007/s00056-019-00203-8.
5. McCarthy J. What is artificial intelligence? 2007 [Access 2022 March 29]. Available from: <http://www-formal.stanford.edu/jmc/whatisai/node1.html>.
6. Alqahtani H. Evaluation of an online website-based platform for cephalometric analysis. *J Stomatol Oral Maxillofac Surg.* 2020;121(1):53–7. doi: 10.1016/j.jormas.2019.04.017.
7. Bor S, Kotan S. The influence of Artificial Intelligence in Orthodontics. *J Dent Fac Usak Univ.* 2023;2(1):14-19. doi: izlik.org/JA89FW49XH.

8. Sangeetha RV, Prasanna TR, Manjula KT, Madhusudan V. Comparison of accuracy of different cephalometric analyses using NemoCeph digital software and hand-traced cephalometric analyses: A cephalometric study. *World J Dent.* 2022;13(6):630-640. doi: 10.5005/jp-journals-10015-2098.
9. Nikneshan S, Mohseni S, Nouri M, Hadian H, Kharazifard MJ. The effect of emboss enhancement on reliability of landmark identification in digital lateral cephalometric images. *Iran J Radiol.* 2015;12(2):e19302. doi: 10.5812/iranjradiol.19302.
10. Agrawal MS, Agrawal JSM, Nanjannawar L, Patni V. An evaluation of the reproducibility of landmark identification in traditional versus computer-assisted digital cephalometric analysis system. *Trends Orthod.* 2015;5(3):103-10. doi: 10.4103/2321-1407.155834.
11. Celik E, Polat-Ozsoy O, Memikoglu TUT. Comparison of Cephalometric measurements with digital versus conventional cephalometric analysis. *Eur J Orthod.* 2009;31(3):241– 246. doi: 10.1093/ejo/cjn105.
12. Uysal T, Baysal A, Yagci A. Evaluation of speed, repeatability, and reproducibility of digital radiography with manual versus computer-assisted cephalometric analyses. *Eur J Orthod.* 2009;31(5):523-528. doi: 10.1093/ejo/cjp022.
13. Prawat JS, Nieberg L, Cisneros GJ, Acs G. A comparison between radiographically and sonically produced cephalometric values. *Angle Orthod.* 1995;65(4):271-276.

14. Naoumova J, Lindman R. A comparison of manual traced images and corresponding scanned radiographs digitally traced. *Eur J Orthod.* 2009;31(3):247– 253. doi: 10.1093/ejo/cjn110.
15. Guedes PA, De Souza JEN, Tuji FM, Nery EM. A comparative study of manual vs. computerized cephalometric analysis. *Dental Press J Orthod.* 2010;15(2):44-51. doi: 10.1590/S2176-94512010000200007.
16. Tan SSW, Ahmad S, Moles DR, Cunningham SJ. Picture archiving and communications systems : a study of reliability of orthodontic cephalometric analysis. *Eur J Orthod.* 2011;33(5):537– 543. doi: 10.1093/ejo/cjq116.
17. Erkan M, Gurel HG, Nur M, Demirel B. Reliability of four different computerized cephalometric analysis programs. *Eur J Orthod.* 2012;34(3):318–321. doi: 10.1093/ejo/cjr008.
18. Paixão MB, Sobral MC, Vogel CJ, Araújo TM. Comparative study between manual and digital cephalometric tracing using Dolphin Imaging software with lateral radiographs. *Dental Press J Orthod.* 2010;15(6):123-130. doi: 10.1590/S2176-94512010000600016.
19. Albarakati SF, Kula KS, Ghoneima AA. The reliability and reproducibility of cephalometric measurements: a comparison of conventional and digital methods. *Dentomaxillofac Radiol.* 2012;41(1):11-17. doi: 10.1259/dmfr/37010910.
20. Chen SK, Chen YJ, Yao CC, Chang HF. Enhanced speed and precision of measurement in a computer-assisted digital cephalometric analysis system. *Angle Orthod.* 2004;74(4):501– 507. doi: 10.1043/0003-3219(2004)074<0501:ESAPOM>2.0.CO;2.

21. Sayinsu K, Isik F, Trakyali G, Arun T. An evaluation of the errors in cephalometric measurements on scanned cephalometric images and conventional tracings. *Eur J Orthod*. 2007;29(1):105-108. doi: 10.1093/ejo/cjl065.
22. Shaheed S, Iftikhar A, Rasool G, Bashir U. Accuracy of linear cephalometric measurements with scanned lateral cephalograms. *Pak Oral Dental J*. 2011;31(1):68-72.
23. Baumrind S, Frantz RC. The reliability of head film measurements: 1. Landmark identification. *Am J Orthod*. 1971;60(2):111-27. doi: 10.1016/0002-9416(71)90028-5.
24. Liu JK, Chen YT, Cheng KS. Accuracy of computerized automatic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop*. 2000;118(5):535-40. doi: 10.1067/mod.2000.110168.
25. Goracci C, Ferrari M. Reproducibility of measurements in tablet-assisted, PC-aided, and manual cephalometric analysis. *Angle Orthod*. 2014;84(3):437-42. doi: 10.2319/061513-451.1.
26. Bruntz LQ, Palomo JM, Baden S, Hans MG. A comparison of scanned lateral cephalograms with corresponding original radiographs. *Am J Orthod Dentofacial Orthop*. 2006;130(3):340-8. doi: 10.1016/j.ajodo.2004.12.029.
27. Kublashvili T, Kula K, Glaros A, Hardman P, Kula T. A comparison of conventional and digital radiographic methods and cephalometric analysis software: II. Soft tissue. *Semin Orthod*. 2004;10(3): 212-9. doi: 10.1053/J.SODO.2004.05.005.

28. Segura FJE, Valverde AS, Melendez A, Angelares PRC. Comparative study between digital and manual cephalometry with digital radiographs. *Rev Mex Ortod*. 2014;2(2):e93-e96. doi: 10.1016/S2395-9215(16)30134-9.
29. Kumar D, Solanki R, Sharma R, Khangwal M, Berwal V, Gupta S. Cephalometric landmark identification by conventional and digital direct radiography in skeletal class II individuals. *Int J Res Edu Sci Methods*. 2015;2(12):6-10.
30. 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. doi: 10.1186/s12903-022-02170-w.
31. Kumar M, Shetty P, Shetty P, Kumari S, Singh R. Comparative evaluation of CEPH Ninja for Android and Nemoceph for computer for Steiner's cephalometric analysis. *Ann Int Med Den Res*. 2019; 5(3):DE52-DE56.
32. Kumar M, Kumari S, Chandna A, Konark, Singh A, Kumar H, Punita. Comparative evaluation of CephNinja for Android and NemoCeph for computer for cephalometric analysis: A study to evaluate the diagnostic performance of CephNinja for cephalometric analysis. *J Int Soc Prev Community Dent*. 2020 Jun 15;10(3):286-291. doi: 10.4103/jispcd.JISPCD_4_20.

33. Duran GS, Gökmen Ş, Topsakal KG, Görgülü S. Evaluation of the accuracy of fully automatic cephalometric analysis software with artificial intelligence algorithm. *Orthod Craniofac Res.* 2023;26(3):481-490. doi: 10.1111/ocr.12633.
34. Çoban G, Öztürk T, 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(4):e222112. doi: 10.1590/2177-6709.27.4.e222112.oar.
35. Prince STT, Srinivasan D, Duraisamy S, Kannan R, Rajaram K. Reproducibility of linear and angular cephalometric measurements obtained by an artificial-intelligence assisted software (WebCeph) in comparison with digital software (AutoCEPH) and manual tracing method. *Dental Press J Orthod.* 2023;28(1):e2321214. doi: 10.1590/2177-6709.28.1.e2321214.oar.