Original Article

Evaluating the Accuracy of Automated Cephalometric Analysis Based on Artificial Intelligence

Erum Behroz Khan[1]*, Muhammad Zaheen, Imran Ullah, Asad Ullah, Neelam Shah Jehan, Huda Ghaffar

[1] Department of Orthodontics, Saidu College of Dentistry, Saidu Sharif Swat- Pakistan

Article Info

Corresponding Author
Erum Behroz Khan
Department of Orthodontics, Saidu College of Dentistry, Saidu Sharif, Swat- Pakistan
Email: erumbehroz@gmail.com

Date Received:
22nd August, 2024
Date Revised:
10th March, 2025
Date Accepted:
11th March, 2025

Abstract

Objective: To determine the correlation and comparison between artificial intelligence (AI) cephalometric tracing and manual human tracing of cephalograms.

Methodology: This cross sectional comparative study was conducted on cephalograms from 100 participants at Saidu College of Dentistry, Swat, using a non-probability sampling technique. Informed consent was not required as the study utilized existing radiographic records with prior patient consent. Age and gender were documented, followed by the manual tracing of 100 cephalograms by a sole operator, which were subsequently analyzed using AudaxCeph software. Pearson correlation testing and paired t-tests were utilized for analysis.

Results: The participants' average age was 19.39 years (SD ± 6.24), with a gender distribution of 56% female and 44% male. The results revealed strong positive correlations (0.91 to 0.98) between manually and software-traced cephalometric parameters, with statistically significant p-values (<0.01). Paired t-tests showed no significant differences across various parameters, affirming the reliability of the software-generated cephalometric measurements.

Conclusion: This study demonstrates the software's accuracy in assessing specific cephalometric parameters, highlighting its potential as a reliable tool in orthodontic diagnosis and treatment planning.

Keywords: Artificial intelligence, Cephalometry, Deep Learning, Neural Networks, Orthodontics, X-rays

This article may be cited as:
Khan EB, Zaheen M, Ullah I, Ullah A, Jehan NS, Ghaffar H. Evaluating the accuracy of automated cephalometric analysis based on artificial intelligence. J Postgrad Med Inst. 2025;39(1):1-9. http://doi.org/10.54079/jpmi.39.1.3473

Introduction

In cephalometric analysis, precise landmark localization is vital for diagnosing orthodontic and orthognathic conditions, planning treatments, and evaluating outcomes. Nevertheless, difficulties emerge during the transition from a three-dimensional skull to a two-dimensional cephalogram.[1] This shift brings about complexities such as superimposition and distortion, resulting in variations in landmark definition.. Even experienced orthodontists find the manual identification of all cephalometric landmarks to be a time-consuming task, taking more than half an hour.[2]

Examining lateral cephalometric radiographs is crucial in orthodontics for the diagnosis and planning of treatments, as well as for assessing skeletal and dental relationships.[3] It predicts growth directions in youths and evaluates orthodontic outcomes, which is crucial in planning adult orthognathic surgery for skeletal malocclusions. Contemporary digital analysis entails the identification of cephalometric landmarks utilizing specialized software. Notably, recent progress integrates artificial intelligence (AI) to streamline analysis processes, thus reducing clinicians' workload. The increasing prevalence of AI-powered image analysis applications represents a significant trend, providing orthodontists with improved tools for conducting cephalometric evaluations more efficiently.[4]

Recognizing the labor-intensive, experience-dependent, and error-prone nature of manual localization; efforts have been directed towards incorporating artificial intelligence (AI) into the process.[5] In recent years, both traditional machine-learning methods and advanced deep-learning techniques, particularly convolutional neural networks (CNNs), have been employed in cephalometric analysis.[6] While conventional approaches favored the use of Random Forest in conjunction with geometric or shape information, the adoption of CNNs has demonstrated significant enhancements in accuracy and robustness. This transition towards AI-based methodologies holds the potential for more streamlined and precise clinical applications in the realm of cephalometric analysis.[7]

The comparison between AI-driven and manual cephalometric tracing focuses on key aspects of diagnostic accuracy, time efficiency, adaptability, cost-effectiveness, and user accessibility. AI, with its advanced image recognition and machine learning capabilities, enhances precision and consistency, minimizing human error. Its automation significantly reduces analysis time, improving workflow efficiency. Moreover, AI continuously adapts to evolving diagnostic criteria, addressing the limitation of manual tracing. While the initial investment in AI may be high, it has the potential for long-term cost savings compared to labor-intensive manual methods. Additionally, AI's user-friendly interfaces offer greater accessibility than expertise-dependent manual tracing. Given the absence of local studies and the anatomical variations across populations, this research provides region-specific insights into the feasibility of AI in cephalometric analysis.

The objective of this study was to evaluate the correlation and compare cephalometric tracing performed by artificial intelligence with manual tracing conducted by humans.

Methodology

This comparative study was carried out on cephalograms obtained from 100 participants at Saidu College of Dentistry, Swat, spanning from October 1, 2023, to January 30, 2024. The cephalograms were collected from the Department of Orthodontics at Saidu College of Dentistry using a non-probability sampling technique. Ethical approval was secured from the relevant ethical review committee (Approval No. 147-ERB/023). Informed consent was not necessary since this study utilized existing radiographic records, and patients had previously consented to the use of their records for research purposes. The inclusion criteria were age above 10 years, both genders and Pakistani nationality. Cephalograms that were unclear, patients with syndromes or malformations, cleft palate cases, and those with other bone pathologies were excluded.

The sample size was calculated using a correlation calculator, considering the correlation of the nasion point between AI and manual methods as 0.836, derived from a previous study1. The calculation was conducted with a 5% Type I error and 90% power. The initially determined sample size was 10; however, to meet the normality assumption, 100 participants were included in the study.

Age and gender were recorded from patients' record files. One hundred cephalograms were manually traced by a single operator using a lightbox, acetate tracing paper, and a lead pencil. Images of 100 lateral cephalograms were obtained from the radiology department and imported into AudaxCeph software. The following cephalometric parameters were assessed (see Table I)

Data analysis was conducted using SPSS version 22. For numeric data, mean and standard deviation (SD) were computed. Pearson correlation tests were utilized to evaluate the correlation between cephalometric parameters traced manually and those traced by the software. Furthermore, a paired t-test was employed to compare cephalometric parameters between the two methods. The significance level was established at $p < 0.05$.

Statistical analysis

Data analysis was conducted using SPSS version 22. For numeric data, mean and standard deviation (SD) were computed. Pearson correlation tests were utilized to evaluate the correlation between cephalometric pa-

rameters traced manually and those traced by the software. Furthermore, a paired t-test was employed to compare cephalometric parameters between the two methods. The significance level was established at $p < 0.05$.

Results

In this study involving 100 participants, the demographic characteristics were analyzed to provide a comprehensive understanding of the study population. The mean age of the participants was 19.39 years (SD ± 6.24). The gender distribution revealed that 56 participants (56%) were female, while 44 participants (44%) were male. This signifies a slight majority of female participants in the study cohort. Further exploration of age groups demonstrated an equal representation, with 56 participants (56%) falling within the 10-20 age range and 44 participants (44%) in the 21-30 age bracket.

Table II presents the correlation analysis between cephalometric parameters traced using both software and manual methods. The correlation coefficient (Correlation*) signifies the strength and direction of the relationship, with associated statistical measures such as the test statistic, p-value, and confidence intervals (Conf. Low and Conf. High). The Pearson correlation test was employed for this analysis. Significantly, robust positive correlations ranging from 0.91 to 0.98 were noted for all parameters, indicating a high level of agreement between the two tracing methods. The associated p-values, all less than 0.01, further support the statistical significance of these correlations.

Utilizing a paired t-test, no statistically significant differences were found between AI and manual tracing across various parameters: SNA (p = 0.23), SNB (p = 0.12), SNPP (p = 0.14), SNMd (p = 0.08), MMA (p = 0.14), PFH/AFH (p = 0.13), UIPP (p = 0.35), IMPA (p = 0.53), and IIA (p = 0.40). (Table III)

Table 1. Operational definitions of cephalometric parameters

Parameter	Measurement
SNA (Sella-Nasion-A Point)	Angle formed by lines through Sella (S), Nasion (N), and A Point.
SNB (Sella-Nasion-B Point)	Angle formed by lines through Sella (S), Nasion (N), and B Point.
SNPP (Sella-Nasion to Palatal Plane)	Angle formed by lines through Sella (SN) and the Palatal Plane (ANS to PNS).
SNMd (Sella-Nasion-Menton)	Angle formed by lines through Sella (SN), Nasion (N), and Menton (M).
PFH (Posterior Facial Height)	Straight distance in mm from Sella (SN) to Menton (M).
AFH (Anterior Facial Height)	Measurement from Nasion (N) to Menton (M).
UIPP (Upper Incisor to Palatal Plane)	Angle formed by the long axis of the upper incisor and the Palatal Plane.
IMPA (Incisor-Mandibular Plane Angle)	Angle formed by the long axis of the lower incisor and the Mandibular Plane.
IIA (Interincisal Angle)	Angle formed by the long axes of the upper and lower incisors.
MMA	The angle formed between the maxillary plane (anterior nasal spine to posterior nasal spine) and the mandibular plane (gonion to menton).

Table 2. Correlation of cephalometric parameters traced with software versus manual

cephalometric parameter	correlation*	statistic	p-value	conf.low	conf.high
SNA	0.94	26.73	0.007	0.91	0.96
SNB	0.91	21.22	0.008	0.86	0.94
SN Md	0.94	26.73	0.009	0.91	0.96
SNPP	0.97	38.7	0.002	0.95	0.98
MMA	0.97	41.44	0.003	0.96	0.98
PFH/AFH	0.95	30.64	0.006	0.93	0.97
UIPP	0.98	44.61	0.001	0.96	0.98
IMPA	0.97	41.13	0.004	0.96	0.98
IIA	0.97	41.29	0.005	0.96	0.98

*Pearson correlation test

J Postgrad Med Inst 2025;39(1):242-46.
http://doi.org/10.54079/jpmi.39.1.3473

Table 3. Correlation of cephalometric parameters traced with software versus manual

Cephalometric parameter	AI, N = 100	mannual, N = 100	p-value*
SNA	80.10 ± 2.90	80.58 ± 2.77	0.23
SNB	79.44 ± 2.72	80.42 ± 2.73	0.12
SNPP	10.38 ± 3.03	9.25 ± 1.75	0.14
SNMd	33.48 ± 3.69	34.78 ± 3.21	0.08
MMA	23.98 ± 4.71	24.64 ± 4.72	0.61
PFH/AFH	65.60 ± 3.91	64.40 ± 2.75	0.13
UIPP	112.87 ± 7.20	113.85 ± 7.62	0.35
IMPA	93.37 ± 7.04	92.77 ± 6.49	0.53
IIA	128.70 ± 11.32	127.40 ± 10.29	0.4

Discussion

The finding of our study is perfect correlation between cephalometric parameters traced manually and those traced using the AudaxCeph AI tool.

Kiełczykowski et al.[3] investigated the growing influence of artificial intelligence (AI) in medicine and dentistry, highlighting its potential to improve diagnostic precision and decrease errors. Their study aimed to evaluate the efficacy of AI in orthodontic diagnostics through the analysis of lateral cephalometric radiographs. Through a review spanning 2009 to 2023 across databases like PubMed and Scopus, they found that AI algorithms generally showed high accuracy in positioning cephalometric landmarks. The study concluded that AI holds promise in streamlining orthodontic treatment planning and improving radiological examinations, potentially surpassing manual analysis in accuracy in the future.[4]

Another systematic review found that convolutional neural networks (CNN)-based AI algorithms achieved point localization accuracy ranging from 64.3% to 97.3%, with a mean error between 1.04 mm ± 0.89 and 3.40 mm ± 1.57, generally within the clinically acceptable range of 2 mm. Despite these promising results, the study noted that AI systems have not yet matched the accuracy levels of experienced orthodontists.[5]

Ristau B conducted a comparison between the accuracy of cephalometric landmark identification using an automated tracing software, AudaxCeph®, based on convolutional neural networks, and human tracers (two board-certified orthodontists). Sixty cephalograms were examined, with thirteen landmarks identified. The research revealed no statistical disparity between the performance of the orthodontists and AudaxCeph®'s automatic tracing software, except for certain dimensions of Porion and the lower incisor apex.[6]

Mahto RK, et al.,[7] conducted a comparison between cephalometric measurements derived from the fully automated AI-driven platform "WebCeph"™ and manual tracing, evaluating the accuracy and consistency of the automated measurements. Thirty pre-treatment lateral cephalograms were utilized, with 18 landmarks plotted for manual tracing and 12 measurements obtained. Digital images were also uploaded to "WebCeph"™ for automated measurements. The study found that all measurements had an Intraclass correlation coefficient (ICC) value above 0.75, indicating good reliability. Seven parameters showed a higher ICC value (>0.9), while five parameters demonstrated an ICC value between 0.75 and 0.90. Overall, the findings suggest that "WebCeph"™ provides valid and reliable automated cephalometric measurements.[8-10] These results support our findings.

Kunz F et al.[11] evaluated the precision of skeletal and dental cephalometric parameters from different commercial providers through AI-assisted automated analysis. Twelve orthodontic experts marked landmarks on 50 X-rays to establish a "humans' gold standard." Statistical analyses were conducted to compare results from four providers, demonstrating the potential of fully automated cephalometric analyses in terms of efficiency and error reduction.

The study is limited by its single-center design and reliance on a single operator for manual tracing, which may introduce bias. Additionally, the use of a non-probability sampling technique limits the generalizability of the findings.

Conclusion

These findings suggested that the software-generated cephalometric measurements closely align with manual tracings, validating the reliability of the software in assessing these specific parameters.

References

1. Jiang F, Guo Y, Yang C, Zhou Y, Lin Y, Cheng F, et al. Artificial intelligence system for automated landmark localization and analysis of cephalometry. Dentomaxillofac Radiol 2023;52(1):20220081.
2. Yao J, Zeng W, He T, Zhou S, Zhang Y, Guo J, et al. Automatic localization of cephalometric landmarks based on convolutional neural network. Am J Orthod Dentofacial Orthop 2022;161(3):e250-9.
3. Kiełczykowski M, Kamiński K, Perkowski K, Zadurska M, Czochrowska E. Application of artificial intelligence (AI) in a cephalometric analysis: a narrative review. Diagnostics 2023;13(16):2640.
4. Hendrickx J, Gracea RS, Vanheers M, Winderickx N, Preda F, Shujaat S, et al. Can artificial intelligence-driven cephalometric analysis replace manual tracing? A systematic review and meta-analysis. Eur J Orthod 2024;46(4):cjae029.
5. Ribas-Sabartés J, Sánchez-Molins M, d'Oliveira NG. The accuracy of algorithms used by artificial intelligence in cephalometric points detection: a systematic review. Bioengineer 2024;11(12):1286.
6. Ristau B, Coreil M, Chapple A, Armbruster P, Ballard R. Comparison of AudaxCeph®'s fully automated cephalometric tracing technology to a semi-automated approach by human examiners. Int Orthod 2022;20(4):100691.
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.
8. 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):1-8.
9. Danisman H. Artificial intelligence web-based cephalometric analysis platform: comparison with the computer assisted cephalometric method. Clin Investig Orthod 2023;82(4):194-203.
10. Bor S, Ciğerim SÇ, Kotan S. Comparison of AI-assisted cephalometric analysis and orthodontist-performed digital tracing analysis. Prog Orthod 2024;25(1):41.
11. Kunz F, Stellzig-Eisenhauer A, Widmaier LM, Zeman F, Boldt J. Assessment of the quality of different commercial providers using artificial intelligence for automated cephalometric analysis compared to human orthodontic experts. J Orofac Orthop 2023;23:1-16.

Authors' Contribution Statement

EBK contributed to the study's conception, design, literature search, and supervision. MZ contributed to data analysis, data collection, writing the manuscript, and literature search. IU contributed to the literature search and data collection. AU contributed to figures and data collection. NSJ contributed to the literature search and data collection. HG contributed to the literature search and data collection. All authors are accountable for their work and ensure the accuracy and integrity of the study.

Confilct of Interest

Authors declared no conflict on interest

Grant Suppport and Financial Disclosure

None

Data Sharing Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.