Title

Comparison Between Computerized and Web-based Fully Automated Cephalometric Analysis in Cleft Lip and Palate Patients

Author
Fabros, Marissa

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Comparison Between Computerized and Web-based Fully Automated Cephalometric Analysis in Cleft Lip and Palate Patients

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science in Oral Biology

by

Marissa Nicole Fabros

2023

ABSTRACT OF THE THESIS

Comparison Between Computerized and Web-based Fully Automated Cephalometric Analysis in Cleft Lip and Palate Patients

by

Marissa Nicole Fabros

Master of Science in Oral Biology

University of California, Los Angeles, 2023

Professor Sanjay M. Mallya, Chair

Cleft lip with or without cleft palate (CL/P) is one of the most common craniofacial anomalies. Patients with CL/P often require extensive and prolonged orthodontic treatment due the skeletal and dental malocclusions that frequently manifest, making the orthodontic burden of care significant compared to the non-cleft patient. Orthodontic treatment may be required at various stages in their dental and skeletal development, during which orthodontic records must regularly be taken, including cephalometric radiographs. These radiographs allow for analysis of skeletal and dental relationships by relating various anatomical landmarks through linear and angular measurements.

Due to the time-consuming nature of the current computer-aided method, which currently involves an orthodontist locating points of a lateral skull radiograph on a computer monitor,

systems have recently been developed to automate the cephalometric process. However, while systems such as CephX have been shown to significantly shorten analyzing time, the accuracy of the measurements is inadequate. Further, radiographs of CL/P-affected patients are often excluded from the sample, as identification of cephalometric landmarks is more complicated due to abnormal anatomy. There is promise that lateral simulated 2D cephalometric projections from CBCTs improve the accuracy of cephalometric measurements over 2D cephalograms.

This study sought to compare the accuracy and analyzing time between web-based fully automated and computer-aided cephalometric analysis of lateral cephalometric images derived from cone-beam computed tomography in unilateral cleft lip and palate-affected patients. Both methods of cephalometric analysis were performed on 36 CBCTs obtained of individuals with unilateral cleft lip and palate in the mixed dentition stage.

Of the 12 measurements obtained, 4 measurements, U1-PP (°), SN-MP (°), U1-SN (°) and U1-NA (mm) were both statistically significant ($P<0.05$) and had mean differences above the clinically acceptable limit of 2 mm or 2°. The agreement interval fell outside the range of clinical acceptability for every measurement. Results also showed that the automated program took significantly longer than the computer-assisted method to produce a cephalometric analysis.

Based on these results, it is advisable to use CephX for cephalometric analysis in patients with cleft lip and palate only with clinician supervision and intervention. Further development is needed, particularly with regard obtaining measurements involving landmarks that are challenging to identify and those with multiple definitions before completely replacing computerized tracing.

The thesis of Marissa Nicole Fabros is approved.

Jimmy Kuanghsian Hu

Yong Kim

Sanjay M. Mallya, Committee Chair

University of California, Los Angeles

2023

TABLE OF CONTENTS

List of Abbreviations

Abbreviation	Definition
2D	2-dimensional
3D	3-dimensional
ABO	American Board of Orthodontics
AI	Artificial intelligence
ANS	Anterior nasal spine
CBCT	Cone-beam computed tomography
EP	E-Plane
FMA	Frankfort-mandibular plane angle
ICC	Intraclass correlation coefficient
L1	Lower incisor
MP	Mandibular plane
MPR	Multi-planar reformatting
PNS	Posterior nasal spine
PP	Palatal plane
RLC	Reconstructed lateral cephalogram
SDR	Successful detection rate
U1	Upper incisor

List of Figures and Tables

Acknowledgements

I would like to thank Dr. Sanjay Mallya for giving me the opportunity to work on this project and continually offering his guidance, mentorship, and expertise as my PI during my academic research. I also want to thank Dr. Jimmy Hu and Dr. Yong Kim, both of whom I am honored to have had as committee members, for their invaluable feedback.

Additionally, I would like to offer my special thanks to Dr. Andrew Paige for his encouragement and advice throughout this process. I would also like to express my sincerest gratitude for my friends and family, especially my parents, Kathy and Bobby Fabros, whose unwavering support has been critical to my success in my academic endeavors.

INTRODUCTION

Orofacial clefts, which include cleft lip with or without cleft palate, are among the most common major congenital defects in humans [1]. It is estimated that about 1 in every 700 babies is born with a cleft lip with or without cleft palate, and its incidence varies among populations, proportionately occurring more often among Asian and Native American populations, and less frequently in African American populations [2].

A cleft lip with or without a cleft palate is characterized by partial or complete clefting of the upper lip, with or without clefting of the alveolar ridge or the hard palate [3]. These congenital malformations result from a partial or complete lack of fusion of the maxillary prominence with the medial nasal prominence on one or both sides, resulting in a unilateral or bilateral cleft, respectively [2]. Depending on the severity, children with a cleft lip with or without a cleft palate can present with feeding, speaking, hearing and/or dental issues [4].

Extensive and prolonged orthodontic treatment is often required for cleft-affected patients and the orthodontic burden of care for patients with cleft lip and/or palate must not be underestimated [5,6]. Cleft-affected patients experience a range of skeletal and dental issues when compared to the non-cleft patient. Differences in skeletal growth patterns and dimensions among cleft patients can manifest in the maxillary arch in the vertical, antero-posterior, and transverse dimensions due to the severity of the initial cleft deformity and post-primary surgical growth, resulting in a skeletal malocclusion [7]. Cleft patients also often present with one or more dental anomalies, such as agenesis, ectopic eruption, and supernumerary teeth, that can contribute to a dental malocclusion [8].

As a result, cleft patients often require complex long-term orthodontic treatment that involves being seen multiple times by an orthodontist from a young age to monitor their dental development and jaw growth [6, 9]. A study assessing the orthodontic burden of care for patients with a cleft lip and/or palate found that patients attended an average of 44 orthodontic appointments and required a duration of treatment averaging over 3 years [6].

Orthodontic treatment may be involved at any or all of four different stages: (1) in infancy before the initial surgical repair of the lip to reposition the maxillary segments (2) during the late primary and early mixed dentition to address orthodontic problems resulting from surgical repair as well as to prepare the teeth and arch for alveolar bone grafting (3) during the late mixed and early permanent dentition to address malalignment and (4) in late adolescence after the completion of facial growth, in conjunction with orthognathic surgery [10]. At various stages in cleft patient management, orthodontic records must regularly be taken, including radiographs [9].

Lateral cephalometric radiographs are the standard for radiographically evaluating dentofacial proportions for comprehensive orthodontic treatment. Cephalometric analysis enables determination of skeletal and dental relationships by relating various cephalometric landmarks through linear and angular measurements. These landmarks represented as a series of points that may be defined as locations on a physical structure, extreme points, or constructed points [10]. However, current methods for identifying landmarks to perform cephalometric analysis are time-consuming [11].

Historically, acetate overlays were used to trace lateral cephalograms manually by identifying landmarks and obtaining angular and linear measurements with a ruler and protractor [12]. With advances in technology, manual tracing methods have been slowly replaced with digital cephalometric analysis software [13]. Computer-aided cephalometric analysis involves an orthodontist to locate landmarks of a lateral skull radiograph on a computer monitor and the software completes the cephalometric analysis by automatically measuring distances and angles [11]. Studies have found that computerized cephalometric analyses are similar in terms of accuracy and reproducibility and enable faster analyses compared to those done manually [14, 15].

More recently, systems have been developed to automate the cephalometric process, which involves automated landmarking of cephalograms [16]. Automated cephalometric analysis involves storing a scanned or digital cephalometric radiograph in the computer and loading it to the software that automatically locates the landmarks using computer vision and artificial intelligence techniques and performs the measurements for analysis [11]. The purpose of this technology is to reduce the time required to obtain an analysis, improve the accuracy of landmark identification, and reduce errors due to clinician subjectivity [17]. Studies have shown that fully automated tracing programs produced by different commercial providers can significantly shorten analyzing time but are not as accurate as the computerized method [13, 18].

While numerous studies have explored the clinical applicability of artificial intelligence systems that recognize cephalometric landmarks, few have specifically assessed whether these systems can be applied in the evaluation of cleft-affected patients [19]. Identification of cephalometric landmarks is even more complicated in young cleft-affected patients due to abnormal anatomy. Localizing the landmarks posterior nasal spine (PNS), anterior nasal spine (ANS) and point A are especially difficult due to reduced radiopacity caused by the cleft. Further complicating locating ANS is the fact that it is not positioned in the midline due to outward rotation of the larger segment of the maxilla [20]. The results of one study found that cleft lip and/or palate was a factor associated with greater identification errors when artificial intelligence was used for automatic recognition of anatomic landmarks compared to gold standard values on lateral cephalograms [19].

However, another study shows promise that lateral simulated 2-dimensional (2D) cephalometric projections from cone-beam computed tomography (CBCT) improved the accuracy of cephalometric measurements over conventional 2D cephalograms. CBCT allows 2D multi-planar reformatting (MPR) and secondary reconstruction of the data. This enables the generation of images in multiple orientations, thus preventing unnecessary exposure of patients to radiation. It also enables evaluation of craniofacial structures without encountering issues related to anatomic superimpositions and differential magnification of bilateral structures. It also enables us to leverage the significant database of information connecting 2D standardized head radiographs to orthodontic treatment planning and outcomes assessments that has been collected for almost a century and apply it to 3D imaging technology. For most 2D cephalometric measurements in the sagittal plane, it was found that lateral simulated 2D cephalometric projections from CBCT had improved accuracy over conventional 2D cephalograms [21].

While CBCTs are not always indicated for orthodontic patients, there is sufficient data supporting the use of CBCT in patients with congenital deformities, the major condition being cleft palate. For these patients, the CBCT provides the information to determine the location and extent of the clefting, the location of any impacted or supernumerary teeth relative to the cleft, and the location and timing of an alveolar bone graft so that erupting teeth near the cleft site can bring new bone with them [10].

CephX is an online platform that uses artificial intelligence (AI) technologies to automate diagnostic and analytical dental imaging tasks, including cephalometric analysis. It offers the capability of performing cephalometric analysis based on conventional 2D cephalometric radiographs as well as CBCT scans, whereby its algorithm converts a 3-dimensional (3D) DICOM file to a 2D cephalometric radiograph, locates landmark coordinates and produces cephalometric tracing and analysis [22].

While numerous studies have validated the clinical applicability of AI systems that recognize cephalometric landmarks, few have specifically assessed whether these systems can be employed in the evaluation of cleft-affected patients. Those studies that do assess the accuracy of automated cephalometric measurements in patients with orofacial clefts did not apply this technology to CBCT-derived lateral cephalograms [19].

OVERALL OBJECTIVES AND SPECIFIC AIMS

The goal of this study is to identify a timesaving and comparably accurate alternative to computerized cephalometric analysis of radiographs obtained from unilateral cleft lip and palate-affected patients.

Aim 1: Compare the accuracy of cephalometric measurements of web-based fully automated to computerized tracing of lateral cephalometric images derived from CBCTs in unilateral cleft lip and palate-affected patients

- Evaluate differences in angular and linear measurements (in °/mm)
- Statistical analysis of data using paired *t* tests ($P<0.05$) to detect significant differences and Bland-Altman plots to quantify the agreement between each measurement

Aim 2: Compare the analyzing time of web-based fully automated to computerized analysis of lateral cephalometric images derived from CBCTs in unilateral cleft lip and palate-affected patients

- Measure analyzing time (in minutes and seconds) using both methods

MATERIALS AND METHODS

Data Collection

For this study, CBCT volumes were used to produce reconstructed lateral cephalograms (RLCs). CBCT scans were obtained from the patient database at the UCLA Orthodontics Clinic, originally collected for the purposes of diagnosis and treatment planning. A sample size of 35 images was determined based on the result of a calculation made in previous study. The effect size was calculated to 0.49 and a statistical power of 80% at a significance level of 0.05 was assumed by the G*power program (version 3.1, Heinrich- Heine-University Dusseldorf, Germany). [23] Records of 930 patients with 3D images were examined, and 36 CBCTs were selected and designated as samples based on inclusion and exclusion criteria.

The following inclusion criteria were applied: (1) presence of unilateral cleft lip and palate, and (2) patients in the mixed dentition stage. Patients with any syndrome with craniofacial abnormalities or who had previously undergone alveolar bone graft surgery as documented in the medical history were excluded. The distributions of specific characteristics within the sample are described in Table 1.

Table 1 Characteristics of the sample

		N (%)
Total		36 (100%)
Age		
Female	9.01 ± 1.32	13 (36%)
Male	9.30 ± 1.03	23 (64%)
Cleft Type		
Left		24 (67%)
Right		12 (33%)
Ethnicity		
Asian		4 (13%)
Caucasian		5 (13%)
Hispanic		24 (75%)
Dentition		
Early Mixed		23 (64%)
Late Mixed		13 (36%)

All CBCT scans were acquired by the same CBCT scanner (NewTom 5G; Cefla, Charlotte, NC) at the UCLA Oral & Maxillofacial Radiology clinic with consistent acquisition parameters and saved in DICOM format. Any scans with artifacts that could interfere with the anatomical point identification were excluded.

Lateral Cephalogram Reconstruction

The DICOM data for each CBCT was imported into the Dolphin Imaging software (version 11.95; Dolphin Imaging & Management Solutions, Chatsworth, CA). The hard-tissue 3D rendering was then oriented by the orthodontist in the sagittal, axial, and coronal views. Once oriented, a right-facing two-dimensional lateral cephalometric radiograph was reconstructed from the 3D volume dataset and saved in the Dolphin Imaging database for cephalometric tracing.

Cephalometric Analysis

The landmarks identified in this study were selected based on those used to generate American Board of Orthodontics (ABO) cephalometric measurements, including 8 angular measurements (SNA, SNB, SN-MP, FMA, ANB, U1-SN, L1-MP) and 4 linear measurements (U1-NA, L1-NB, LL-EP, UL-EP) [24]. Two additional landmarks regarded as difficult to identify in cleft-affected patients (ANS and PNS) were selected and used to generate a single angular measurement (U1-PP) [20]. The landmarks and measurements are represented on a tracing in Figure 1 and defined in Table 2 and Table 3, respectively.

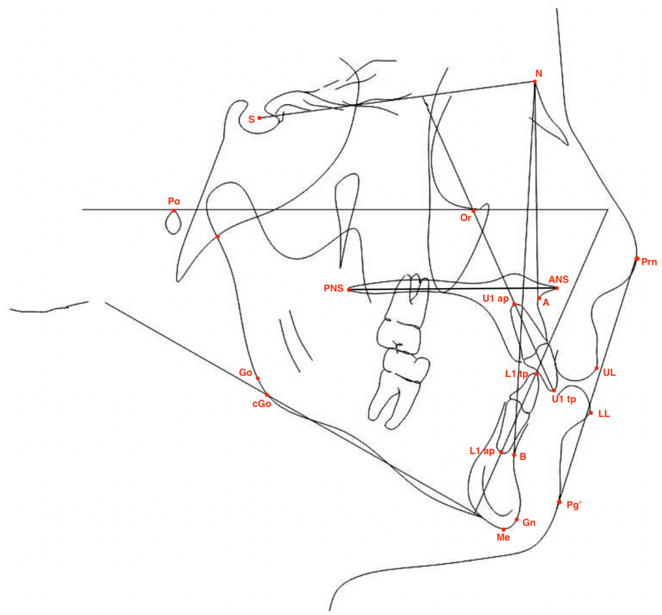


Figure 1 Cephalometric landmarks identified and measurements compared in this study. The definitions are listed in Table 1 and Table 2

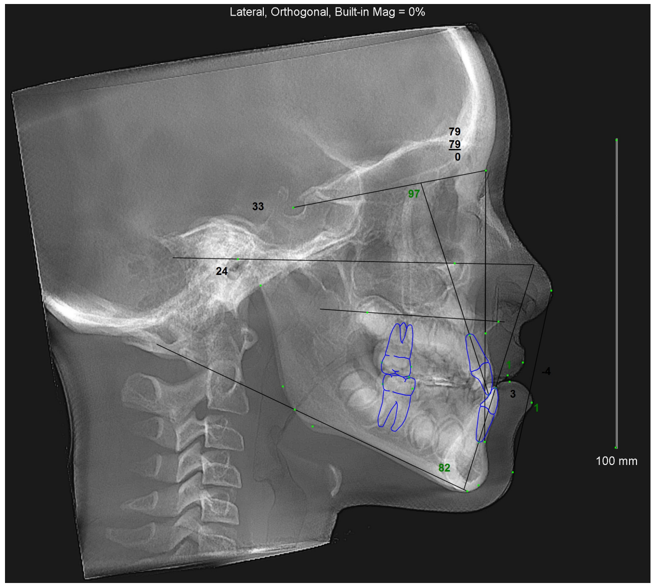
Table 2 Cephalometric landmark definitions [20, 23-25]

Landmark	Definition
Sella (S)	Midpoint of the sella turcica (hypophyseal or pituitary fossa)
Nasion (N)	Most anterior point of nasofrontal suture in the median plane
Porion (Po)	The upper- and outer-most point on the external auditory meatus
Orbitale (Or)	The most inferior and anterior point of the orbital margin
Point A (A)	The deepest point on the curved profile of the anterior portion of the maxilla, between the anterior nasal spine and alveolar crest
Menton (Me)	Lowest point of the mandibular symphysis in the midline
Gnathion (Gn)	Most anterior and inferior point of the bony chin (midpoint between pogonion and menton)
Point B (B)	The deepest point on the curved profile of the mandible, between the chin and the alveolar crest
Articulare (Ar)	The point of the intersection of the posterior margin of the ascending mandibular ramus and the outer margin of the posterior cranial base
Gonion (Go)	The most posterior and inferior point on the angle of the mandible
Upper incisor apex (U1 ap)	The root apex of the most anterior maxillary central incisor
Upper incisor tip (U1 tp)	The tip of the crown of the most anterior maxillary central incisor
Lower incisor tip (L1 tp)	The tip of the crown of the most anterior mandibular central incisor
Lower incisor apex (L1 ap)	The root apex of the most anterior mandibular central incisor
Anterior nasal spine (ANS)	The tip of the bony anterior nasal spine
Posterior nasal spine (PNS)	The posterior end of the hard palate, if visible; otherwise at the point of intersection of the dorsal maxillary contour and the soft palate contour
Pronasale (Prn)	The most prominent point of apex nasi
Upper lip (UL)	The most prominent point of the border of the upper lip
Lower lip (LL)	The most prominent part of the border of the lower lip
Soft tissue pogonion (Pg')	The most anterior soft tissue point of the chin in the midsagittal plane
Constructed Gonion (cGo)	The intersection of the mandible and the line that bisects the angle formed by Menton-Gonion and Articulare-Ramus Point

Table 3 Cephalometric measurement definitions [23, 25]

Landmark	Definition
SNA (°)	The angle formed between points S, N and A
SNB (°)	The angle formed between points S, N and B
SN-MP (°)	The angle formed by mandibular plane (cGo-Me) and line S-N
FMA (°)	The angle formed by mandibular plane and Frankfort horizontal plane (P-Or)
ANB (°)	The angle formed between points A, N and B
U1-NA (mm)	The distance between point upper incisor tip and line N-A
U1-SN (°)	The angle formed by upper incisor axes and line S-N
U1-PP (°)	The distance between upper incisor axes and palatal plane (line ANS-PNS)
L1 to NB (mm)	The distance between point lower incisor tip and line N-B
L1 to MP (°)	The angle formed by lower incisor axes and mandibular plane (cGo-Me)
LL-EP (mm)	The distance between point LL and E plane (line Prn-Pg')
UL-EP (mm)	The distance between point UL and E plane (line Prn-Pg')

Using a mouse-driven cursor, two dentists in the third year of their post-graduate education residents having undergone cephalometric training at the UCLA School of Dentistry performed landmark identification for each RLC based on the definitions in Table 2 on each of the cephalograms, based up on which a cephalometric tracing (Fig. 1) and analysis (Fig. 2) were automatically produced. The mean of the measurements was used to form the "anatomic truth" values. To verify intra-operator reliability, 10 radiographs were randomly selected and retraced digitally by the same orthodontist. Inter-operator precision was validated by comparing the cephalometric values between the two clinicians.



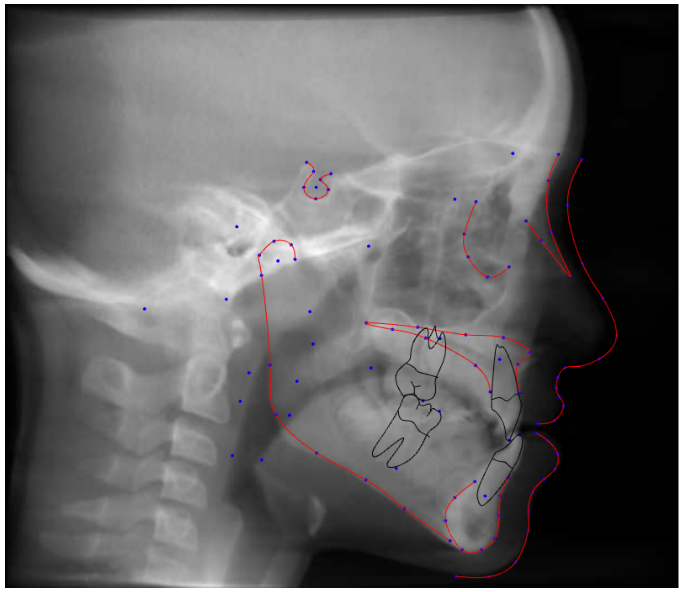
(a)

	Value	Norm	Std Dev	Dev Nor
Maxilla to Cranial Base				
SNA (°)	79.0	82.0	3.5	-0.9
Mandible to Cranial Base				
SNB (°)	78.9	80.9	3.4	-0.6
SN - MP (°)	33.2	32.9	5.2	0.1
FMA (MP-FH) (°)	24.2	23.9	4.5	0.1
Maxillo-Mandibular				
ANB (°)	0.1	1.6	1.5	-1.0 *
Maxillary Dentition				
U1 - NA (mm)	1.1	4.3	2.7	-1.2 *
U1 - SN (°)	96.9	102.8	5.5	-1.1 *
U1 - Palatal Plane (°)	111.8	112.0	6.0	0.0
Mandibular Dentition				
L1 - NB (mm)	3.5	4.0	1.8	-0.3
L1 - MP (°)	82.0	95.0	7.0	-1.9 *
Soft Tissue				
Lower Lip to E-Plane (mm)	1.3	-2.0	2.0	1.7 *
Upper Lip to E-Plane (mm)	-4.2	-6.0	2.0	0.9

(b)

Figure 2 Sample (a) digitized lateral cephalometric tracing and (b) analysis produced by Dolphin software

Anonymized DICOM data for each CBCT exported from the Dolphin Imaging software was uploaded to the CephX (ORCA Dental AI, Las Vegas, NV) online service via www.cephx.com using a standard web browser. The CephX algorithm, AlgoCeph, automatically converted the 3D DICOM file to a 2D cephalometric radiograph, located the landmarks, and produced a cephalometric tracing (Fig. 3) and analysis (Fig. 4). No reliability test was required, as landmark identification was performed automatically and thus yielded identical repeat measurements.

(a) 

(b)

Descriptor	Type	Mean	Sd	Patient	Graph
SNA	Deg	82.0	2.0	78.85	-(*|)+
SNB	Deg	80.0	2.0	79.66	-(*)+
SN-MP	Deg	33.0	4.0	35.26	-(*)+
FMA	Deg	28.0	3.0	21.89	-(* |)+
ANB	Deg	2.0	2.0	-0.81	-(*|)+
U1-NA	mm	4.3	2.7	1.22	-(*|)+
U1-SN	Deg	102.8	5.5	95.32	-(*|)+
U1-PP	Deg	112.0	6.0	107.07	-(*)+
L1-NB	mm	4.0	1.8	2.78	-(*)+
L1-MP	Deg	92.0	3.4	83.48	-(* |)+
Lower Lip to E-Plane	mm	-2.0	2.0	-1.48	-(*)+
Upper Lip to E-Plane	mm	-6.0	2.0	-5.95	-(*)+

Figure 3 Sample (a) digitized lateral cephalometric tracing and (b) analysis produced by CephX software

Analyzing time for each analysis was measured in minutes and seconds using a timer. For computerized tracing, time between the initiation of importing the 3D volume and the point at which the cephalometric measurements were made available by the software was recorded. For web-based fully automated tracing, the analyzing time was designated as the time it took for the software to upload, process, and produce the cephalometric measurements from the DICOM file.

Statistical Analysis

Statistical analysis was conducted using IBM SPSS Statistics (version 28.0; IBM Corp., Armonk, NY). The intraclass correlation coefficient (ICC) with a confidence interval of 95% was used to assess intra-rater reliability, where values below 0.5 indicate poor reliability, between 0.5 and 0.75 moderate reliability, between 0.75 and 0.9 good reliability and any value about 0.9 excellent reliability [26].

The mean, standard deviation, and successful detection rate (SDR) at precision ranges of 1, 2, 3 and 4 (millimeters for linear measurements and degrees for angular measurements) of the absolute differences between each measurement were analyzed to evaluate inter-rater reliability and assess the accuracy of CephX. Measurements were considered clinically acceptable when the difference in the angular and linear measurements was less than 2° or 2 mm, respectively [23, 27].

Paired *t*-test analysis was conducted to assess the differences between the means of each cephalometric measurement based on the landmarks identified by the orthodontists and those identified by CephX. Statistical significance was defined as a P value < 0.05. Bland-Altman plots were generated for each measurement to visualize the agreement between the computer-aided and CephX analysis, with the difference between CephX and the anatomic truth (y-axis) plotted against the mean of the measurements derived from CephX and clinician localizations (x-axis). An agreement interval was constructed, within which 95% of the differences lie. The upper and lower limits of agreement were set by multiplying the standard deviation by 1.96 and adding this value to and subtracting this value from the mean difference, respectively.

RESULTS

The ICC values showed a mean intra-operator reproducibility of 0.988 (0.953-0.997) and 0.986 (0.939-0.997) for Clinician A and Clinician B, respectively (Table 4), indicating excellent intra-examiner reliability.

Table 4 Intra-rater reliability for clinician A and clinician B

	Clinician A		Clinician B	
	ICC	**95%**	**ICC**	**95%**
Skeletal measurement				
SNA (°)	.976	.909-.994	.984	.94-.996
SNB (°)	.995	.982-.999	.996	.986-.999
SN-MP (°)	.997	.989-.999	.992	.967-.998
FMA (°)	.993	.975-.998	.995	.976-.999
ANB (°)	.995	.979-.999	.994	.975-.998
Dental measurement				
U1-NA (mm)	.953	.823-.988	.970	.884-.992
U1-SN (°)	.990	.962-.997	.992	.97-.998
U1-PP (°)	.974	.891-.994	.939	.774-.984
L1-NB (mm)	.996	.984-.999	.992	.969-.998
L1-MP (°)	.994	.978-.999	.985	.917-.996
Soft tissue measurement				
LL-EP (mm)	.997	.988-.999	.995	.979-.999
UL-EP (mm)	.992	.972-.998	.997	.988-.999

The absolute difference between the mean of each cephalometric measurement acquired by Clinician A and Clinician B was below the clinically acceptable limit of 2 mm or 2° except for U1-PP (°), which had a slightly increased mean error of 2.09° (Table 5). Given the otherwise high consistency between clinician measurements, the mean values of Clinician A and Clinician B's analyses were used in this study and designated as "anatomic truth" values.

Table 5 Inter-rater reliability between clinician A and clinician B

	Absolute Error		Successful Detection Rate (%)			
	Mean	SD	≤1 unit	≤2 units	≤3 units	≤4 units
Skeletal measurement						
SNA (°)	0.76	0.51	69.4	100.0	100.0	100.0
SNB (°)	0.59	0.40	86.1	100.0	100.0	100.0
SN-MP (°)	0.98	0.74	63.9	88.9	97.2	100.0
FMA (°)	0.88	0.60	72.2	94.4	100.0	100.0
ANB (°)	0.38	0.28	100.0	100.0	100.0	100.0
Dental measurement						
U1-NA (mm)	0.59	0.37	88.9	100.0	100.0	100.0
U1-SN (°)	1.53	1.05	41.7	69.4	88.9	97.2
U1-PP (°)	2.09	1.62	30.6	58.3	72.2	83.3
L1-NB (mm)	0.22	0.18	100.0	100.0	100.0	100.0
L1-MP (°)	1.05	0.71	50.0	88.9	100.0	100.0
Soft tissue measurement						
LL-EP (mm)	0.28	0.23	100.0	100.0	100.0	100.0
UL-EP (mm)	0.31	0.23	100.0	100.0	100.0	100.0

Considering the clinical relevance of any difference above 2 mm or 2°, significant differences were observed between the mean of 5 measurements, including U1-PP (°) (6.17±5.40), U1-SN (°) (5.72±6.94), SN-MP (°) (2.47±1.62), U1-NA (mm) (2.45±2.07) and L1-MP (°) (2.32±1.98). However, successful detection rate shows that there were individual instances of each measurement greater than 2 mm or 2° (Table 6).

Table 6 Comparison of the absolute differences of cephalometric measurements between computer-aided and CephX analyses

	Absolute Error		Successful Detection Rate (%)			
	Mean	SD	≤1 unit	≤2 units	≤3 units	≤4 units
Skeletal measurement						
SNA (°)	1.52	1.25	47.2	75.0	86.1	97.2
SNB (°)	0.79	0.63	72.2	91.7	100.0	100.0
SN-MP (°)	2.47	1.62	19.4	41.7	66.7	83.3
FMA (°)	1.82	1.42	38.9	58.3	75.0	94.4
ANB (°)	1.58	1.13	33.3	69.4	91.7	97.2
Dental measurement						
U1-NA (mm)	2.45	2.07	36.1	50.0	63.9	77.8
U1-SN (°)	5.72	6.94	19.4	27.8	36.1	50.0
U1-PP (°)	6.17	5.40	2.8	16.7	30.6	36.1
L1-NB (mm)	0.94	0.54	61.1	94.4	100.0	100.0
L1-MP (°)	2.32	1.98	33.3	55.6	66.7	75.0
Soft tissue measurement						
Lower lip to E-plane (mm)	1.75	0.80	22.2	55.6	94.4	100.0
Upper lip to E-plane (mm)	0.95	0.73	55.6	91.7	97.2	100.0

Table 7 shows measurement bias between computer-aided and CephX analyses. According to a paired t test, 9 cephalometric measurements, SNA (°), SNB (°), SN-MP (°), ANB (°), U1-NA (mm), U1-SN (°), U1-PP (°), L1-NB (mm) and LL-EP (mm), significantly differed between computer-aided and CephX analysis.

Table 7 Comparison of the differences of cephalometric measurements between computer-aided and CephX analyses

	Mean	SD	95% limits of agreement		*P* value	Proportion within -/+ 2 units (%)
			Lower	Upper		
Skeletal measurement						
SNA (°)	-.89	1.76	-4.35	2.56	.004	75.00
SNB (°)	.34	.95	-1.52	2.21	.037	91.67
SN-MP (°)	2.45	1.65	-.78	5.68	<.001	41.67
FMA (°)	-.39	2.29	-4.88	4.09	.311	58.33
ANB (°)	-1.23	1.52	-4.21	1.74	<.001	69.44
Dental measurement						
U1-NA (mm)	2.33	2.20	-1.98	6.65	<.001	50.00
U1-SN (°)	2.93	8.55	-13.82	19.68	.047	27.78
U1-PP (°)	3.65	7.39	-10.83	18.14	.005	16.67
L1-NB (mm)	-.74	.79	-2.30	.81	<.001	94.44
L1-MP (°)	-.54	3.03	-6.48	5.40	.294	55.56
Soft tissue measurement						
LL-EP (mm)	-1.64	1.01	-3.62	0.33	<.001	55.56
UL-EP (mm)	-.24	1.19	-2.56	2.09	.235	91.67

The Bland-Altman plots represent the agreement of the results between the computer-aided and fully automated methods for skeletal (Fig. 5), dental (Fig. 6) and soft tissue (Fig. 7) measurements. For every measurement, the agreement interval fell outside the range of clinical acceptability.

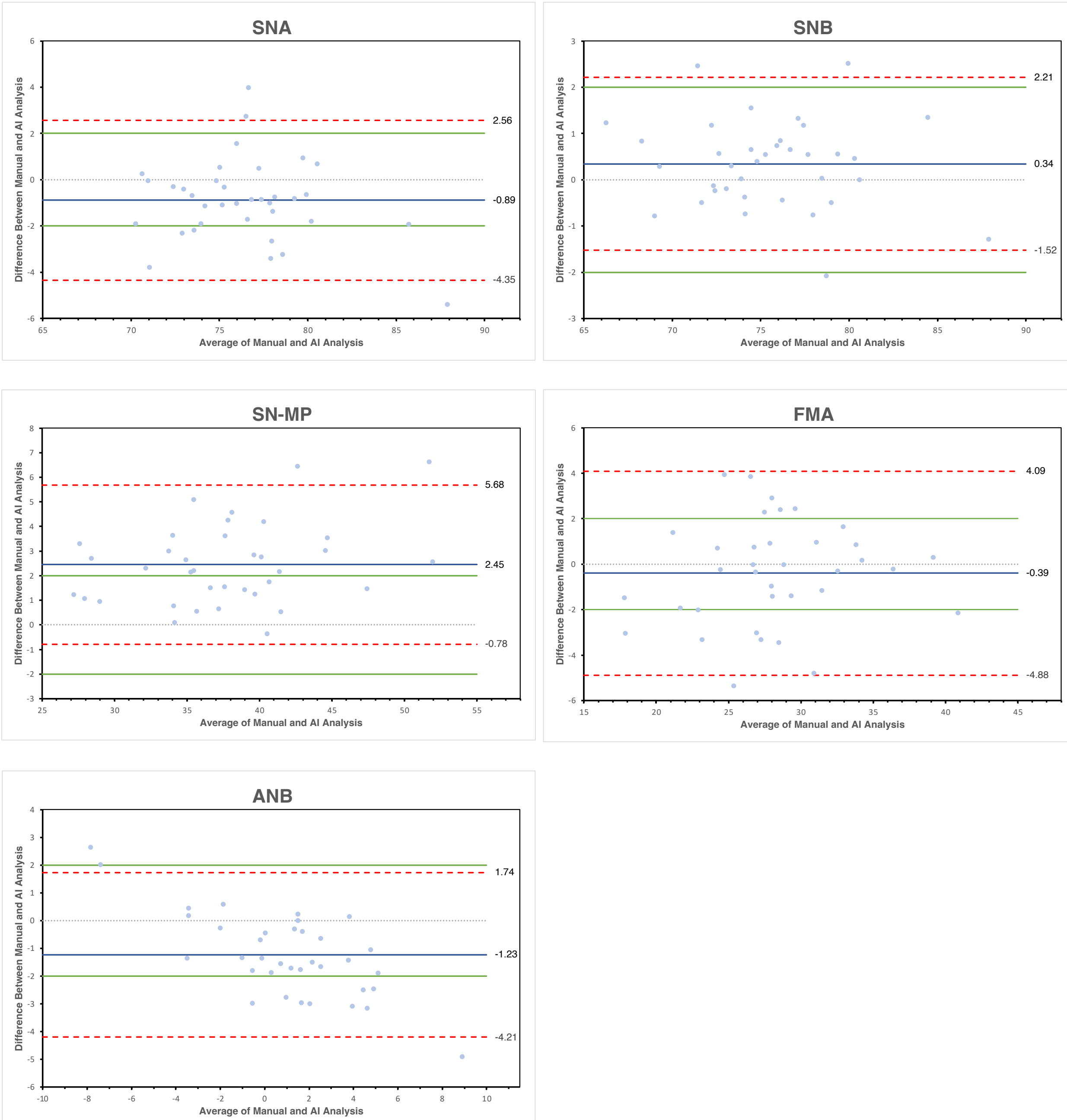


Figure 4 Bland-Altman plots of skeletal cephalometric measurements. For each plot, the x-axis represents the the average of measurements derived from CephX and clinician localizations, while the y-axis represents the difference between the measurements derived from CephX and clinician landmark identificaiton. The blue line represents the mean bias, the red dashed lines represent the upper and lower 95% limits of agreeement and the green lines represenet the upper and lower clinically acceptable limits. The scale values on the x- and y-axis vary in accordance with the mean and bias of each measurement comparison

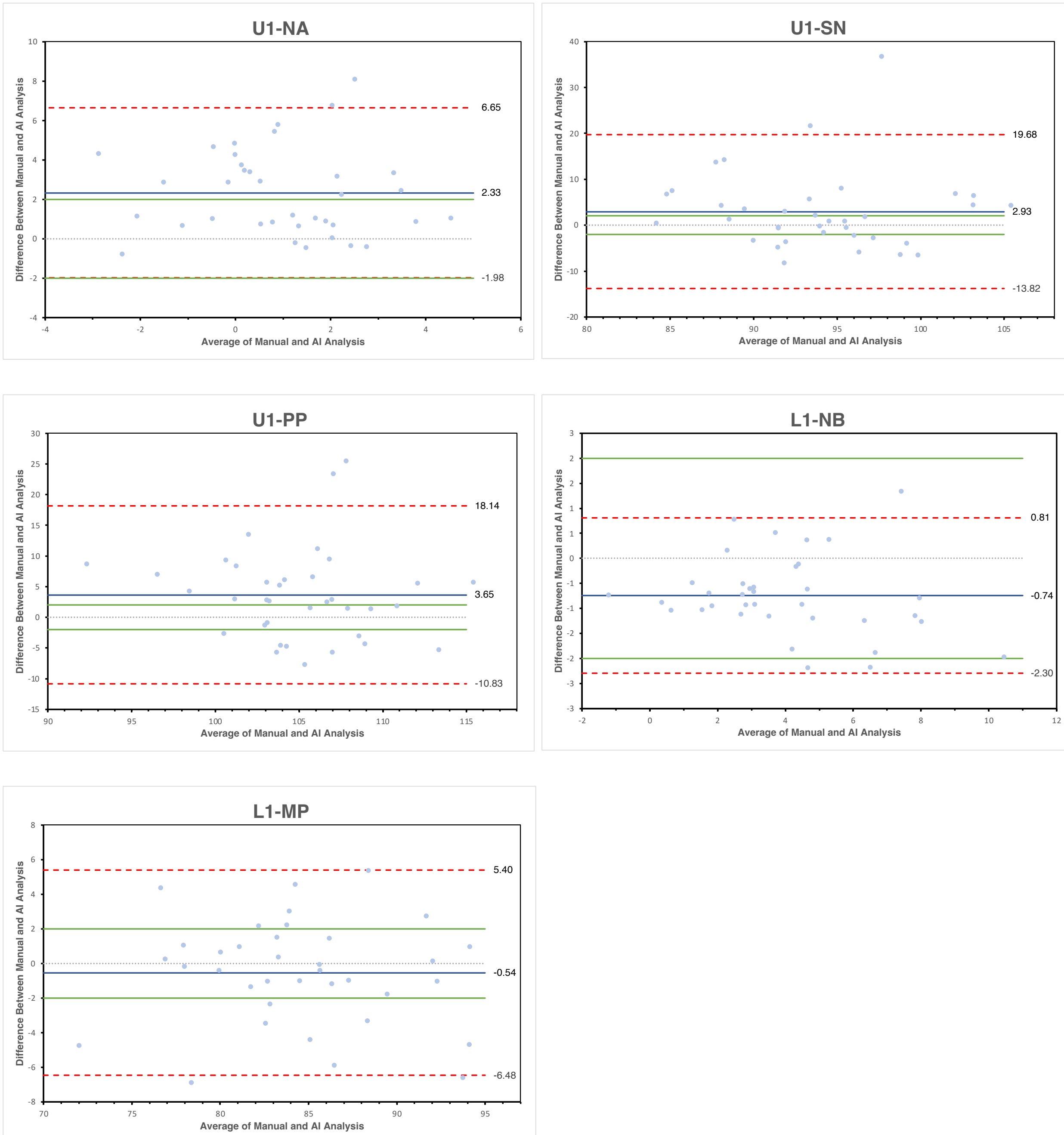


Figure 5 Bland-Altman plots of dental cephalometric measurements. For each plot, the x-axis represents the the average of measurements derived from CephX and clinician localizations, while the y-axis represents the difference between the measurements derived from CephX and clinician landmark identificaiton. The blue line represents the mean bias, the red dashed lines represent the upper and lower 95% limits of agreeement and the green lines represenet the upper and lower clinically acceptable limits. The scale values on the x- and y-axis vary in accordance with the mean and bias of each measurement comparison

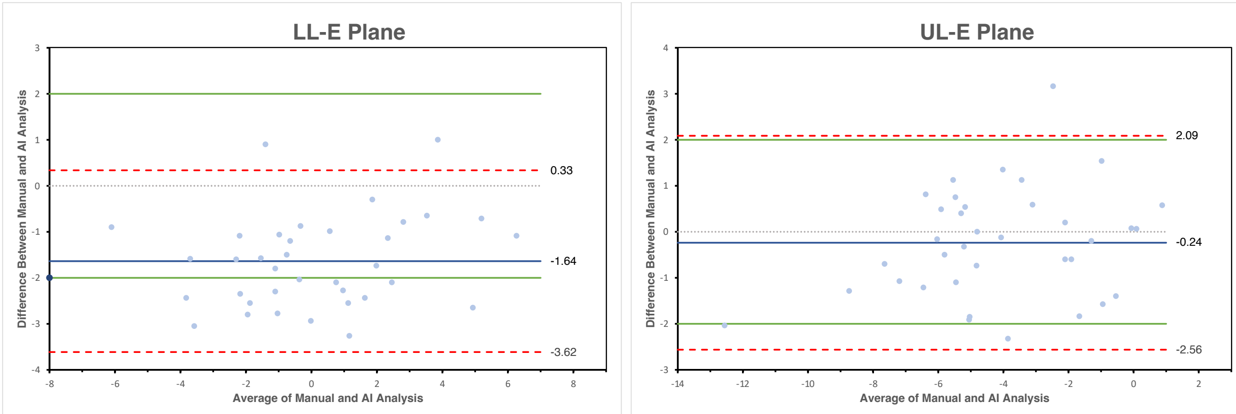


Figure 6 Bland-Altman plots of soft tisue cephalometric measurements. For each plot, the x-axis represents the the average of measurements derived from CephX and clinician localizations, while the y-axis represents the difference between the measurements derived from CephX and clinician landmark identificaiton. The blue line represents the mean bias, the red dashed lines represent the upper and lower 95% limits of agreeement and the green lines represenet the upper and lower clinically acceptable limits. The scale values on the x- and y-axis vary in accordance with the mean and bias of each measurement comparison

As shown in Table 8, the average analyzing time required for cephalometric analysis was 338 seconds for computer-aided tracing and 2236 seconds for CephX analysis.

Table 8 Comparison of analyzing time between computer-aided and CephX analyses

	Computer-aided			CephX		
	Min	**Max**	**Mean ± SD**	**Min**	**Max**	**Mean ± SD**
Time (seconds)	326	473	388 ± 40	1185	2236	1610 ± 262
Time (minutes)	5.43	7.88	6.46 ± .67	19.75	37.27	26.84 ± 4.36

DISCUSSION

Orofacial clefting is the most common human craniofacial congenital condition [3]. Oral clefts are often associated with skeletal and dental abnormalities, including anteroposterior, vertical and transverse maxillary deficiency and missing, malposed or additional teeth [7, 8]. Patients having received a diagnosis of cleft lip and palate require regular monitoring by the orthodontist of the dental and facial growth and development, allowing for the determination of the optimal time for intervention [28]. This includes obtaining CBCTs, which provide valuable information for proper orthodontic management, including the site, extent, and severity of the cleft-related craniofacial dysmorphology [29]. Lateral cephalograms can be reconstructed from CBCT volumes and cephalometric analysis, conventionally used to describe facial growth and development in cleft lip and palate patients, can be performed by manual, computer-aided and automated approaches [11, 20, 21].

In the present study, we compared computer-aided and automated analysis of lateral cephalometric images derived from cone-beam computed tomography in unilateral cleft lip and palate-affected patients. The mean absolute differences for the dental measurements (3.52 ± 3.39) were greater than skeletal (1.63 ± 1.21) and soft tissue measurements (1.35 ± 0.77). This is consistent with the finding of a study assessing RLCs processed from CBCT scans, though patients with cleft lip and palate syndromes were notably excluded [23].

When further dividing dental measurements into maxillary (4.78 ± 4.80) and mandibular (1.63 ± 1.26), maxillary dental measurements showed the highest mean error. All maxillary dental measurements involve the upper incisor which has two associated landmarks: upper incisor apex and upper incisor tip. One study that excluded patients with cleft lip and palate observed the highest mean radial error for the upper incisor apex, noting that this deviation may be due to indistinct outlines and mixed density of incisor roots, as well as obstruction by other dental roots [23]. Given that all patients in this study were in the mixed dentition, many had unerupted or partially erupted teeth, also impacting the accuracy of incisor apex identification. Another study that assessed cleft sub-groups had a lower success rate for upper incisors in comparison with the non-cleft subgroups, which was considered related to the fact that upper incisors tend to show rotation in patients with cleft lip and/or palate, which may cause difficult identification of the tip of the upper incisors [19].

All 3 maxillary dental measurements were found to be both clinically and statistically significant: U1-PP (°) (6.17±5.40), U1-SN (°) (5.72±6.94) and U1-NA (mm) (2.45±2.07). The measurement with the highest mean absolute difference was U1-PP (°), which, in addition to upper incisor apex, involves the landmarks ANS and PNS. Results from a study that compared patients with cleft lip and/or palate to non-cleft subgroups revealed relatively lower success rates for ANS and PNS in addition to upper incisor apex, associating these results with reduced radiopacity due to the cleft and the fact that ANS is not positioned in the midline due to the outward rotation of the larger segment of the maxilla [19].

The only other measurement that was found to be clinically and statistically significant was SN-MP (°) (2.47±1.62). A possible explanation for this deviation is that there are a variety of definitions that exist for mandibular plane other than the one used by the clinicians listed in Table 3, including: a plane joining gonion and gnathion; a plane joining gonion and menton; and a tangent to the lower border of the mandible and menton [30]. Several definitions also exist in the literature for the gonion landmark [18].

Cephalometric analysis is an essential diagnostic tool for analyzing orthodontic problems, maxillofacial deformities, evaluating growth and planning treatment and thus the reliability of cephalometric measurements is imperative. In patients with cleft lip and palate, growth of the maxilla is often restricted, leading to a hypoplastic maxilla and Class III malocclusion. To address this discrepancy, orthopedic correction involving the use of protraction facemask therapy at ~7-9 years of age may be considered [29]. To transmit the orthopedic force from the protraction facemask to the maxilla, intraoral devices that use the upper dentition as anchorage have been used. However, this can lead side effects, including labioversion of the upper incisors and eventual clockwise rotation of the mandible due to extrusion of the upper molars and counterclockwise rotation of the upper occlusal plane. Thus, a diagnosis of labially inclined maxillary incisors and/or a vertical facial growth pattern would be contraindications for facemask therapy with tooth-borne anchorage, and other treatment options should be considered [31]. Given that U1-PP (°), U1-SN (°) and U1-NA (mm) provide information on the position of upper incisors and SN-MP (°) provides information on facial growth pattern, inaccuracy of these measurements can lead to a compromised treatment plan.

While only 4 measurements were identified as both clinically and statistically significant, the Bland-Altman plots show that, for every measurement, limits of agreement exceeded maximum acceptable differences.

Several studies have concluded that the radiograph analyzing time with artificial intelligence was substantially shorter than the computer-assisted method. However, these studies measured the analyzing time when tracing and uploading conventional or reconstructed lateral cephalograms processed from CBCT scans into automated analysis software [13, 23, 32]. In our study, the computer-assisted analyzing time measured involved reconstruction of a lateral cephalograms from each CBCT, by human operator for the computer-assisted method and by the software for the automated method. It was found that the automated program took an average of over 4 times as long as the computer-assisted method to complete the task. However, it is important to note that the necessary tasks for computer-assisted tracing must be performed by an experienced orthodontist, while clinical experience is not required for the operator tasks involved in automated analysis.

Despite its strengths, this study also has some limitations. The recognition of cephalometric landmarks in patients with cleft lip and palate is challenging, particularly point A, ANS and PNS, and can be more heavily subject to observer bias, potentially compromising the accuracy of the referenced anatomic truth [29]. Moreover, unlike most studies that assessed landmark recognition accuracy, this study compared the accuracy of angular and linear measurements. Even if anatomical landmarks are marked incorrectly, the resulting angular and linear measurements may still be accurate, and vice versa. Thus, the ability for the software to accurately identify landmarks cannot be assessed from this study. However, it has been argued that effective clinical accuracy of orthodontic parameters is not only given by the metric deviation but also the direction of the deviation, and thus it is the resulting orthodontic parameters that should be compared [18]. Additionally, patient-oriented differences, including dental age, palate repair history and presence or absence of orthodontic appliances may impact the software's ability to recognize relevant landmarks and obtain accurate measurements. More detailed classification of image samples would improve the comprehensiveness of the conclusions. Also, in comparing the analyzing time between computer-assisted and automated methods, while the time recorded for the computer-assisted method involved operator involvement the entire time, the actual operator time was not recorded for the automated method, the majority of which was dedicated to processing the CBCT. It could be argued that the automatic program still saves operator time, increasing the efficiency of repetitive work and freeing up the orthodontist for other tasks. Future studies should be designed to compare landmark accuracy, have an increased sample size classifying image samples based on patient-oriented differences and measure time of actual operator involvement for each of the two methods.

CONCLUSION

Most cephalometric measurements evaluated in this study in unilateral cleft lip and palate patients obtained by CephX were, on average, within the clinically acceptable range as compared to those obtained by orthodontists using computer-assisted cephalometric analysis software. However, none of the limits of agreement were found to be clinically acceptable. Three of the four least reliable measurements obtained by CephX involved maxillary dental landmarks that are difficult to localize in cleft-affected patients. The other measurement found to be highly unreliable had alternative definitions. Thus, until it is developed further, especially to accommodate measurements involving difficult-to-trace landmarks and those with multiple definitions, cephalometric analysis performed by CephX in patients with unilateral cleft lip and palate should be accompanied by manual supervision and adjustment by a clinician.

REFERENCES

1. Neville BW, Al E. *Oral and Maxillofacial Pathology*. Elsevier, Inc.; 2016:1-4.

2. Sadler TW, Langman J. *Langman's Medical Embryology*. 14th ed. Wolters Kluwer; 2019:305-308.

3. Mossey P, Castilla E. *Global Registry and Database on Craniofacial Anomalies*. World Health Organization; 2003:11.

4. Vyas T, Gupta P, Kumar S, Gupta R, Gupta T, Singh H. Cleft of lip and palate: A review. *Journal of Family Medicine and Primary Care*. 2020;9(6):2621-2625.

5. Bennun RD, Harfin JF, SándorGKB, Genecov D. *Cleft Lip and Palate Management : A Comprehensive Atlas*. John Wiley & Sons; 2015.

6. Hameed O, Amin N, Haria P, Patel B, Hay N. Orthodontic burden of care for patients with a cleft lip and/or palate. *Journal of Orthodontics*. 2019;46(1):63-67.

7. Antonarakis GS, Tompson BD, Fisher DM. Preoperative Cleft Lip Measurements and Maxillary Growth in Patients with Unilateral Cleft Lip and Palate. *The Cleft Palate-Craniofacial Journal*. 2016;53(6):198-207.

8. Nicholls W. Dental anomalies in children with cleft lip and palate in Western Australia. *European Journal of Dentistry*. 2016;10(2):254-258.

9. Sharma G. Orthodontic Management of Cleft Lip and Palate Patients. In: *Current Treatment of Cleft Lip and Palate*. Intechopen; 2020.

10. Proffit WR, Fields HW, Larson BE, Sarver DM. *Contemporary Orthodontics*. 6th ed. Elsevier, Inc.; 2019.

11. Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis. *The Angle Orthodontist*. 2008;78(1):145-151.

12. Chen SK, Chen YJ, Yao CCJ, Chang HF. Enhanced speed and precision of measurement in a computer-assisted digital cephalometric analysis system. *Angle Orthodontist*. 2004;74(4):501-507.

13. Meric P, Naoumova J. Web-based Fully Automated Cephalometric Analysis: Comparisons between App-aided, Computerized, and Manual Tracings. *Turkish Journal of Orthodontics*. 2020;33(3):142-149.

14. Sayinsu K, Isik F, Trakyali G, Arun T. An evaluation of the errors in cephalometric measurements on scanned cephalometric images and conventional tracings. *European Journal of Orthodontics*. 2007;29(1):105-108.

15. Celik E, Polat-Ozsoy O, Toygar Memikoglu TU. Comparison of cephalometric measurements with digital versus conventional cephalometric analysis. *European Journal of Orthodontics*. 2009;31(3):241-246.

16. Mosleh MA, Sapiyan Baba M, Malek S, Almaktari RA. Ceph-X: development and evaluation of 2D cephalometric system. *BMC Bioinformatics*. 2016;17(Suppl 19):193-201.

17. Hutton TJ, Cunningham S, Hammond P. An evaluation of active shape models for the automatic identification of cephalometric landmarks. *European Journal of Orthodontics*. 2000;22(5):499-508.

18. Kunz F, Stellzig-Eisenhauer A, Marie Widnaier L, Zeman F, Boldt J. Assessment of the quality of different commercial providers using artificial intelligence for automated cephalometric analysis compared to human orthodontic experts. *Journal of Orofacial Orthopedics*. Published online August 29, 2023.

19. Tanikawa C, Lee C, Lim J, Oka A, Yamashiro T. Clinical applicability of automated cephalometric landmark identification: Part I—Patient-related identification errors. *Orthodontics & Craniofacial Research*. 2021;24(Suppl 2):43-52.

20. Bongaarts C, van't Hof MA, Prahl-Andersen B, Kuijpers-Jagtman AM. Identification of cephalometric landmarks in unilateral cleft lip and palate patients: are there alternatives for point A, ANS, and PNS? *Cleft Palate-Craniofacial Journal*. 2008;45(1):81-86.

21. Moshiri M, Scarfe WC, Hilgers ML, Scheetz JP, Silveira AM, Farman AG. Accuracy of linear measurements from imaging plate and lateral cephalometric images derived from cone-beam computed tomography. *American Journal of Orthodontics and Dentofacial Orthopedics*. 2007;132(4):550-560.

22. ORCA Dental AI. CEPHX. Published 2023. https://cephx.com/

23. Bao H, Zhang K, Yu C, et al. Evaluating the accuracy of automated cephalometric analysis based on artificial intelligence. *BMC Oral Health*. 2023;23(191).

24. American Board of Orthodontics. Cephalometric Review. Published 2007. https://ntiiiby.americanboardortho.com/media/t1ypvc51/202007_cephworkshop.pdf

25. American Board of Orthodontics. Construction of Mandibular Plane. https://www.americanboardortho.com/media/4omnm4kg/construction-of-the-mandibular-plane.pdf

26. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. 1994;6(4):284-290.

27. Yassir YA, Salman AR, Nabbat SA. The accuracy and reliability of WebCeph for cephalometric analysis. *Journal of Taibah University Medical Sciences*. 2022;17(1):57-66.

28. American Cleft Palate-Craniofacial Association. Parameters for evaluation and treatment of patients with cleft lip/palate or other craniofacial differences. *The Cleft Palate-Craniofacial Journal*. 2018;55(1):137-156.

29. Kocchar AS, Nucci L, Sidhu MS, et al. Reliability and reproducibility of landmark identification in unilateral cleft lip and palate patients: digital lateral vis-a-vis CBCT-derived 3D cephalograms. *Journal of Clinical Medicine*. 2021;10(3):535.

30. Jacobson A, Jacobson R. *Radiographic Cephalometry: From Basics to 3-D Imaging*. 2nd ed. Quintessence Publishing, Co, Inc; 2006:66.

31. Baek SH, Kim KW, Choi JY. New treatment modality for maxillary hypoplasia in cleft patients. *The Angle Orthodontist*. 2010;80(4):783-791.

32. Duran GS, Gökmen Ş, Topsakal KG, Görgülü S. Evaluation of the accuracy of fully automatic cephalometric analysis software with artificial intelligence algorithm. *Orthodontics & Craniofacial Research*. 2023;26(3):481-490.