Evaluation Of Accuracy And Reliability Of Artificial Intelligence-Based Fully Automated And Semi-Automated Cephalometric Analysis Software In Comparison With Manual Cephalometric Analysis

Arshya Kumar[1], Sirirengalakshmi Muthuswamy Pandian[2]

[1]*Postgraduate Student, Department of Orthodontics and Dentofacial Orthopedics, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai 600077*
[2]*Professor Department of Orthodontics and Dentofacial Orthopedics, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai 600077*

KEYWORDS

artificial intelligence, cephalometric analysis, digital software, digital manual software

ABSTRACT

Introduction: Lateral cephalometric analysis stands as a pivotal diagnostic tool within the realm of Orthodontics, often regarded as the gold standard. Over time, several computerised programs have surfaced, aiming to streamline the process of digital cephalometric analysis, thereby offering efficiency gains and time savings. The current study sought to assess the diagnostic accuracy and reliability of four computerised cephalometric analysis programs. These encompassed three digital software solutions alongside an online artificial intelligence platform. The objective was to compare the digital approaches against the traditional method of manual cephalometric tracing.

Materials and methodology: This study analysed thirty pre-treatment lateral cephalometric radiographs of individuals undergoing orthodontic treatment. Four computer programs—Dolphin Imaging software, Nemoceph, and FACAD for semi-automated cephalometric analysis, and WebCeph for artificial intelligence-based cephalometric analysis—were evaluated and compared with manual cephalometric tracing. Measurements were taken for five skeletal, four dental, and one soft tissue parameters, including six angular and four linear measurements derived from Steiner, McNamara, and Tweed cephalometric analysis. Fifteen cephalometric radiographs were remeasured to assess the intra-examiner correlation coefficient. One-way ANOVA was used to compare different cephalometric points among the various tracing methods, and multiple comparisons were performed using the Post Hoc Tukey's test.

Results: No significant difference was noted in SNA, SNB, ANB, FMA, IMPA, 1 to Pt A, 1 to N-A and nasolabial angle between the software. While 1 to A Pog and 1 to N-B revealed significant differences between the software. FACAD vs. NemoCeph showed a significant difference with a mean difference of -0.30 and -0.30 and a p-value of 0.03 and 0.02 for 1 to A Pog and 1 to N-B values respectively and this variability was noted between FACAD and NemoCeph.

Conclusion: Digital and AI-based software have shown accuracy comparable to manual tracing. Among the software, Dolphin Imaging proved to be the most reliable, followed by WebCeph and NemoCeph, with FACAD showing the most variability. The variability, which lacked clinical significance, was primarily due to challenges in identifying certain landmarks.

1. Introduction

Broadbent pioneered cephalometric analysis in 1931, and it remains an indispensable aspect of modern clinical orthodontics and orthognathic surgery. [1] This method is vital for assessing craniofacial structures and identifying anatomical elements that contribute to malocclusion, serving as a pivotal aspect of orthodontic diagnosis. [2–4] Moreover, cephalometric analysis aids in growth prediction, treatment planning, and evaluating treatment outcomes. [5–7] The conventional approach entails marking radiographic points on transparent acetate sheets and measuring linear and angular values with a protractor and ruler—a manual method that persists as the gold standard and stands out as one of the most efficient and commonly utilized diagnostic tools in orthodontics. [8–10] Nevertheless, traditional cephalometric analysis faces constraints such as magnification of radiographic films, difficulties in identifying landmarks, laborious and time-consuming manual procedures, and increased susceptibility to measurement inaccuracies. [11,12]

Advances in computing technology have resulted in the extensive integration of computer software into orthodontic practices for cephalometric analysis. [13–15] Healthcare professionals can derive benefits from employing different approaches to integrate new technology. [16–18] The use of computer software and applications has replaced the conventional manual method for cephalometric analysis. [19] Computer-assisted cephalometric analysis automates angle and distance calculations, eradicating errors inherent in traditional methods and streamlining data acquisition. Moreover, digital enhancements of lateral cephalogram images improve visualization, aiding observers in landmark identification. [20]

Cephalometric analysis can be conducted using a variety of mobile applications and computer programs, including those tailored for fully automated or artificial intelligence (AI)-driven analysis, as well as digital manual (DM) or semi-automated cephalometric analysis. [14,21]

AI-powered applications, like the WebCeph web program, can automatically identify landmarks and perform measurements once the digital radiograph is imported. [22] In DM, the operator is required to manually identify landmarks, after which measurements are automatically calculated, as demonstrated in software such as Dolphin Imaging, NemoCeph, and FACAD. [23,24] Though crucial for orthodontic diagnosis and treatment planning, these applications and computer software are hindered by limited accessibility stemming from their expensive nature and the requirement for a laptop or desktop computer. [25,26] Considering their pivotal role, it is important to regularly validate the accuracy of these programs.

Unlike the conventional method of manually tracing cephalometric images, the present study evaluated the diagnostic accuracy and reliability of four computerized cephalometric analysis tools. These tools comprised three digital manual programs and one web-based platform utilizing Artificial intelligence. The goal was to aid clinicians in determining the most precise option for cephalometric evaluations. According to the null hypothesis, no significant difference was expected between the cephalometric measurements generated by Dolphin Imaging, Nemoceph, FACAD, and WebCeph software and those acquired through manual cephalometric tracing.

2. Methodology

Thirty pre-treatment lateral cephalometric X-rays of individuals seeking orthodontic care were analyzed at the Department of Orthodontics and Dentofacial Orthopedics at Saveetha Dental College and Hospital in Chennai. High-quality images were captured with patients in their natural head position using the Carestream 9600 X-ray unit (New York, USA) operating at 140 KHz, 15 mA current, and an exposure time of 8 seconds. The selection of these X-rays was based on the subjects' skeletal malocclusion. The study included patients aged 18–32 years, with a mean age of 22.4 ± 3 years, while those exhibiting significant asymmetry, syndromes, low-quality radiographs, incorrect head positioning, or any other factors impeding landmark identification were excluded from the study.

The reliability of measurements was checked by three examiners, with precautions taken to avoid errors due to examiner fatigue. Various skeletal, dental, and soft tissue parameters were measured, including linear and angular measurements. To ensure accuracy, one operator conducted both manual and digital tracings, while two additional manual operators were involved to establish a "manual ground truth." The mean measurements from the three observers were used as the reference standard. Three examiners assessed the interexaminer reliability for all tracings, and steps were taken to restrict daily cephalogram tracings to five to mitigate errors caused by fatigue. Measurements were taken for a total of 5 skeletal, 4 dental, and 1 soft tissue parameters, which included 4 linear and 6 angular measurements as mentioned in table 1. To improve landmark identification, a sole operator (AK) performed both manual and digital tracings, while two extra manual operators were enlisted to establish a "manual ground truth," using the average measurements from the three observers as the reference standard.

Table 1: Description of the cephalometric measurements

SKELETAL PARAMETERS	
SNA	The angle between Sella, Nasion, and point A

SNB	The angle between Sella, Nasion, and point B
ANB	Difference between SNA and SNB angle.
FMA	The angle formed by Frankfort's horizontal plane and mandibular plane.
IMPA	The angle formed by the mandibular plane angle and the long axis of the mandibular incisor.
DENTAL PARAMETERS	
1- Pt A	Maxillary central incisor to Point A
1-A pog	Mandibular Central incisor to Apo Line
1-N-A	The linear distance between the anteriormost point of the maxillary central incisor and the NA line
1- N-B	The linear distance between the anteriormost point of the mandibular central incisor and the NB line
SOFT TISSUE PARAMETER	
Nasolabial Angle	The angle formed by a line tangent to the base of the nose and a line tangent to the upper lip

Manual Tracing

High-quality digital cephalograms, obtained in JPG format, were printed from the radiographs. Manual tracings were performed in a dimly lit room on an illuminated view box. These tracings were manually drawn using a 0.5 mm lead pencil on fine grade 36 m matte acetate tracing paper, firmly affixed over the superior X-ray printout. Linear and angular measurements corresponding to Steiner's, McNamara's, and Tweed's cephalometric analyses were then recorded.

Digital manual Tracing

To conduct digital manual cephalometric evaluations, digital images of the selected cephalograms in JPG format were imported into FACAD, Dolphin, and NemoCeph digital software. After calibrating the images using a reference scale, the operator identified skeletal, dental, and soft tissue landmarks on the digital images. Upon completing landmark identification, linear and angular measurements were obtained from the software. All recorded cephalometric measurements were documented in an Excel spreadsheet.

Web-based fully automated tracing

To conduct AI-generated cephalometric evaluations, digital images of the lateral cephalograms in JPG format were imported into Webceph software. Once the images are calibrated using a reference scale, the software automatically detects the skeletal, dental, and soft tissue landmarks.

Statistical analysis

Statistical analysis was carried out using SPSS software (IBM SPSS Statistics, Version 20.0, Armonk, NY: IBM Corp.). For each tracing system, the mean, minimum, maximum and standard deviation of all measurements were computed. Inter-group comparisons were conducted through a one-way ANOVA test, followed by post hoc testing using the Tukey test. Intra-class and inter-class variations were examined using intra-class correlation coefficients (ICC) with a 95% confidence interval.

3. Result and Discussion

The ICC values across various cephalometric analysis software demonstrated good reproducibility for skeletal, dental, and soft tissue measurements. For SNA, SNB, ANB, FMA, IMPA, 1 to Pt A, 1 to N-A and nasolabial angle, the p-values indicated no significant differences between the software. In contrast, 1 to A Pog and 1 to N-B revealed significant differences between the software as shown in Table 2.

Table 2: Comparison of mean and standard deviation of different cephalometric measurements between various cephalometric analysis software

Landmark	Manual tracing	FACAD	Dolphin Imaging	NemoCeph	WebCeph	One-way ANOVA Test Value	P Value
SNA	81.70 ± 1.46	81.70 ± 1.35	81.97 ± 1.59	82.23 ± 1.16	82.23 ± 1.43	1.30	0.27
SNB	80.07 ± 1.13	80.07 ± 1.17	80.37 ± 1.14	80.61 ± 1.26	80.59 ± 1.48	1.96	0.09
ANB	1.63 ± 0.6	1.63 ± 0.57	1.61 ± 0.44	1.62 ± 0.71	1.64 ± 0.9	0.81	0.52
FMA	25.33 ± 1.68	25.65 ± 1.29	25.67 ± 1.37	25.59 ± 1.78	25.13 ± 1.86	0.65	0.63
IMPA	96.33 ± 1.68	96.53 ± 1.58	96.85 ± 1.34	96.73 ± 1.84	96.26 ± 1.92	0.66	0.62
1 to Pt A	3.7 ± 0.53	3.69 ± 0.46	3.75 ± 0.53	3.57 ± 0.5	3.75 ± 0.49	0.64	0.63
1 to A Pog	1.90 ± 0.42	1.79 ± 0.39	1.93 ± 0.29	2.09 ± 0.40	1.83 ± 0.44	2.65	0.03*
1 to N- A	4.20 ± 0.52	4.19 ± 0.45	4.23 ± 0.52	4.07 ± 0.50	4.25 ± 0.49	0.61	0.65
1 to N- B	2.47 ± 0.35	2.29 ± 0.39	2.43 ± 0.29	2.59 ± 0.4	2.33 ± 0.44	3.00	0.02*
Nasolabial angle	101 ± 3.1	101.5 ± 2.9	100.9 ± 3.1	101.2 ± 2.87	101.3 ± 2.89	0.16	0.96

The post hoc Tukey test concluded that FACAD vs. NemoCeph showed a significant difference with a mean difference of -0.30 and -0.30 and a p-value of 0.03 and 0.02 for 1 to A Pog and 1 to N-B values, respectively (Table 3). These differences indicate variability between the methods, with FACAD and NemoCeph

Table 3: Multiple comparison of mean differences of different cephalometric measurements between various cephalometric analysis softwares

Landmark	Groups	Manual Tracing	FACAD	Dolphin Imaging	NemoCeph	WebCeph	Mean Difference
1 to A Pog line	Manual tracing	-	0.11	-0.03	-0.19	0.07	Mean difference
		-	0.83	1.00	0.32	0.95	P value
	FACAD	-0.11	-	-0.13	-0.30	-0.03	Mean difference
		0.83	-	0.68	0.03*	1.00	P value

	Dolphin Imaging	0.03	0.13	-	-0.17	0.10	Mean difference
		1.00	0.68	-	0.47	0.86	P value
	NemoCeph	0.19	0.30	0.17	-	0.27	Mean difference
		0.32	0.03*	0.47	-	0.07	P value
	WebCeph	-0.07	0.03	-0.10	-0.27	-	Mean difference
		0.95	1.00	0.86	0.07	-	P value
1 to NB	Manual tracing	-	0.17	0.04	-0.13	0.14	Mean difference
		-	0.39	0.99	0.69	0.60	P value
	FACAD	-0.17	-	-0.13	-0.30	-0.03	Mean difference
		0.39	-	0.65	0.02*	1.00	P value
	Dolphin Imaging	-0.04	0.13	-	-0.17	0.10	Mean difference
		0.99	0.65	-	0.43	0.84	P value
	NemoCeph	0.13	0.30	0.17	-	0.27	Mean difference
		0.69	0.02*	0.43	-	0.05	P value
	WebCeph	-0.14	0.03	-0.10	-0.27	-	Mean difference
		0.60	1.00	0.84	0.05	-	P value

Discussion

The precision of cephalometric analysis is crucial for accurately diagnosing malocclusion and developing effective treatment plans. Due to the quick advancements in computer technology, digital software is being employed to obtain cephalometric values. In addition to this software, cephalometric analysis is accessible through web or smartphone applications that allow for automated tracing. [27] Accuracy and a high rate of reproducibility are the most crucial requirements, regardless of the technique employed. Therefore, the present study evaluated the reliability of digital manual and artificial intelligence-based software when compared to manual tracing in order to aid the clinician in obtaining an accurate cephalometric measurement faster.

Erkan et al. in 2011 evaluated the reliability of Dolphin Imaging and NemoCeph when compared to manual tracing by measuring SNA, SNB, ANB, and 1-NB parameters. [20] No significant difference was noted between the software; however, the least difference was noted with Dolphin Imaging software. Hence stating these cephalometric analysis softwares were reliable. Abreu et al. in 2016 compared the reliability of Dolphin software with manual tracing. SNA, SNB and 1-NB values showed a significant difference, indicating Dolphin software was least effective and produced systematic errors. [10] Kublashivili et al. noted similar results where Dolphin software showed a significant difference in the soft tissue parameters when compared to manual tracing and other software but not in hard tissue parameters. [28,29] Paixão et al in 2010 and Meriç et al in 2020 noted no significant difference in any parameters measured between Dolphin imaging software and manual tracing. [24,30] The present study noted good reliability between Dolphin Imaging software and manual tracing.

Farooq et al. in 2016 assessed the reliability of FACAD software in comparison to the manual tracing method. [31] A high corelation was noted in all the parameters except 1-NA, interincisal angle, and y-axis, which was attributed to difficulty in identification of the anatomical landmarks. The present study noted a significant difference in 1-NB value between FACAD and NemoCeph software, though the values were clinically insignificant. It may be concluded that the apparent lack of accuracy of measurements may not be related to the software used but rather to the likely uncertainty at the landmark identification since it is semi-digitized software. Similar to the present study, Sangeetha et al evaluated the accuracy between NemoCeph software and manual tracing and noted a good agreement between the values measured by both methods, suggesting NemoCeph is a reliable method for obtaining the cephalometric values. [32]

Çoban et al. compared the cephalometric values between digital manual (Dolphin Imaging software) and artificial intelligence-based WebCeph software. [33] The results indicated that most measurements showed significant differences between the two methods ($p<0.05$), except for 7 parameters. The authors noted that while some measurements showed significant differences, not all were clinically significant. It was suggested that the AI-based method should be further developed to better address specific malocclusions. Overall, the AI-based WebCeph platform provided comparable results to digital manual tracing for certain parameters, but more refinement is needed to improve accuracy across different malocclusion types. The present study noted good reliability between WebCeph software and manual tracing. Hence, artificial intelligence-based and digital manual software are reliable for obtaining cephalometric values

4. Conclusion and future scope

Digital manual as well as artificial intelligence-based software presented to be accurate when compared to manual tracing. Dolphin Imaging software emerged to be the most reliable software, followed by WebCeph and NemoCeph, and the most variability was noted in FACAD software. No clinical significance was noted, and the variability was due to uncertainty in certain landmark identification.

Reference

[1] Gateno J, Xia JJ, Teichgraeber JF. New 3-dimensional cephalometric analysis for orthognathic surgery. J Oral Maxillofac Surg. 2011 Mar;69(3):606–22.

[2] Baxi S, Shadani K, Kesri R, Ukey A, Joshi C, Hardiya H. Recent Advanced Diagnostic Aids in Orthodontics. Cureus. 2022 Nov;14(11):e31921.

[3] Kapila SD, Nervina JM. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. Dentomaxillofac Radiol. 2015;44(1):20140282.

[4] Swennen GRJ, Schutyser F, Hausamen JE. Three-Dimensional Cephalometry: A Color Atlas and Manual. 2006th ed. Swennen GRJ, Schutyser FAC, Hausamen JE, editors. Berlin, Germany: Springer; 2005. 366 p.

[5] Kolokitha OE, Topouzelis N. Cephalometric methods of prediction in orthognathic surgery. J Maxillofac Oral Surg. 2011 Sep;10(3):236–45.

[6] Jain RK, Sowmithradevi S, Shantha SKK. Correlation of forehead type with maxillary incisor inclination in Dravidian south Indian population: A prospective study. World J Dent. 2022 Aug 26;13(6):606–10.

[7] Prasanna Arvind T.R, Sumathi Felicita A. Correlation between collum angle and lower lip position in different Class II malocclusions - A retrospective cephalometric study. Orthod Waves [Internet]. 2021 Apr 3 [cited 2024 Sep 27]; Available from: https://www.tandfonline.com/doi/abs/10.1080/13440241.2021.1924417

[8] 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 Apr 3;28(1):e2321214.

[9] Hlongwa P. Cephalometric analysis: manual tracing of a lateral cephalogram. SADJ [Internet]. 2019;74(6). Available from: http://ref.scielo.org/fc245s

[10] Paini de Abreu D, Salvatore Freitas KM, Nomura S, Pinelli Valarelli F, Hermont Cançado R. Comparison among manual and computerized cephalometrics using the softwares dolphin imaging and dentofacial planner. Dent Oral Craniofac Res [Internet]. 2016;2(6). Available from: https://oatext.com/Comparison-among-manual-and-computerized-cephalometrics-using-the-softwares-dolphin-imaging-and-dentofacial-planner.php

[11] Agrawal MS, Manish Agrawal JA, 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 Apr 27;5(103):103–10.

[12] Albarakati SF, Kula KS, Ghoneima AA. The reliability and reproducibility of cephalometric measurements: a comparison of conventional and digital methods. Dentomaxillofac Radiol. 2012 Jan;41(1):11–7.

[13] Liu J, Zhang C, Shan Z. Application of Artificial Intelligence in Orthodontics: Current State and Future Perspectives. Healthcare (Basel) [Internet]. 2023 Oct 18;11(20). Available from: http://dx.doi.org/10.3390/healthcare11202760

[14] Subramanian AK, Chen Y, Almalki A, Sivamurthy G, Kafle D. Cephalometric Analysis in Orthodontics Using Artificial Intelligence-A Comprehensive Review. Biomed Res Int. 2022 Jun 16;2022:1880113.

[15] Vaid NR, Adel SM. Contemporary Orthodontic Workflows: A Panacea for Efficiency? Semin Orthod. 2023 Mar 1;29(1):1–3.

[16] Katyal V, Vaid N. Virtual-First: A virtual workflow for new patient consultation, engagement and education in orthodontics. Semin Orthod. 2023 Mar 1;29(1):109–15.

[17] Pandian SM, Gandedkar NH, Palani SK, Kim YJ, Adel SM. An integrated 3D-driven protocol for surgery first orthognathic approach (SFOA) using virtual surgical planning (VSP). Semin Orthod. 2022 Dec 1;28(4):320–33.

[18] Pandian SM, Subramanian AK, Ravikumar PA, Adel SM. Biomaterial Testing in Contemporary Orthodontics: Scope, Protocol and Testing Apparatus. Semin Orthod. 2023 Mar 1;29(1):101–8.

[19] Livas C, Delli K, Spijkervet FKL, Vissink A, Dijkstra PU. Concurrent validity and reliability of cephalometric analysis using smartphone apps and computer software. Angle Orthod. 2019 Nov;89(6):889–96.

[20] Erkan M, Gurel HG, Nur M, Demirel B. Reliability of four different computerized cephalometric analysis programs. Eur J Orthod. 2011 Apr 18;34(3):318–21.

[21] Tsolakis IA, Tsolakis AI, Elshebiny T, Matthaios S, Palomo JM. Comparing a Fully Automated Cephalometric Tracing Method to a Manual Tracing Method for Orthodontic Diagnosis. J Clin Med Res [Internet]. 2022 Nov 20;11(22). Available from: http://dx.doi.org/10.3390/jcm11226854

[22] The accuracy and reliability of WebCeph for cephalometric analysis. Journal of Taibah University Medical Sciences. 2022 Feb 1;17(1):57–66.

[23] Naoumova J, Lindman R. A comparison of manual traced images and corresponding scanned radiographs digitally traced. Eur J Orthod. 2009 Jun;31(3):247–53.

[24] Paixão MB, Sobral MC, Vogel CJ, Araujo TM de. Estudo comparativo entre traçados cefalométricos manual e digital, através do programa Dolphin Imaging em telerradiografias laterais. Dental Press J Orthod. 2010 Dec;15(6):123–30.

[25] Mahto RK, Kharbanda OP, Duggal R, Sardana HK. A comparison of cephalometric measurements obtained from two computerized cephalometric softwares with manual tracings. J Indian Orthod Soc. 2016 Jul;50(3):162–70.

[26] Kunz F, Stellzig-Eisenhauer A, Boldt J. Applications of Artificial Intelligence in Orthodontics—An Overview and Perspective Based on the Current State of the Art. NATO Adv Sci Inst Ser E Appl Sci. 2023 Mar 17;13(6):3850.

[27] Adel SM, Alwafi AA, Pandian SM, Bichu YM, Abuljadayel LW, Alansari RA, et al. What are orthodontic residents perusing on social media? A cross-sectional survey. Semin Orthod. 2023 Dec 1;29(4):382–9.

[28] 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 Sep;10(3):212–9.

[29] Gregston MD, Kula T, Hardman P, Glaros A, Kula K. A comparison of conventional and digital radiographic methods and cephalometric analysis software: I. hard tissue. Semin Orthod. 2004 Sep;10(3):204–11.

[30] Meriç P, Naoumova J. Web-based Fully Automated Cephalometric Analysis: Comparisons between App-aided, Computerized, and Manual Tracings. Turk J Orthod. 2020 Sep;33(3):142–9.

[31] Farooq MU, Khan MA, Imran S, Sameera A, Qureshi A, Ahmed SA, et al. Assessing the Reliability of Digitalized Cephalometric Analysis in Comparison with Manual Cephalometric Analysis. J Clin Diagn Res. 2016 Oct;10(10):ZC20–3.

[32] 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 Aug 26;13(6):630–40.

[33] 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 Aug 15;27(4):e222112.