Open Access Article

Ann. Acad. Med. Siles. (online) 2025; 79: 222–225
eISSN 1734-025X
DOI: 10.18794/aams/205193
www.annales.sum.edu.pl

PRACA ORYGINALNA
ORIGINAL PAPER

Manual tracing and artificial intelligence tracing of lateral cephalograms: A critical comparative assessment of performance

Qazi Jawad Hayat, Barq Rehman Khan, Ziad Khan, Muhammad Kashif, Muhammad Abdullah Zuhair

Khyber College of Dentistry, Peshawar, Pakistan / Khyber Medical University, Peshawar, Pakistan

ABSTRACT

INTRODUCTION: Cephalometric analysis, a cornerstone of orthodontics and craniofacial surgery, traditionally involves manual radiograph tracing, a time-consuming and potentially variable process. Artificial intelligence (AI) offers a potential alternative for faster, more consistent analysis. This study compared AI-driven and manual cephalometric methods to assess agreement and identify discrepancies.

MATERIAL AND METHODS: This quantitative, comparative cross-sectional study was conducted in a private practice in Peshawar, Pakistan (August–November 2024), including 29 orthodontic patients who met specific criteria (good-quality cephalograms and absence of facial clefts/intra-oral appliances). Cephalometric radiographs were analyzed by two experienced dentists using manual tracing and by AI software (Audaxceph 6.0.50.3887). Five key angular measurements (SNA, SNB, ANB, FMA, and SN-Mp), used in Steiner's and Tweed's analyses, were compared. Inter-rater reliability for the manual tracings was assessed using intraclass correlation coefficients (ICCs).

RESULTS: Excellent inter-rater reliability was observed for manual tracings (ICCs > 0.90). Paired t-tests revealed no significant differences between manual and AI methods for SNA, SNB, ANB, and FMA. However, a statistically significant difference ($p = 0.006$) was found for SN-Mp.

CONCLUSIONS: This study, comparing manual and AI-driven cephalometric analysis, found strong agreement for most key measurements (SNA, SNB, ANB, and FMA), suggesting AI's potential to enhance clinical efficiency. The significant difference in SN-Mp, however, emphasizes the need for continued clinical oversight. A combined approach, integrating AI with clinical expertise, is recommended for optimal diagnostic accuracy and treatment planning.

KEYWORDS

artificial intelligence, AI, cephalography, cephalogram, tracing

Received: 19.01.2025 **Revised:** 15.05.2025 **Accepted:** 15.05.2025 **Published online:** 14.08.2025

Address for correspondence: Dr. Qazi Jawad Hayat, Khyber College of Dentistry, Peshawar, Khyber Pakhtunkhwa, Pakistan, tel. 00923339443452, e-mail: qazijawadhayat@gmail.com

Publisher: Medical University of Silesia, Katowice, Poland

INTRODUCTION

Cephalometric analysis has a rich history dating back to the late 1800s, when radiographs were first employed to study the head and neck. In the 1930s, Holly Broadbent, a professor of orthodontics at the University of Michigan, analyzed the correlation between the teeth and the skull. This pioneering work involved measuring various angles and distances on radiographic images, establishing the foundations of cephalometric analysis [1]. Cephalometric analysis was initially coined to describe manually locating landmarks on acetate overlays over a light table and measuring the linear and angular values with a protractor, which is tedious, time-consuming, and subjective [2]. Cephalometric skeletal analysis plays a pivotal role in orthodontics and craniofacial surgery, serving as a critical diagnostic tool for understanding craniofacial structure and function. This method is essential for diagnosing skeletal discrepancies, evaluating growth patterns, and planning treatment strategies for individuals with orthodontic and craniofacial conditions [3]. Traditionally, this process has relied on manual techniques, demanding a high level of expertise and attention to detail from clinicians. However, recent advancements in artificial intelligence (AI) have introduced a new dimension to cephalometric analysis, challenging the dominance of manual methods. AI is concerned with developing programs and computers that can gather data, apply reason to it, and then translate it into intelligent actions. AI is a broad area that includes reasoning, typical linguistic dispensation, machine learning, and planning. The manual approach to cephalometric skeletal analysis involves identifying and marking anatomical landmarks on radiographs, followed by precise measurements to assess craniofacial relationships. While this approach has been the standard for decades, it is not without its challenges [4]. The proces is time-consuming, labor-intensive, and prone to variability, as the accuracy of the results often depends on the skill and experience of the practitioner [5,6]. Inter- and intra-observer inconsistencies can lead to discrepancies in measurements, affecting both diagnosis and treatment outcomes [7]. Convolutional neural networks, a type of deep learning model, have various applications, including image classification and segmentation, natural language processing, facial landmark detection, and lane detection [8].

AI-driven tools can analyze cephalometric images with remarkable speed and precision, minimizing human error and variability. AI-based methods of cephalometric analysis can be semi-automatic or fully automatic. The fully automatic method uses AI to trace, identify landmarks, and calculate the cephalometric measurements, whereas the semi-automatic method involves a combination of manual selection of landmarks followed by automated calculation of values [9,10].

These systems are trained on large datasets, enabling them to recognize complex patterns and deliver consistent results across different cases [11]. As such, AI holds the promise of revolutionizing cephalometric skeletal analysis by enhancing accuracy, efficiency, and accessibility. Some studies have found statistically significant differences between manual and AI-based methods. Hwang et al. [12] reported AI to be more accurate than a manual method for 14 out of the 46 landmarks measured in their study, while another 14 variables were found to be more accurately measured by the manual method as compared to the AI-based method, and similar results were obtained by Agrawal et al. [13].

The aim of this study was to comprehensively compare AI-driven and manual methods for cephalometric skeletal analysis. By examining the strengths, limitations, and practical applications of both approaches, this research can provide a nuanced understanding of their respective roles in clinical practice. Key aspects will be explored – such as accuracy, reliability, and efficiency – along with the implications of integrating AI into orthodontic and craniofacial workflows. The introduction of AI into cephalometric skeletal analysis marks a significant step forward in the evolution of diagnostic methodologies. As the field continues to evolve, understanding the interplay between traditional expertise and technological innovation becomes increasingly important. By critically analyzing these two approaches, this research sheds light on their potential to complement each other and drive advancements in patient care and clinical outcomes.

MATERIAL AND METHODS

This is a quantitative, comparative cross-sectional study, conducted at a private practice in Peshawar, Khyber Pakhtunkhwa, Pakistan between August 2024 and November 2024. Prior to its commencement, proper informed consent was taken from the patients. Initially, the study involved 55 patients; after applying the inclusion and exclusion criteria, 29 patients were enrolled in the study. The inclusion criteria consisted of subjects seeking orthodontic treatment whose records included cephalometric X-rays. The exclusion criteria were a lack of consent, poor quality cephalograms, cephalograms showing artifacts, any history of facial clefts, and use of intra-oral appliances. No restrictions were placed on the gender, age, or ethnicity of the patients.

All the measurements were based on the American Board of Orthodontics Analysis and included the angles SNA and SNB, as well as ANB, which is SNA minus SNB. The AI tracing was done with the help of Audaxceph version 6.0.50.3887. The manual tracing and evaluation of the cephalograms were conducted by two proficient dentists. For this study, five specific readings from the traced cephalograms were included, focusing on the angles SNA, SNB, SN-Mp, and FMA as per Steiner's and Tweed's skeletal analysis. Descriptive statistics such as mean, median, and standard deviation were obtained for age and percentages for gender. For intergroup comparisons, a sample paired t-test was used. All the data are described in tables and charts.

RESULTS

The frequencies of male and female patients participating in this study were in a ratio of 31.03% to 68.96%, respectively (9 females and 20 males). The mean age of the sample was 18.10 years with standard deviation of 5.690 and a median of 16.0, as shown in Table I.

Table I. Statistics with regard to the participants' age

Age (years)		
N	valid missing	29 0
Mean		18.10
Median		16.00
Std. deviation		5.690
Minimum		7
Maximum		30

The paired sample t-test showed no significant differences between the values for SNA, SNB, ANB, and FMA that were traced manually versus those traced with the help of AI, although a p-value of 0.006 was obtained for SNMP, as shown in Table II.

Table II. Paired sample t-test

	Pairs	Paired differences							
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference		t	df	Sig. (2-tailed)
					lower	upper			
Pair 1	Angle SNA Manual – Angle SNA Ai	-.9724	3.7192	.6906	-2.3871	.4423	-1.408	28	.170
Pair 2	Angle SNB Manual – Angle SNB Ai	-.0966	2.7197	.5050	-1.1311	.9380	-.191	28	.850
Pair 3	ANB Values Manual – ANB Values Ai	-.8724	3.5360	.6566	-2.2174	.4726	-1.329	28	.195
Pair 4	SN-MP Angle Manual – SN-MP Angle Ai	-2.6552	4.8426	.8992	-4.4972	-.8132	-2.953	28	.006
Pair 5	FMA Angle Manual – FMA Angle Ai	-1.1793	6.3147	1.1726	-3.5813	1.2227	-1.006	28	.323

Manual cephalometric tracings were performed independently by two experienced dentists. Inter-rater reliability was assessed using intraclass correlation coefficients (ICCs). High levels of agreement were observed for all measurements (ICCs > 0.90 for all variables), indicating excellent inter-rater reliability.

DISCUSSION

This study presents a detailed analysis comparing five key cephalometric parameters measured manually by two experienced dentists and using AI-driven software. The excellent inter-rater reliability observed for the manual tracings (ICCs > 0.90 for all variables) confirms the consistency and accuracy of the manual method and provides a strong baseline for comparison with the AI-based approach. This high level of agreement between the dentists (manual raters) strengthens the validity of the study's findings.

The core finding of this research is the general agreement between AI-driven and manual cephalometric analysis. For the majority of the parameters assessed (SNA, SNB, ANB, and FMA), no statistically significant differences were found between the two methods. This suggests that AI-based tools and applications can provide comparable results to traditional manual tracing, offering a faster and more efficient alternative in clinical practice. The efficiency gained by using AI is a significant advantage, particularly in busy clinical settings where time constraints are a major issue. AI can process images and generate measurements much faster than manual methods, freeing up clinicians' time for other essential tasks, such as interacting with patients and planning treatments.

However, a statistically significant difference was observed for SNMP ($p = 0.006$), indicating a potential discrepancy between the manual and AI tracing methodologies specific to this measurement. This finding is crucial and requires further investigation and research. It suggests that the AI algorithm may have difficulty accurately identifying the specific landmarks or performing the calculations involved in determining SNMP angle. This discrepancy could be due to several factors, including the complexity of the anatomical structures involved in SNMP measurement, variations in image quality, or limitations in the AI's training data. Further research is needed to pinpoint the exact cause of this difference and to explore potential solutions, such as refining the AI algorithm or improving image acquisition protocols.

This finding aligns with the research published by Mercier et al. in 2024 [11], which suggests that current AI technology has not yet reached a level of 100% accuracy in landmark detection. While AI has made significant improvement in image analysis, challenges remain in accurately identifying complex anatomical landmarks in all cases. This highlights the importance of continued research and development in AI-driven cephalometric analysis to improve its accuracy and reliability. It also emphasizes the need for clinicians to take great caution when interpreting AI-generated results, particularly for parameters where discrepancies have been identified.

The implications of these findings for clinical practice are significant. While AI offers the potential for more efficiency and less variability in cephalometric analysis, it is not yet a perfect replacement for manual methods. Clinicians should be aware of the potential limitations of AI-based tools, particularly for parameters like SNMP. A hybrid approach that combines AI-driven analysis with clinical expertise and judgment may be the most effective strategy for the foreseeable future. This approach would leverage the speed and efficiency of AI while ensuring accurate, reliable results through a careful review and interpretation by experienced clinicians.

Authors' contribution

Study design – Q.J. Hayat, B.R. Khan
Data collection – B.R. Khan, M. Kashif
Data interpretation – Q.J. Hayat, Z. Khan
Statistical analysis – Q.J. Hayat, M. Kashif
Manuscript preparation – Q.J. Hayat, Z. Khan, M.A. Zuhair
Literature research – Q.J. Hayat, B.R. Khan

REFERENCES

1. Subramanian A.K., Chen Y., Almalki A., Sivamurthy G., Kafle D. Cephalometric analysis in orthodontics using artificial intelligence–A comprehensive review. Biomed. Res. Int. 2022; 2022: 1880113, doi: 10.1155/2022/1880113.

2. Houston W.J., Maher R.E., McElroy D., Sherriff M. Sources of error in measurements from cephalometric radiographs. Eur. J. Orthod. 1986; 8(3): 149–151, doi: 10.1093/ejo/8.3.149.

3. Nijkamp P.G., Habets L.L., Aartman I.H., Zentner A. The influence of cephalometrics on orthodontic treatment planning. Eur. J. Orthod. 2008; 30(6): 630–635, doi: 10.1093/ejo/cjn059.

4. Dinesh A., Mutalik S., Feldman J., Tadinada A. Value-addition of lateral cephalometric radiographs in orthodontic diagnosis and treatment planning. Angle Orthod. 2020; 90(5): 665–671, doi: 10.2319/062319-425.1.

5. Kim D.I., Lagravère Vich M., Mayoral P., Miguez M. Three-dimensional changes in skeletal/ dental landmarks with use of mandibular advancement devices. J. Dent. Sleep Med. 2020; 7(2): 1–18, doi: 10.15331/jdsm.7120.

6. Ibragimov B., Likar B., Pernus F., Vrtovec T. A game-theoretic framework for landmark-based image segmentation. IEEE Trans. Med. Imaging 2012; 31(9): 1761–1776, doi: 10.1109/TMI.2012.2202915.

7. Azad M., Elaiwat S., Alam M.K. A computational tool for detection of soft tissue landmarks and cephalometric analysis. Electronics 2022; 11(15): 2408, doi: 10.3390/electronics11152408.

8. Lee H.T., Chiu P.Y., Yen C.W., Chou S.T., Tseng Y.C. Application of artificial intelligence in lateral cephalometric analysis. J. Dent. Sci. 2024; 19(2): 1157–1164, doi: 10.1016/j.jds.2023.12.006.

9. Albarakati S.F., Kula K.S., Ghoneima A.A. 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.

10. Yassir Y.A., Salman A.R., Nabbat S.A. The accuracy and reliability of WebCeph for cephalometric analysis. J. Taibah Univ. Med. Sci. 2021; 17(1): 57–66, doi: 10.1016/j.jtumed.2021.08.010.

11. Mercier J.P., Rossi C., Sanchez I.N., Renovales I.D., Sahagún P.M.P., Templier L. Reliability and accuracy of artificial intelligence-based software for cephalometric diagnosis. A diagnostic study. BMC Oral Health 2024; 24(1): 1309, doi: 10.1186/s12903-024-05097-6.

12. Hwang H.W., Park J.H., Moon J.H., Yu Y., Kim H., Her S.B. et al. Automated identification of cephalometric landmarks: *Part 2-Might it be better than human?* Angle Orthod. 2020; 90(1): 69–76, doi: 10.2319/022019-129.1.

13. Agrawal M.S., Agrawal J.M., Patni V., Nanjannawar L. An evaluation of the reproducibility of landmark identification in traditional versus computer--assisted digital cephalometric analysis system. APOS Trends Orthod. 2015; 5(3): 103–110, doi: 10.4103/2321-1407.155834.