TURKISH JOURNAL of ORTHODONTICS

DOI: 10.5152/TurkJOrthod.2020.20062

Original Article

Web-based Fully Automated Cephalometric Analysis: Comparisons between App-aided, Computerized, and Manual Tracings

Pamir Meriç[1], Julia Naoumova[2]

[1]Department of Orthodontics, Faculty of Dentistry, Trakya University, Edirne, Turkey
[2]Specialist Clinic of Orthodontics, Public Dental Service, Västra Götaland Region, Gothenburg, Sweden

Cite this article as: Meriç P, Naoumova J. Web-based Fully Automated Cephalometric Analysis: Comparisons between App-aided, Computerized, and Manual Tracings. Turk J Orthod 2020; 33(3): 142-9.

Main points:

- The fully automatic cephalometric analysis program CephX needs improvement to enhance its reliability for the majority of the dental measurements and for GoGn-SN (°).
- Manual correction of CephX gives similar results compared with CephNinja and Dolphin.
- CephX is significantly faster than CephNinja, Dolphin, and manual cephalometric analysis.

ABSTRACT

Objective: To compare the accuracy of cephalometric analyses made with fully automated tracings, computerized tracing, and app-aided tracings with equivalent hand-traced measurements, and to evaluate the tracing time for each cephalometric analysis method.

Methods: Pre-treatment lateral cephalometric radiographs of 40 patients were randomly selected. Eight angular and 4 linear parameters were measured by 1 operator using 3 methods: computerized tracing with software Dolphin Imaging 13.01(Dolphin Imaging and Management Solutions, Chatsworth, Calif, USA), app-aided tracing using the CephNinja 3.51 app (Cyncronus LLC, WA, USA) , and web-based fully automated tracing with CephX (ORCA Dental AI, Las Vegas, NV). Correction of CephX landmarks was also made. Manual tracings were performed by 3 operators. Remeasurement of 15 radiographs was carried out to determine the intra-examiner and inter-examiner (manual tracings) correlation coefficient (ICC). Inter-group comparisons were made with one-way analysis of variance. The Tukey test was used for post hoc testing.

Results: Overall, greater variability was found with CephX compared with the other methods. Differences in GoGn-SN (°), I-NA (°), I-NB (°), I-NA (mm), and I-NB (mm) were statistically ($p<0.05$) and clinically significant using CephX, whereas CephNinja and Dolphin were comparable to manual tracings. Correction of CephX landmarks gave similar results to CephNinja and Dolphin. All the ICCs exceeded 0.85, except for I-NA (°), I-NB (°), and I-NB (mm), which were traced with CephX. The shortest analyzing time was obtained with CephX.

Conclusion: Fully automatic analysis with CephX needs to be more reliable. However, CephX analysis with manual correction is promising for use in clinical practice because it is comparable to CephNinja and Dolphin, and the analyzing time is significantly shorter.

Keywords: Apps, artificial intelligence, automated identification, automatic tracing, cephalometric, computerized tracing, web-based

INTRODUCTION

Since its introduction in 1931, cephalometric analysis has become a widely used diagnostic and clinical tool in orthodontics (1). With the rapid advancement in technology, the manual method is gradually being replaced by digital cephalometric analysis software, which has numerous benefits, such as reduction in radiation doses, facilitated image acquisition, archiving and sharing, faster measurements, and easily determined treatment plans, as well as the elimination of chemical and associated environmental hazards. In addition, superimposition of serial radiographs can be performed faster, and it also allows the user to obtain several analyses at a time (2).

Address for Correspondence: Pamir Meriç, Department of Orthodontics, Faculty of Dentistry, Trakya University, Edirne, Turkey
E-mail: pamirmeric@trakya.edu.tr

Received: May 27, 2020
Accepted: July 16, 2020
Available Online Date: August 11, 2020

Specially designed medical and dental apps are one of the fastest growing categories of software, and as of today, more than 350 orthodontic apps exist, many of which can be accessed for free (3). Smartphones as an electronic training resource are useful for clinical decision support and to prevent medication errors (4); however, there is a lack of a systematic approach to evaluate the accuracy and evidence resulting from the use of mobile apps. There are contradictory findings regarding the validity of cephalometric analysis apps compared with the manual and digital tracing programs (5). The aspect that these digital tracing systems have in common, regardless of whether they are used on a tablet, smartphone, or a computer, is that the anatomical points need to be marked individually by the orthodontist during the tracing, making the cephalometric program only semiautomated. Since the main source of error in cephalometric analysis is landmark identification, it is important to assess whether the use of completely automated tracing programs, which have been developed lately, is reliable (6, 7). Computerized software and smartphone apps can save time compared with manual tracing (1, 8); however, physicians aim to use even lesser time for tracing. This might be possible with fully automated cephalometric analysis methods, such as web-based CephX, which is an artificial-intelligence (AI)–based algorithm that performs automatic, immediate cephalometric analyses (9). Automatic cephalometric analysis has been a topic of interest during the past years; however, the software algorithms developed did not seem accurate enough for clinical purposes. Whether a digital smartphone app or automatic analysis is selected, it should be precise, reliable, and highly reproducible. Given the increasing number of apps and computer-assisted cephalometric tracing programs and the lack of accuracy of the commercially available software, there is a need for comparative studies to allow the physicians to make an informed choice of suitable software and analysis methods (5, 10).

To our knowledge, there are no published data comparing all the 4 systems: fully automated, computerized, app-aided, and manual tracing. Therefore, this study aimed to compare the accuracy of manual cephalometric analyses to cephalometric analysis using AI, computerized method, and app-aided systems. In addition, the time required to perform the analysis using the 4 different methods was also assessed. The null hypothesis was that there are no statistical differences among the cephalometric analysis methods regarding the accuracy and the tracing time.

METHODS

The study was conducted according to the principles of the Declaration of Helsinki and was approved by the Research Ethics Committee at Trakya University, Faculty of Medicine (Approval No: TÜTF-BAEK 2017/318). Written informed consent was obtained from all the participants before their enrollment.

According to the power analysis, a minimum of 39 patients were needed to detect the correlations deviating from 0.5°/mm and above between the groups (with a significance level of 0.05 and power of 80%). The effect size was based on a previous study (11).

Pre-treatment lateral cephalometric radiographs of 40 patients (7 males, 33 females, mean age: 16.0 ± 4.6 years) were randomly selected from the archive of the Trakya University, Faculty of Dentistry, Department of Orthodontics. The cephalometric images were taken with the patient in the upright standing position with the Frankfort plane parallel to the floor, keeping the teeth in centric relation and the lips relaxed. All the lateral cephalometric radiographs were taken using the same cephalometric radiography machine (PaX-Flex; Vatech Inc. NJ) by the same technician with a magnification factor of 1.1.

The exclusion criteria were poor quality of cephalograms with artifacts that could interfere with the anatomical point identification, no unerupted or partially erupted teeth preventing incisor apex identification, and craniofacial deformities.

To optimize landmark identification, the same operator (PM) undertook all the digital and manual tracings, and to obtain a "manual ground truth," 2 additional manual operators were included; thus, a total of 3 observers performed the manual tracings. The mean measurements of the 3 observers represented the "manual ground truth."

No more than 5 tracings were made at a time to avoid operator fatigue. The same 8 angular and 4 linear parameters were measured on each radiograph (Figure 1, Table 1) except for the GoGn-SN value because CephX uses the GoMe plane to calculate the GoGn-SN angle.

To determine the intra-operator error, 15 radiographs were retraced digitally by the same operator (PM) after an interval of 1 month. For the intra-operator error of the "manual truth," 15

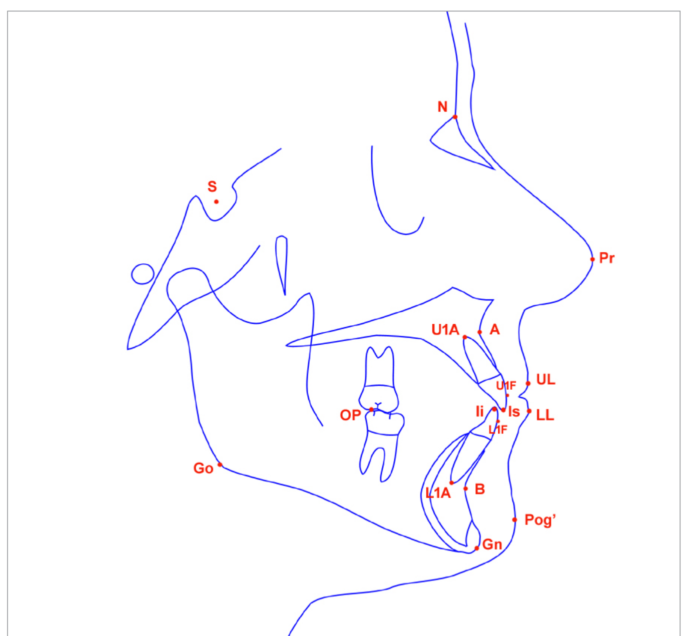

Figure 1. Cephalometric landmarks used in the study

S: Sella; N: Nasion; A: Point A; B: Point B; Go: Gonion; Gn: Gnathion; Pog: Soft tissue pogonion; Pr: Pronasale; UL: Upper lip; LL: lower lip; U1A: Upper incisor root apex; Is: Incision superior; Ii: Incision inferior; U1F: Upper incisor labial face; L1F: Lower incisor labial face

Turk J Orthod 2020; 33(3): 142-9

radiographs were retraced by the 3 observers and the mean formed the "manual truth" values.

Analyzing time for each analysis was measured in seconds using a stopwatch. The start- and end-points for the manual cephalometric measurements included plotting of the landmarks and measuring the angles and distances. The manual measurements were made by 3 operators, and the mean analyzing time was calculated. Analyzing time for computerized and app-aided tracing included plotting of the landmarks by 1 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.

Table 1. Description of the cephalometric landmarks and measurements used in the study

LANDMARK	DESCRIPTION
Sella (S)	The center of sella turcica
Nasion (N)	The most anterior point of the frontonasal suture
Point A (A)	The innermost point on the contour of the maxilla between the anterior nasal spine and the alveolar crest
Point B (B)	The most posterior point in the concavity along the anterior border of the symphysis
Gonion (Go)	The most prominent point on the angle of the mandible formed by the junction of the ramus and the body of the mandible
Gnathion (Gn)	Midpoint between menton and pogonion
Pog'	Soft tissue pogonion
Pronasale (Pr)	Tip of the nose
Upper lip (UL)	Most anterior point of the upper lip
Lower lip (LL)	Most anterior point of the lower lip
Incision superior incisal (Is)	The midpoint of the incisal edge of the most prominent maxillary central incisor
Incision inferior (Ii)	The midpoint of the incisal edge of the most prominent mandibular central incisor
U1F	Most anterior point of the maxillary central incisor
L1F	Most anterior point of the mandibular central incisor
SNA (°)	Angle determined by points S, N, and A
SNB (°)	Angle determined by points S, N, and B
ANB (°)	Angle determined by points A, N, and B
I-I (°)	Angle formed by the intersection of the upper incisor axis and the lower incisor axis
I-NA (°)	Angle formed by the intersection of the upper incisor axis and the NA line
I-NA (mm)	Linear distance between the most anterior point of the maxillary central incisor (U1F) and the NA line
I-NB (°)	Angle formed by the intersection of the lower incisor axis and the NB line
I-NB (mm)	Linear distance between the most anterior point of the mandibular central incisor (L1F) and the NB line
Occlusal Plane	The line joining the distal occlusal contact point of the first molars to midway of the anterior overbite
GoMe plane	A line between gonion and menton
OCC-SN (°)	Angle between the SN line and the occlusal plane
GoGn-SN (°)	Angle between the Go-Gn and SN lines
E-line	Esthetic line joining the soft tissue pogonion and pronasale
UL E-line (mm)	Linear distance between the most anterior point of the upper lip (UL) and the E-Line
LL E-line (mm)	Linear distance between the most anterior point of the lower lip (LL) and the E-Line

Manual Tracing

For manual tracing, digital images imported to Adobe Photoshop 7.0 (Adobe Systems, San Jose, California, USA) and resized to scale 1:1 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 amount 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 amount 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 2.232×2.304 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 Dolphin Imaging 13.01 software (Dolphin Imaging and Management Solutions, Chatsworth, Calif, USA). The files were in grayscale format, and the image properties of the film were 2.232×2.304 pixels, 150 dpi, and 8 bits. The digital films were calibrated by digitizing 2 points (20 mm) on the ruler within the digital cassette. Landmark identification was carried out manually using a mouse-driven cursor. The screen used for computerized analysis was 21.5" in size. All measurements were performed automatically by the software (Figure 2).

App-aided Tracing

For the app tracing method, the CephNinja 3.51 app (Cyncronus LLC, WA, USA) was used. All the digital radiographs were uploaded as .jpeg files to Microsoft OneDrive using a standard computer. The files were in grayscale format, and the image properties of the film were 2.232×2.304 pixels, 150 dpi, and 8 bits. The radiographs were

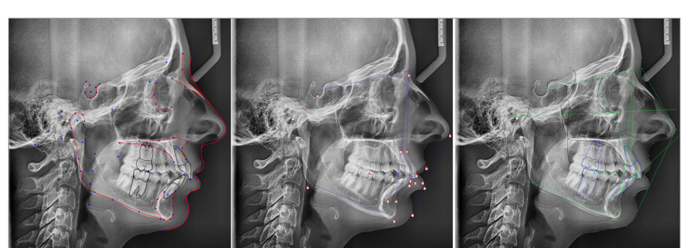

Figure 2. The same cephalometric radiograph traced with CephX (left), CephNinja (midmost) and Dolphin (right)

Table 2. Intra-class correlation coefficient (ICC) for reproducibility of each cephalometric analysis method

	CephX (cx)		CephX corrected (cxc)		CephNinja (cn)		Dolphin (d)		Manual (m) "ground truth"	
Measurements	**ICC**	**95%**	**ICC**	**95%**	**ICC**	**95%**	**ICC**	**95%**	**ICC**	**95%**
SKELETAL										
SNA (°)	0.965	0.902-0.988	0.914	0.768-0.970	0.893	0.713-0.963	0.940	0.833-0.979	0.973	0.913-0.991
SNB (°)	0.992	0.976-0.997	0.989	0.967-0.996	0.988	0.965-0.996	0.990	0.972-0.997	0.993	0.979-0.998
ANB (°)	0.983	0.949-0.995	0.977	0.935-0.992	0.988	0.964-0.996	0.975	0.902-0.992	0.992	0.976-0.997
GoGn-SN/GoMe-SN (°)	0.990	0.971-0.997	0.993	0.979-0.998	0.977	0.936-0.992	0.974	0.922-0.991	0.992	0.975-0.997
DENTAL										
I-NA (°)	0.750	0.409-0.908	0.838	0.594-0.942	0.964	0.898-0.988	0.970	0.916-0.990	0.985	0.955-0.995
I-NA (mm)	0.912	0.289-0.979	0.862	0.645-0.951	0.923	0.792-0.973	0.950	0.858-0.983	0.864	0.610-0.954
I-NB (°)	0.733	0.063-0.921	0.906	0.705-0.969	0.974	0.926-0.991	0.980	0.932-0.993	0.977	0.932-0.992
I-NB (mm)	0.824	0.436-0.943	0.947	0.853-0.982	0.969	0.911-0.989	0.982	0.943-0.994	0.950	0.857-0.983
I-I (°)	0.872	0.646-0.956	0.948	0.854-0.982	0.980	0.943-0.993	0.985	0.956-0.995	0.983	0.950-0.994
OCC-SN (°)	0.936	0.823-0.978	0.948	0.854-0.982	0.884	0.697-0.959	0.938	0.829-0.979	0.993	0.979-0.998
SOFT TISSUE										
UL E-line (mm)	0.890	0.553-0.967	0.968	0.557-0.993	0.988	0.963-0.996	0.987	0.963-0.996	0.996	0.976-0.997
LL E-line (mm)	0.990	0.973-0.997	0.991	0.974-0.997	0.987	0.962-0.995	0.993	0.981-0.998	0.993	0.981-0.998

(For the manual measurements, the "ground truth" values were used as the mean remeasurements of the 3 examiners)

Table 3. Intra-class correlation coefficient calculated for inter-examiner reliability of the manual tracings of 3 examiners

	Manual (m)	
Measurements	**ICC**	**95%**
SKELETAL		
SNA (°)	0.930	0.849 -0.965
SNB (°)	0.936	0.720 -0.976
ANB (°)	0.934	0.850-0.968
GoGn-SN (°)	0.950	0.824-0.980
DENTAL		
I-NA (°)	0.939	0.895-0.966
I-NA (mm)	0.900	0.830-0.944
I-NB (°)	0.980	0.966-0.989
I-NB (mm)	0.912	0.748-0.962
I-I (°)	0.928	0.878-0.959
OCC-SN (°)	0.931	0.881-0.962
SOFT TISSUE		
UL E-line (mm)	0.980	0.947-0.991
LL E-line (mm)	0.977	0.962-0.987

imported to the CephNinja app using an iPhone 6S (IOS 11.4) smartphone. The same calibration procedure (20 mm) was performed for the cephalometric films. Landmark identification was carried out manually on a smartphone screen using the index finger. The zoom-in/zoom-out function was used when needed (Figure 2).

Web-based Fully Automated Tracing

An online automatic cephalometric tracing and analysis service named CephX (ORCA Dental AI, Las Vegas, NV) was used. After entering the system with www.cephx.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 2.232×2.304 pixels, 150 dpi, and 8 bits. Once the images were uploaded, the AlgoCeph system automatically identified all the anatomical points. The screen used for the analysis was 21.5" in size. Calibration was set to 20 mm, and the analysis was downloaded to the computer without any correction (Figure 2). The same set of data, after the automatic tracing, was also manually corrected for landmark position and downloaded to the computer.

Statistical Analysis

Statistical analysis was conducted using the Statistical Package for 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. Inter-group comparisons were made with one-way analysis of variance (ANOVA), and the Tukey test was used for post hoc testing. Intra-class and inter-class (manual tracings) variations were studied using intra-class correlation coefficients (ICC) with a confidence interval of 95%.

RESULTS

The ICC values calculated for repeated measurements to detect the method of error with each tracing technique are reported in Table 2. For manual measurements, the "ground truth" values were used, that is, the mean remeasurements of the 3 operators. All the ICCs exceeded 0.85, except for the dental landmarks: I-NA (°), I-NB (°), and I-NB (mm) traced with CephX. These landmarks showed a higher ICC when manual correction was performed. Most of the other values were above 0.9, regardless of the tracing method used, thereby providing an indication of very high intra-rater reliability.

Table 4. Comparison of skeletal, dental and soft tissue measurements by 4 different cephalometric analysis methods

Measurements	CephX (cx)			CephX corrected (cxc)			CephNinja (cn)			Dolphin (d)			Manual (m) "ground truth"			p-value
	Min	Max	Mean ±SD	Min	Max	Mean ±SD	Min	Max	Mean ±SD	Min	Max	Mean ±SD	Min	Max	Mean ±SD	
SKELETAL																
SNA (°)	73.6	87.2	79.9 ± 3.2	73.8	87.2	80 ± 3.2	73.7	86.7	79.6 ± 3	72.4	89.7	79.6 ± 3.4	75	86	79.8±2.8	NS
SNB (°)	69.6	84.4	76.6 ± 3.2	70	84.4	76.7 ± 3.6	69.6	84.6	76.3 ± 3.6	69.1	87.7	76.5±4.4	70.5	84.5	76.6±3.4	NS
ANB (°)	-2.2	8.2	3.2 ± 2.5	-2.3	8	3.3 ± 2.6	-1.1	7.9	3.3 ± 2.5	-1.4	8.1	3.3 ± 2.6	-1	7,6	3.2±2.3	NS
GoGn-SN (°)/ GoMe-SN (°)	28.3	55.4	40.3 ± 6.5	25.1	54.4	38.7 ± 6.6	20	53.6	34.7 ± 7	24.3	54.4	35.7 ± 7.1	20.8	51	33.8±6.9	cx-cn**, cx-d*, cx-m**
DENTAL																
I-NA (°)	17.6	38.5	27.2 ± 4.5	16.3	38.4	27.6 ± 5.5	11.2	34.8	24.3 ± 5.5	14.2	35	24 ± 5.9	15	36.3	25.3±5.4	cx-d * cxc-m *
I-NA (mm)	1	9.4	4.8 ± 1.8	1.7	9.2	5.1 ± 1.8	2.7	8.5	6 ± 1.4	1.1	8.4	4.9 ± 1.9	3	9.33	6.3±1.6	cx-cn*, cx-m** cn-d*, d-m**, cxc-m **
I-NB (°)	10	33.5	22.5 ± 5.9	9.8	36.4	24.7 ± 6.4	11.3	43.2	27 ± 6.6	9.3	43.1	26.7 ± 7.5	12.3	40	26.8±6.5	cx-cn*, cx-d*, cx-m*
I-NB (mm)	-0.7	9.4	4.9 ± 2.5	-0.8	9	4.8 ± 2.4	1.2	9.2	5.7 ± 1.9	-0.5	10.2	5.3 ± 2.5	0.8	9.5	5.8±1.9	cxc-m *
I-I (°)	112.7	144.9	126.9 ± 8.2	107.9	152	124.2±9.4	103.7	152.4	125.3 ± 10	101.2	155.2	125.8 ± 11.2	106.6	149.6	124.3±9.2	NS
OCC-SN (°)	9	24.7	16.4 ± 4.3	9.9	25.3	16.6±4.3	9.9	30.1	17.6 ± 4.1	6	27.7	17.7 ± 4.7	8.3	26.3	17.1±4.1	NS
SOFT TISSUE																
UL E-line (mm)	-10.6	1.2	-4.5 ± 3.4	-9.9	0.9	-4.2±3.1	-9.3	1.9	-4.1 ± 2.9	-9.6	1.5	-4.4 ± 3	-9.33	1.33	-4.24±2.9	NS
LL E-line (mm)	-6.7	3.1	-1.2 ± 3	-7.2	3.7	-1.1±3	-7.5	4.1	-1.1 ± 2.7	-7.7	4.4	-1.2 ± 2.9	-8.1	3.8	-1.4±2.8	NS

NS: Non-significant; CX: CephX; CN: CephNinja; D: Dolphin; M: Manual; Min: Minimum; Max: Maximum; SD: Standard deviation
*p<0.05, **p<0.01, ***p<0.001
GoGn-SN (°) / GoMe-SN (°): CephX uses the GoMe plane instead of GoGn
For the manual measurements, the "ground truth" values were used: the mean measurements of 3 examiners

Table 5. Comparison of the analysis time between the 4 different cephalometric analysis methods

	CephX (cx)			CephX corrected (cxc)			CephNinja (cn)			Dolphin (d)			Manual (m) "ground truth"			p
	Min	Max	Mean±SD	Min	Max	Mean±SD	Min	Max	Mean±SD	Min	Max	Mean±SD	Min	Max	Mean±SD	
Duration (sec)	20	46	29.5 ± 5.4	35	81	58.7 ± 11.8	60	155	100.1 ±17.4	105	172	129.4±19.9	418	841	551±105	cx-cn***, cx-d***, cx-m***, cn-m***, d-m***, cxc-cx**, cxc-cn ***, cxc-d ***, cxc-m ***

**p<0.01, ***p<0.001, For the manual analysis, the mean time of measurements made by three examiners are given.
Min: Minimum; Max: Maximum; SD: Standard deviation

The ICC values for the inter-examiner correlation for manual measurements are shown in Table 3. All the values were above 0.9, indicating very high inter-examiner reliability between the 3 operators.

For the inter-group comparisons of the cephalometric values and the tracing times, the results of the one-way ANOVA and Tukey test are shown in Table 4 and Table 5, respectively.

Regarding the skeletal parameters, no statistically significant differences for SNA, SNB, and ANB were detected among the 4 tracing systems. The mean values for the GoGn-SN measurements were significantly higher in the CephX group than in the other 3 groups ($p<0.05$), but when manual correction was performed, the GoGn-SN value became similar to the values obtained by the other tracing methods (Table 4).

Regarding the dental parameters, there were no statistically significant differences among the 4 tracing systems for I-I and Occ-SN measurements. Significantly higher means of I-NA (°) were observed using CephX compared with Dolphin ($p<0.05$) and CephX corrected compared with manual tracing, whereas the mean I-NB (°) was significantly lower in CephX than in CephNinja, Dolphin, and manual tracing ($p<0.05$). The mean I-NA (mm) and I-NB (mm) were significantly lower in CephX than in manual tracing ($p<0.05$) regardless of manual correction, whereas higher values were obtained in CephNinja than in CephX and Dolphin ($p<0.05$).

The soft tissue measurements were similar in all 4 tracing systems ($p>0.05$).

The shortest analyzing time was obtained using CephX, followed by CephX corrected, CephNinja and Dolphin, whereas manual tracing took the longest time (Table 5).

DISCUSSION

Digital systems are increasingly used in cephalometry because of rapid advances in computer technology. Cephalometric analysis is not only available as computer software but also as applications on smartphones or online, where automatic tracing is possible. Regardless of the method used, the most important criteria for tracing are accuracy and a high rate of reproducibility. Therefore, the focus of this study was to compare the accuracy of manually traced lateral cephalograms with automatic, digitized, or app-aided tracings. The principal finding of this study was that automatic tracing with CephX is significantly faster than the other methods, but the software needs improvement to become more reliable in the majority of dental measurements and also for GoGn-SN (°). After manual correction of the landmarks on CephX, measurements similar to digitized and app-aided tracings can be obtained in a significantly shorter time.

The threshold for clinically relevant differences of cephalometric measurements varies in the literature; however, a difference that is statistically significant but is smaller than 2 units of measurement (millimeters or degrees) is considered to be within the clinically acceptable limits (12, 13). Therefore, all statistically significant differences for dental and skeletal measurements found in this study using CephX were also clinically significant. Thus, the null hypothesis that all the 4 methods were no different can be rejected regarding the skeletal measurement of GoGn-SN (°) and the dental measurments, including I-NA (°), I-NA (mm), I-NB (°), and I-NB (mm). The null hypothesis can be accepted for SNA (°), SNB (°), ANB (°), I-I (°), OCC-SN (°), UL E-line (mm), and LL E-line (mm).

Previous researches have shown that the inter-operator error is greater than the intra-operator error and that the experience of the observer when locating the landmarks also affects the random error (13-15). To avoid such error, the digital measurements in this study were carried out by a single experienced examiner. However, since fully automatic tracing systems are deterministic i.e., the same image will give the same result every time, unlike manual tracings, which are known for having high inter-observer error, the CephX measurements in this study were compared with a "manual ground truth" that was obtained with 3 manual observers instead of a single observer. The inter-examiner reproducibility for the manual tracings was very high, and the majority of the ICC values for the repeated measurements were also high, irrespective of the tracing method, indicating that a high intra-operator reliability (Tables 2 and 3).

The use of cephalometric software may diminish the errors that occur during manual tracings obtained by drawing and measuring with a ruler and a protractor (16, 17). However, some measurements, particularly those involving the maxillary and mandibular incisors, are difficult to identify; hence, such structures have been shown to have low reliability not only in manual but also in digital tracings, despite the possibility of using filtering and zooming (14, 18). These results are in accordance with our study, i.e., the measurements related to the landmarks including incisors revealed significant differences, and the most unreliable system was the CephX, which was also reflected by the lower ICC values. This error may be due to where the incisal landmark is placed. CephNinja and manual tracing use the most prominent facial point of the incisor, whereas CephX uses the incisal point of the tooth. Additionally, automatic identification of landmarks is always associated with an error that will increase if a line consisting of 2 points is measured because this will be affected by the errors of the 2 points rather than a single one (19). However, when CephX landmarks were manually corrected, dental linear and angular measurements on incisors were similar to app-aided and computerized tracings (Table 4). In general, angular measurements were more reliable than linear ones, especially when CephNinja, Dolphin, and manual tracing were used. These results are in accordance with the findings obtained by other investigators (20, 21).

Previous studies have reported that nasion and gonion are inconsistent points and sources of mistakes, which is in line with the findings of our study, as measurements related to these points revealed significant differences using CephX. Another source of error that may explain the higher GoGn-SN (°) measurements obtained by CephX and compared to the other methods, is that the program uses the GoMe plane instead of the GoGn plane when the GoGn-SN value is measured (12, 22). This shortcoming can be adjusted by manually correcting the landmark.

Correction of automatically traced points on CephX has been performed by other investigators, resulting in clinically insignificant FMA angle (Frankfurt plane/Mandibular plane angle) obtained by the CephX group compared with computerized tracing group (23).

The resolution of the images is an important criterion for the validity of the results. Digital images of 150 dpi, 8 bits, have been reported to be sufficient for clinical purposes (7). In this study, a resolution of 150 dpi was used for all the 4 tracing methods to allow for comparison and also because it is recommended by the software manufacturers as it facilitates identification of the landmarks. The specific anatomical landmarks used in this study were chosen partly because the app-based application did not offer more parameters and partly because of the commonalities among the 4 methods. Moreover, these landmarks were chosen also because they are commonly used for orthodontic diagnosis and treatment planning.

The time required to identify and trace the anatomical structures differs significantly between experienced and inexperienced operators when digital tracings are used but cannot be reduced with manual tracing (21). In this study, the time required to make the digital measurements was substantially shorter than for the manual method, which is in line with the findings by other investigators (7, 11). Analyzing time with CephX was found to be 13 times faster than manual tracing and about 3 times faster than with CephNinja and Dolphin (Table 5). Also, when manual correction of CephX landmarks was made, the analyzing time was significantly shorter compared with the other methods. The time required to make a cephalometric analysis should not include the time required to make a diagnosis or a treatment plan. Even if the shortest tracing time was obtained with CephX, it was also the method that was less reliable in the majority of the dental measurements. The reliability and the validity of the tracing method should, therefore, always be superior to the tracing time; however, it should be pointed out that the manual correction of CephX landmarks results in similar measurements to CephNinja and Dolphin and seems to be a promising method for use in the clinical practice.

CONCLUSION

With the development of fully automated methods, cephalometric analyses can be performed faster and more reliably in the near future. On the basis of the results from this study, it can be concluded that CephX requires improvement to provide similar results as the other methods that were assessed. However, manual correction of CephX landmarks gives equivalent results to digital tracings using CephNinja and Dolphin with significantly shorter analyzing time.

Ethics Committee Approval: This study was approved by the Research Ethics Committee of Trakya University, Faculty of Medicine (Approval No: TÜTF-BAEK 2017/318).

Informed Consent: Written informed consent was obtained from the patients who agreed to take part in the study.

Peer-review: Externally peer-reviewed.

Author Contributions: Supervision – P.M., J.N.; Design – P.M., J.N.; Resources – P.M.; Materials – P.M.; Data Collection and/or Processing – P.M.; Analysis and/or Interpretation – P.M., J.N.; Literature Search – P.M., J.N.; Writing Manuscript – P.M., J.N.; Critical Review – P.M., J.N.

Conflict of Interest: The authors have no conflict of interest to declare.

Financial Disclosure: This study was supported by Trakya University Scientific Research Projects (Project No: 2018/39).

Acknowledgements: We would like to thank Melis Seki and Sevgi Ersay who helped in tracing the manual cephalometric films.

REFERENCES

1. Broadbent BH. A new x-ray technique and its application to orthodontia. Angle Orthod 1931; 1: 45-66.
2. Naoumova J, Lindman R. A comparison of manual traced images and corresponding scanned radiographs digitally traced. Eur J Orthod 2009; 31: 247-53. [Crossref]
3. Phatak SM, Daokar SS. Orthodontic apps: A stairway to the future. Int J Orthod Rehabil 2019; 10: 75-81. [Crossref]
4. 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. [Crossref]
5. 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; 89: 889-96. [Crossref]
6. Forsyth DB, Shaw WC, Richmond S, Robert CT. Digital imaging of cephalometric radiographs. Part 2: image quality. Angle Orthod 1996; 66: 43-60.
7. Sayinsu K, Isik F, Trakyali G, Arun Tülin. An evaluation of the errors in cephalometric measurements on scanned cephalometric images and conventional tracings. Eur J Orthod 2007; 29: 105-8. [Crossref]
8. Celik E, Polat-Ozsoy O, Toygar-Memikoglu TU. Comparison of cephalometric measurements with digital versus conventional cephalometric analysis. Eur J Orthod 2009; 31: 241-6. [Crossref]
9. Mosleh MAA, Baba MS, Malek S, Almaktari RA. Ceph-X: Development and evaluation of 2D cephalometric system. BMC Bioinformatics 2016; 17: 499. [Crossref]
10. Shettigar P, Shetty S, Naik RD, Basavaraddi SM, Patil AK. A Comparative evaluation of reliability of an android-based app and computerized cephalometric tracing program for orthodontic cephalometric analysis. Biomed Pharmacol J 2019; 12: 341-6. [Crossref]
11. 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: 523-8. [Crossref]
12. 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.
13. Schulze RKW, Gloede MB, Doll GM. Landmark identification on direct digital versus film based cephalometric radiographs: A human skull study. Am J Orthod Dentofacial Orthop 2002; 122: 635-42. [Crossref]

14. Paixão MB, Sobra MC, Vogel CJ, de Araujo TM. Comparative study between manual and digital cephalometric tracing using Dolphin Imaging software with lateral radiographs. Dent Press J Orthod 2010; 15: 123-30. [Crossref]
15. Lagravere MO, Low C, Flores-Mir C, Chung R, Carey JP, Heo G, et al. Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. Am J Orthod Dentofacial Orthop 2010; 137: 598-604. [Crossref]
16. Baumrind S, Frantz RC. The reliability of head film measurements: 1. Landmark identification. Am J Orthod 1971; 60: 111-27. [Crossref]
17. Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmarks identification and reproducibility: A meta analysis. Am J Orthod Dentofacial Orthop 1997; 112: 165-70. [Crossref]
18. Chan CK, Tng TH, Hägg U, Cooke MS. Effects of cephalometric landmark validity on incisor angulation. Am J Orthod Dentofacial Orthop 1994; 106: 487-95. [Crossref]
19. Chen YJ, Chen SK, Chang HF, Chen KC. Comparison of landmark identification in traditional versus computer-aided digital cephalometry. Angle Orthod 2000; 70: 387-92.
20. Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis. Angle Orthod 2008; 78: 145-51. [Crossref]
21. Geelen W, Wenzel A, Gotfredsen E, Kruger M, Hansson LG. Reproducibility of cephalometric landmarks on conventional film, hardcopy, and monitor-displayed images obtained by the storage phosphor technique. Eur J Orthod. 1998; 20: 331-40. [Crossref]
22. Aksakallı S, Yılancı H, Görükmez E, Ramoğlu Sİ. Reliability assessment of orthodontic apps for cephalometrics. Turk J Orthod 2016; 29: 98-102. [Crossref]
23. Alqahtani H. Evaluation of an online website-based platform for cephalometric analysis. J Stomatol Oral Maxillofac Surg 2020; 121: 53-7. [Crossref]