



Original Article

## Evaluation of the accuracy and reliability of WebCeph – An artificial intelligence-based online software

Deepika Katyal<sup>1</sup>, Nivethigaa Balakrishnan<sup>1</sup>

<sup>1</sup>Department of Orthodontics, Saveetha Dental College, Chennai, Tamil Nadu, India.



**\*Corresponding author:**

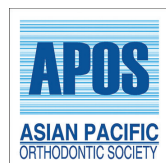
Dr. Deepika Katyal,  
Resident, Department of  
Orthodontics, Saveetha Dental  
College and Hospital, Chennai,  
Tamil Nadu, India.

dr.deepikakat96@gmail.com

Received: 27 September 2021  
Accepted: 13 February 2022  
Epub Ahead of Print: 09 June 2022  
Published: 29 November 2022

DOI  
10.25259/APOS\_138\_2021

**Quick Response Code:**



### ABSTRACT

**Objectives:** Landmark identification is of utmost importance in cephalometric analysis but it turns out to be the main source of error. With modern inventions in the field of artificial intelligence (AI), it becomes essential to assess the reliability of computer-automated programs. A greater deal of time can be conserved with fully automated programs such as WebCeph, which uses an AI-based algorithm that performs automated and immediate cephalometric analysis. This study aimed to evaluate the accuracy, reliability, and duration of tracing cephalometric radiographs with WebCeph, an AI-based software in comparison to digital tracing with FACAD and manual tracing. The null hypothesis proposed is that there is no statistically significant difference among the three methods with regard to accuracy of cephalometric analysis.

**Material and Methods:** Pre-treatment cephalometric radiographs of 25 patients (14 males and 11 females, mean age of  $18 \pm 3.2$  years) were selected randomly from the dental information archiving software of Saveetha University, Department of Orthodontics, Chennai. Composite analysis with skeletal, dental and soft-tissue parameters was selected and cephalometric analysis was done with all three methods – Manual tracing (Group 1), digital tracing using FACAD (Group 2), and fully automated AI-based software WebCeph (Group 3). The timing for each method of analysis was calculated using a stopwatch in seconds. Values were tabulated in an Excel sheet and statistical analysis including one-way analysis of variance and *post hoc* Tukey test were performed.

**Results:** No statistically significant difference was found between the three methods for cephalometric analysis,  $P > 0.05$ . The time taken for measurement using the three different methods was the least while using WebCeph ( $30.2 \pm 6.4$  s) and the maximum while manual tracing ( $472 \pm 40.4$  s).

**Conclusion:** WebCeph is a reliable, faster and practical tool for analyzing cephalometric analysis in comparison to digital tracing using FACAD and manual tracing.

**Keywords:** Cephalometry, Artificial intelligence, WebCeph, Orthodontics

### INTRODUCTION

Cephalometry has been instrumental in orthodontic diagnosis, treatment planning, and craniofacial growth prediction.<sup>[1]</sup> Manual tracing of lateral cephalograms has been in practice for many years. Angular and linear measurements on the lateral cephalogram are carried out with an acetate tracing paper, scale, and protractor. However, manual tracing comes with its disadvantages. It is time-consuming, prone to errors and has a risk of misreading values due to faulty landmark identification or radiographic magnification.<sup>[1-4]</sup> All these drawbacks lead to the advancement of digital and computerized cephalometry in recent times which have now

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 License, which allows others to remix, transform, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

©2022 Published by Scientific Scholar on behalf of APOS Trends in Orthodontics

replaced manual cephalometry with rapid advancement in technology.<sup>[5]</sup>

Digital cephalometric analysis has numerous advantages such as facilitated image acquisition, faster measurements, sharing and archiving, faster treatment planning and reduced chemically associated hazards. Furthermore, several analyses can be performed at once, with superimposition of serial radiographs possible at a faster rate.<sup>[5-7]</sup>

Around 350 specially designed orthodontic applications exist as of today indicating the boom in software development and technology.<sup>[8]</sup> Smartphones are a useful entity for digital analysis and treatment planning; however, there is a lack of a standardized method of evaluation of the reliability and accuracy of mobile phone applications for cephalometric analysis.<sup>[9]</sup>

The common factor among all the smartphone applications, computerized, and digital software that exist for cephalometric tracing is that the anatomical landmarks need to be digitally located by the orthodontist. This aspect makes these applications only semi-automated. Since landmark identification is of utmost importance in cephalometry and also as it turns out, is the main source of error, it is important to assess the reliability of recently developed computer automated programs.<sup>[10,11]</sup> A great deal of time can be saved with fully automated programs such as WebCeph, which uses artificial intelligence (AI)-based algorithm that performs automated and immediate cephalometric analysis.

With the increasing need for faster and more accurate digital cephalometric software, comparative studies are required to help physicians make an informed choice regarding the accuracy and reliability of AI-based software.<sup>[10,11]</sup>

There is no published literature comparing the accuracy and reliability of WebCeph – a fully automated AI-based software, with FACAD which is a semi-automated digital cephalometric software and manual tracing. Therefore, this study aimed to evaluate the accuracy and reliability of WebCeph in comparison to FACAD, taking manual tracing as a gold standard for comparison. Furthermore, the time taken for analysis using each method was also calculated. The null hypothesis proposed is that there is no statistically significant difference among the three methods concerning the accuracy of cephalometric analysis.

## MATERIAL AND METHODS

The study was conducted in a University setting with approval from the Ethical Committee at Saveetha University, Chennai. Informed consent was taken from all the participants before their participation. IRB approval number - IHEC/SDC/ORTHO-2005/22/389.

The sample size calculation was done with a significance level of 0.05 and a power value of 95%. A sample of a minimum

of 25 patients was needed. The effect size was based on a previous study.<sup>[12]</sup>

Pre-treatment cephalometric radiographs of 25 patients (14 males and 11 females, mean age  $18 \pm 3.2$  years) were selected randomly from the dental imaging and archiving software of Saveetha University, Department of Orthodontics, Chennai. The lateral cephalograms of all patients were taken with the patient in an upright standing position with the Frankfurt plane parallel to the floor, keeping the teeth in centric relation and lips relaxed. All lateral cephalograms were taken by the same cephalometric radiography machine by the same operator.

The criteria for including the radiographs in the study were that the radiographs should be of good quality for non-growing individuals with a permanent set of dentition. The exclusion criteria involved poor quality or distorted radiographs with artifacts that could prevent anatomical landmark identification, unerupted or missing teeth, dental deformities that could prevent incisor apex identification and gross skeletal deformities.

To reduce errors in landmark identification, the same operator took all the digital and manual tracings. However, two more observers were included in the study to reperform the manual tracings at a different time interval and the mean measurements of the manual tracings performed by three observers were taken as the “manual ground truth.”

Only five manual tracings were performed by an operator at a given time interval to avoid errors due to operator fatigue.

The same 10 angulars, 10 linear, and two soft-tissue parameters were measured on each radiograph.

To determine the intraoperator error, five radiographs out of 25 were randomly selected and retraced digitally by the same operator after 1 month. For intraoperator error “manual truth,” five radiographs were traced by the three observers and the mean formed the “manual truth” value.

The timing for each method of analysis was calculated using a stopwatch in seconds.

The start- and end-points for the manual cephalometric measurements included plotting the landmarks and measuring the angles and distances. The manual measurements were made by three operators and the mean analyzing time was calculated. Analyzing time for computerized and app-aided tracing included plotting of the landmarks by one operator as measurements of angles and distances were performed by the software. For the web-based fully automated tracing, the analyzing time was the time it took for the system to automatically identify the anatomical points. Manual correction of the landmark positions was also made, which was added to the total analyzing time.

Calibration of the images for all the systems was also included in the analyzing time.

### Manual tracing

For manual tracing, digital images were imported to Adobe Photoshop 7.0 (Adobe Systems, San Jose, California, USA) and resized to scale 1:1, and were printed. Using the rectangular marquee tool, a distance of 10 mm was measured on the vertical calibration ruler on the cephalogram. The selected area was copied and pasted into a new file. The number of vertical pixels of the created file was noted. After returning to the original file, the image menu—image size tab was entered. Resample image box was unchecked, the number of vertical pixels recorded from the previous image was written in the resolution box (pixels/cm), and the image was scaled. The image properties of the film were  $2.232 \times 2.304$  pixels, 150 dpi, and 8 bits. Manual tracing was performed on the printed image using a 0.35 mm lead pencil. All the hard tissue and soft-tissue landmarks were traced, and double images were centered to form a single landmark. A ruler and protractor were used to measure the angular and linear parameters.

### Computerized tracing

For the computerized tracing method, digital radiographs saved as .jpeg files were imported to the FACAD Imaging. The files were in grayscale format, and the image properties of the film were  $2.232 \times 2.304$  pixels, 150 dpi, and 8 bits. The digital films were calibrated by digitizing 2 points (10 mm) on the ruler within the digital cassette. Landmark identification was carried out manually using a laptop-driven cursor. The screen used for computerized analysis was 14" in size. All measurements were performed automatically by the software [Figure 1].

### Web-based fully automated tracing

An online automatic cephalometric tracing and analysis service named WebCeph was used. After entering the system with [www.webceph.com](http://www.webceph.com), using a standard web browser (Google Chrome 64 bit), a new patient was created and a "jpeg" formatted cephalometric X-ray image was uploaded. The files were in grayscale format and the image properties of the film were  $2.232 \times 2.304$  pixels, 150 dpi, and 8 bits. Once the images were uploaded, the system automatically identified all the anatomical points. The screen used for the analysis was 14" in size. Calibration was set to 10 mm and the analysis was downloaded to the computer without any correction. The same set of data, after the automatic tracing, was also manually corrected for landmark position and downloaded to the computer [Figure 2].



Figure 1: Semi-automated cephalometric tracing using 'FACAD' software.

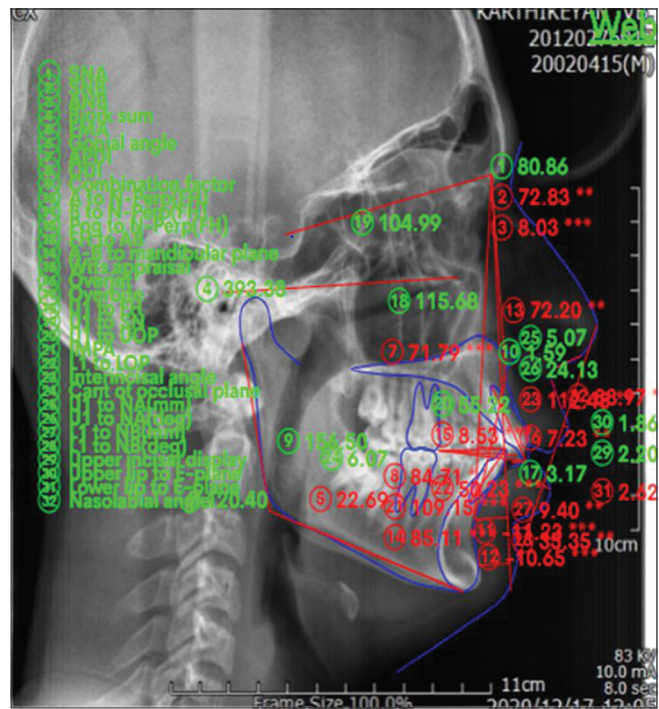


Figure 2: Cephalometric analysis using 'WebCeph'—An AI based software.

### Statistical analysis

Statistical analysis was conducted using the Statistical Package for the Social Sciences version 23.0 software (IBM Corp.; Armonk, NY, USA). The mean, minimum, maximum, and SD of all the measurements were calculated for each tracing system. Intergroup comparisons were made



with a one-way analysis of variance with a significance level of 0.05.

**RESULTS**

Regarding the SNA, SNB and ANB angles comparatively higher values were found with WebCeph software compared to FACAD and manual methods. On the contrary, the mean values of gonial angle and Frankfurt mandibular plane angle were higher with FACAD when compared to WebCeph and manual methods. However, statistically no significant difference was found between the skeletal measurements between the three methods of measurement ( $P > 0.05$ ) [Table 1].

Similarly, with the dental measurements, no significant difference was found statistically between the three methods of measurements ( $P > 0.05$ ) [Table 1].

The soft-tissue parameters which were compared using the three different methods also showed no significant difference statistically ( $P > 0.05$ ) [Table 1].

The time taken for measurement using the three different methods was the least while using WebCeph ( $30.2 \pm 6.4$  s) and maximum while manual tracing ( $472 \pm 40.4$  s) [Table 2].

The ICC values for interexaminer reliability for manual tracings were above 0.90 for all the measurements indicating very high interexaminer reliability between the three operators [Table 3].

**DISCUSSION**

Digital methods of cephalometry are fast gaining popularity. However, the accuracy and reproducibility of the results are the important factors which need to be considered before adapting to any one digital method of cephalometry. Irrespective of whether the application is semi-automated, mobile phone based or AI based, the tracing should be accurate and highly reproducible. The principal finding of this study was that WebCeph, an AI-based software, is as reliable and accurate for cephalometric analysis as manual tracing.

The previous literature has indicated that a difference of lesser than 2 mm or degrees to be clinically insignificant.<sup>[13]</sup> Therefore, all measurements taken in this study which had no statistical difference are also clinically relevant. Therefore, the null hypothesis is accepted, that is, there is no statistical difference between the three methods with regard to cephalometric accuracy.

The results of this study are in concordance with the results obtained by Alqahtani,<sup>[12]</sup> wherein the accuracy and reliability of cephalometric measurements of CephX – an online-based platform were assessed in comparison to FACAD. Alqahtani *et al.* found no statistically significant differences in the angular and linear measurements except for SNA, FMA, and Pg to B values between the two methods. However, overall, no clinically significant difference was found and both the methods were found to be comparable for cephalometric measurements.

**Table 1:** Mean and standard deviations of skeletal, dental, and soft-tissue cephalometric measurements done by the three methods with P values. The three methods reported a non-significant difference about all the measurements ( $P > 0.05$ ).

Measurement	WebCeph		FACAD		Manual ground truth		P-value
	Mean	S.D	Mean	S.D	Mean	S.D	
SNA angle	85.2	3.5	83.1	6.15997	83.8222	5.62000	0.693
SNB angle	80.4	3.1	79.0	4.09312	80.0444	4.78856	0.780
ANB angle	5.9	2.8	4.1	2.01232	5.9111	3.58519	0.477
BJORK sum	389.7	3.0	376.9	40.3	383.7	16.5	0.570
Frankfurt mandibular plane angle	23.2	2.8	28.1	3.5	25.4	1.6	0.430
Gonial angle	119.4	3.8	125.8	2.8	123.0	2.5	0.534
A to N perpendicular	2.3	2.8	-0.8922	4.8	1.7	3.7	0.196
POG to N perpendicular	-6	5.3	-5.5	6.9	-5.5	5.1	0.990
Wits appraisal	3.9	3.7	1.6	3.3	3.4	4.7	0.471
Overjet	2.5	0.845	6.5	2.7	6.7	2.4	0.959
Overbite	2.4	2.0	3.2	3.5	3.4	1.8	0.689
U1 to SN	112.0	6.9	117.6	6.8	116.0	7.4	0.237
Incisor mandibular plane angle	102.2	4.7	103.7	7.4	105.6	5.9	0.515
Interincisal angle	116.0	7.3	109.4	8.1	111.2	7.0	0.182
U1 to NA in mm	6.2	2.2	9.1	2.2	6.6	3.5	0.083
U1 to NA in degree	26.2	5.2	32.3	7.8	30.4	7.6	0.189
L1 to NB in mm	8.3	2.7	8.2	2.7	7.3	2.7	0.668
L1 to NB in degree	31.8	5.1	33.9	8.0	35.4	6.3	0.517
Upper incisor exposure	3.2	1.1	3.9	2.3	4.1	1.1	0.430
Lower lip to E angle	1.6	3.6	0.80	3.6	1.0	5.0	0.894
Nasolabial angle	88.3	4.4	95.6	25.7	89.3	12.08	0.611

**Table 2:** The mean and standard deviation of the time taken in seconds for cephalometric analysis for the three methods.

Duration	WebCeph	FACAD	Manual
	30.2±6.4 s	115±32.5 s	472±40.4 s

**Table 3:** The intraclass coefficient for the interexaminer reliability of manual tracings.

Measurements	Manual methods – ICC value
SNA	0.940
SNB	0.932
ANB	0.930
Frankfurt mandibular plane angle	0.952
Gonial angle	0.939
A to N perpendicular	0.900
POG to N perpendicular	0.980
Wits appraisal	0.912
Overjet	0.928
Overbite	0.931
U1 to SN	0.921
Incisor mandibular plane angle	0.945
Interincisal angle	0.911
U1 to NA in mm	0.900
U1 to NA in degrees	0.983
LI to NB in mm	0.977
LI to NB in degree	0.988
Upper incisor exposure	0.912
Lower lip to e angle	0.945
Nasolabial angle	0.913

The leap in technology has resulted in the invention of applications which can be used on smartphones and tablet PC's.<sup>[14]</sup> Studies have been done in the past wherein the accuracy of smartphone applications was assessed in comparison to manual tracing. A study by Sayar and Kilinc was done to evaluate the accuracy of CephNinja mobile phone application with manual tracing.<sup>[15]</sup> They found no statistically significant difference in eight out of the 12 measurements and concluded that CephNinja was comparable and a more faster method than manual tracing. A similar study was done by Aksakalli *et al.* where iPad applications smart Ceph Pro and CephNinja were tested for reliability against the measurements obtained by Dolphin software.<sup>[16]</sup> The study revealed that the two applications were better at angular measurements in comparison to linear measurements. The study concluded that the applications were not as good as the Dolphin software and needed to be developed more to be comparable.

Fully automated cephalometric programs using AI are rapidly becoming popular as one of the most important criteria and source of error in cephalometrics in landmark identification.<sup>[11]</sup> AI allows landmarks to be detected automatically instead of manually having to locate landmarks thus reducing the probability of error.

Only few studies have been done in the past evaluating the accuracy of AI-based applications. One such study was done by Meriç and Naoumova wherein a fully automated, AI-based program, CephX was evaluated for accuracy and reliability.<sup>[17]</sup> However, the study concluded that the application needed improvements to be comparable to the other two methods assessed. CephX, however, was comparable to Dolphin and CephNinja softwares if the landmarks were manually corrected. It also had a faster analyzing time compared to the other two applications.

Ours is the only study evaluating the accuracy and reliability of WebCeph – a fully automated, AI-based application in comparison to that of FACAD software, taking manual tracing as a gold standard.

All lateral cephalograms incorporated in this study were obtained directly in a digital format. This step eliminated the demand for scanning conventional radiographic films. Scanning of lateral cephalograms is a time-consuming process, prone to poor quality and may cause magnification errors.<sup>[18]</sup> AI-based fully automated cephalometry is more precise because once the images are detected on-screen, measurements and data processing occur automatically. In contrast, the conventional method requires rulers and protractors to accomplish the same job.<sup>[13]</sup> Semi-automated methods such as FACAD are prone to some errors. This can be caused by analog cephalometric radiographs with improper quality that displays poorly on a computer screen and results in lower quality images.<sup>[19,20]</sup> The number of pixels (gray scale) is affected by the compression technique often found in digital images. In this study, the jpeg format with standard compression settings was used, which does not affect the image diagnostic quality.

Almost all measurements showed moderate to almost perfect agreement between the three methods of measurement. Therefore, this study confirms that WebCeph – fully automated program can perform cephalometric analysis with comparable accuracy and reproducibility as manual tracing and in a shorter period of time.

## CONCLUSION

The cephalometric measurements obtained from both WebCeph and FACAD are highly reliable and accurate. The advantages of an online-based, AI-based software include Cloud-based storage, online archiving, quick analysis, no need for specific installation or software, and compatibility with any operating system. All these factors make WebCeph a reliable, faster, and practical tool for cephalometric analysis.

## Declaration of patient consent

The authors certify that they have obtained all appropriate patient consent.

## Financial support and sponsorship

Nil.

## Conflicts of interest

There are no conflicts of interest.

## REFERENCES

1. Arponen H, Elf H, Evalahti M, Waltimo-Siren J. Reliability of cranial base measurements on lateral skull radiographs. *Orthod Craniofac Res* 2008;11:201-10.
2. Cohen AM, Linney AD. A low cost system for computer based cephalometric analysis. *Br J Orthod* 1986;13:105-8.
3. Baumrind S, Frantz RC. The reliability of head film measurements. *Am J Orthod* 1971;60:111-27.
4. Baumrind S, Miller D, Molthen R. The reliability of head film measurements. *Am J Orthod* 1976;70:617-44.
5. Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop* 1998;113:173-9.
6. Naoumova J, Lindman R. A comparison of manual traced images and corresponding scanned radiographs digitally traced. *Eur J Orthod* 2009;31:247-53.
7. Huja SS, Grubaugh EL, Rummel AM, Fields HW, Beck FM. Comparison of hand-traced and computer-based cephalometric superimpositions. *Angle Orthod* 2009;79:428-35.
8. Daokar S, Phatak S. Orthodontic apps: A stairway to the future. *Int J Orthod Rehabil* 2019;10:75.
9. Prgomet M, Georgiou A, Westbrook JI. The impact of mobile handheld technology on hospital physicians' work practices and patient care: A systematic review. *J Am Med Inform Assoc* 2009;16:792-801.
10. Forsyth DB, Davis DN. Assessment of an automated cephalometric analysis system. *Eur J Orthod* 1996;18:471-8.
11. Sayinsu K, Isik F, Traklyali G, Arun T. An evaluation of the errors in cephalometric measurements on scanned cephalometric images and conventional tracings. *Eur J Orthod* 2007;29:105-8.
12. Alqahtani H. Evaluation of an online website-based platform for cephalometric analysis. *J Stomatol Oral Maxillofac Surg* 2020;121:53-7.
13. Chen YJ, Chen SK, Yao JC, Chang HF. The effects of differences in landmark identification on the cephalometric measurements in traditional versus digitized cephalometry. *Angle Orthod* 2004;74:155-61.
14. Singh P. Orthodontic apps for smartphones. *J Orthod* 2013;40:249-55.
15. Sayar G, Kilinc DD. Manual tracing versus smartphone application (app) tracing: A comparative study. *Acta Odontol Scand* 2017;75:588-94.
16. Aksakalli S, Yilanci H, Görükmez E, Ramoğlu Sİ. Reliability assessment of orthodontic apps for cephalometrics. *Turk J Orthod* 2016;29:98-102.
17. Meriç P, Naoumova J. Web-based fully automated cephalometric analysis: Comparisons between app-aided, computerized, and manual tracings. *Turk J Orthod* 2020;33:142-9.
18. Duterloo HS, Planche PG. *Handbook of Cephalometric Superimposition*. United States: Quintessence Publishing Company; 2011.
19. Oliver RG. Cephalometric analysis comparing five different methods. *Br J Orthod* 1991;18:277-83.
20. Macri V, Wenzel A. Reliability of landmark recording on film and digital lateral cephalograms. *Eur J Orthod* 1993;15:137-48.

**How to cite this article:** Katyal D, Balakrishnan N. Evaluation of the accuracy and reliability of WebCeph – An artificial intelligence-based online software. *APOS Trends Orthod* 2022;12:271-6.