Automatic cephalometric analysis: is it time to switch to a hands-free method?

Mariane Michels[1], Rodrigo Leitão de Assunção[2], Matheus Lima Oliveira[1], Marcela Rodrigues Alves[3], Maria Augusta Portella Guedes Visconti[4], Fábio Ribeiro Guedes[4]

[1] Department of Oral Diagnosis, Piracicaba Dental School, State University of Campinas, Piracicaba, SP, Brazil.

[2] Department of Orthodontics, Faculty of Dentistry, Federal Fluminense University, Niterói, RJ, Brazil.

[3] Department of Prosthodontics and Dental Materials, Faculty of Dentistry, Federal University of Rio de Janeiro, RJ, Brazil.

[4] Department of Pathology and Oral Diagnosis, Faculty of Dentistry, Federal University of Rio de Janeiro, RJ, Brazil.

Conflicts of interest: none.

Corresponding author:
Mariane Michels
Piracicaba Dental School - University of Campinas - Division of Oral Radiology
Limeira Avenue, 901
Zip Code: 13414-903
Piracicaba, SP, Brazil
Phone: (+55) 19 21065327
E-mail: marianeradiodonto@gmail.com

Aim: To evaluate the reliability of the automatic cephalometric analysis in relation to the semi-automatic method. **Methods:** Fifty lateral cephalometric radiographs were selected and two dental surgeons performed the Steiner and Tweed analyses independently using the semi-automatic method on the Radiocef Studio 2® software suite (Radiomemory, Belo Horizonte, MG, Brazil), and the automatic method on the Kodak Dental Imaging Software (Carestream Health, Rochester, NY, USA). After thirty days, 30% of the sample was re-evaluated to assess intra-observer agreement. Ten angular and linear measurements of both analyses were selected, averaged for both observers and compared using Student's t-test with a significance level of 5% (α=0.05). Intra and inter-observer agreement were assessed through Intraclass Correlation Coefficient. **Results:** Intra-observer reproducibility was excellent for all measurements and inter-observer reproducibility was excellent for most of them. Significant differences (p<0.05) were found between automatic and semi-automatic methods for all measurements. Most of the measurements were significantly higher (p<0.05) with the automatic method. **Conclusion:** Semi-automatic cephalometric analysis can not be replaced with a completely automatic method.

Keywords: Cephalometry. Software. Benchmarking.

Introduction

Lateral cephalometric radiography is a diagnostic imaging method routinely used to study craniofacial growth changes of an individual. This technique helps the professional to determine the diagnosis, treatment plan and prognosis of a patient based on a cephalometric analysis and, together with physical examination, represents an essential condition for the success of orthodontic and/or surgical treatment[1].

A reliable cephalometric analysis is produced after the identification of anatomical landmarks, which must be recognized and marked accurately on the cephalometric radiograph. Commonly, using a specific software, an observer subjectively indicates the landmarks and a tracing is automatically produced (semi-automatic method); however, considering the inherent superimposition of structures in radiographic exams, some difficulties are faced for the precise identification[2]. Inaccurate identification of such landmarks may lead to misinterpretation and inadequate treatment[3].

Software applications have been developed with the purpose of identifying cephalometric landmarks automatically using a built-in tool to identify cephalometric landmarks, generate cephalograms, and carry out observer-independent analyses (automatic method)[4]. Among the advantages of automatic method in relation to the semi-automatic method, the reduction of the working time and the expectation for more precise cephalometric tracing are the most important. However, several automatic software applications available on the market have demonstrated to be unable to accurately determine cephalometric landmarks, which could affect orthodontic planning when combined with clinical examination and model analysis[5]. Thus, the aim of this study was to evaluate the reliability of the automatic cephalometric analysis in relation to the semi-automatic method.

Materials and methods

Ethical aspects and selection criteria

The present study was performed after the approval of the local Research Ethics Committee (protocol# 949.939). An informed consent form was obtained from all the participants of the study.

Fifty lateral cephalometric radiographs were selected from patients of both sexes, aged between 18 to 30 years, who had permanent dentition. Exclusion criteria included the absence of any tooth, except for third molars, assessed through panoramic radiography.

Image acquisiton

The same technician obtained the lateral cephalometric radiographs using the Kodak 9000C 3D digital extraoral imaging system (Carestream Health, Rochester, NY, USA) with an accessory arm for cephalometric image acquisition. All patients were in maximum habitual intercuspation, with the Frankfurt horizontal plane parallel to the horizontal plane and the midsagittal plane perpendicular to the horizontal plane, con-

firmed with the concentric superimposition of the olives. The X-ray exposure settings were adjusted based on the patient body type but could range between 78−82 kVp and 10−12 mA for 0.5 seconds. The images obtained were exported and stored digitally as TIFF files at a resolution of 200 dpi.

Image assessment

Two dental surgeons with experience in cephalometric analysis (an oral radiologist and an orthodontist) performed the Steiner and Tweed analyses independently using the semi-automatic method on the Radiocef Studio 2® software suite (Radiomemory, Belo Horizonte, MG, Brazil), and the automatic method on the Kodak Dental Imaging Software (KDIS® - Carestream Health, Rochester, NY, USA), which is the native software of the imaging system. It is worth mentioning that the observers were not allowed to subjectively correct the location of the landmarks automatically determined on the KDIS software. After thirty days, 30% of the sample was re-evaluated to assess intra-observer agreement.

Eleven cephalometric landmarks were selected from the Steiner and Tweed analyses, as follows: Sella (S), Nasion (N), A-point or subspinale (A), B-point or supramentale (B), Pogonion (Pg), Gonion (Go), Gnathion (Gn), Porion (PoR), Orbitale (Or), Incisor superior (Is) and Incisor inferior (Ii). Then, ten angular and linear measurements were listed to compare the automatic and semi-automatic methods, as follows: SNA, SNB, 1-NA, 1-NB, 1.NB, Pg-NB, SN.GoGn, FMA, FMIA, IMPA. These measurements were selected for being the most frequent among cephalometric analyzes. Figure 1 illustrates both the semi-automatic (Fig. 1A) and automatic (Fig. 1B) cephalometric analyses. Attention should be given to the superimposed tracings over the structures between methods as, for instance, the base of the mandible, which directly affects the Gonion landmark.

Statistical analysis

The angular and linear measurements obtained from both landmark selection methods were tabulated and compared using Student's t-test. Intra and inter-observer agreement was assessed through Intraclass Correlation Coefficient (ICC). All statistical analyzes were performed using the BioEstat 5.0 statistical software (Belem, Para, Brazil) with a significance level of 5% (α=0.05).

RESULTS

ICC was interpreted according to the Cichetti[6] categorization system where <0.40 is a poor agreement, 0.40−0.59 a fair agreement, 0.60−0.74 a good agreement, and 0.75−1 an excellent agreement. The intra-observer reproducibility ranged from good to excellent for both observers, except for the 1-NA and Pg-NB, which revealed poor reproducibility, as exposed in table 1. Regarding the inter-observer reproducibility, it was excellent for most of the measurements, except for the 1-NA and Pg-NB, as shown in table 2.

Table 3 shows significant differences (p<0.05) that were found between automatic and semi-automatic methods for all measurements. Most of the measurements were sig-

Table 1. Intraclass Correlation Coefficient (ICC) for intra-examiner agreement and reproducibility level

Measurements	ICC Examiner 1	Reproducibility Level	ICC Examiner 2	Reproducibility Level
SNA (°)	0.8748	Excellent	0.8847	Excellent
SNB (°)	0.9498	Excellent	0.9524	Excellent
1.NB (°)	0.8117	Excellent	0.8306	Excellent
SN-GoGn (°)	0.9185	Excellent	0.9884	Excellent
FMA (°)	0.9002	Excellent	0.7236	Good
FMIA (°)	0.8898	Excellent	0.6683	Good
IMPA (°)	0.8470	Excellent	0.8131	Excellent
1-NA (mm)	0.2575	Poor	0.7194	Good
1-NB (mm)	0.8866	Excellent	0.8132	Excellent
Pg-NB (mm)	0.3799	Poor	0.7481	Good

Table 2. Intraclass Correlation Coefficient (ICC) for inter-examiner agreement and reproducibility level

Measurements	ICC	Reproducibility Level
SNA (°)	0.8509	Excellent
SNB (°)	0.8741	Excellent
1.NB (°)	0.8395	Excellent
SN-GoGn (°)	0.8133	Excellent
FMA (°)	0.9111	Excellent
FMIA (°)	0.8383	Excellent
IMPA (°)	0.806	Excellent
1-NA (mm)	0.3633	Poor
1-NB (mm)	0.8885	Excellent
Pg-NB (mm)	0.2562	Poor

Table 3. Mean values (standard deviation) of the angular (in degree) and linear (in millimeter) measurements using the automatic and semi-automatic methods, followed by difference of means and p-value.

Measurements	Tracing Method		Difference of means	p-value
	Automatic	Semi-automatic		
SNA (º)	82.57 (3.89)	80.82 (3.34)	1.76	0.0007*
SNB(º)	78.65 (4.70)	77.39 (3.50)	1.30	0.0262*
1.NB (º)	29.64 (6.78)	27.73 (7.05)	1.91	0.0110*
SN.GoGn (º)	32.04 (4.54)	33.60 (6.77)	1.56	0.0071*
FMA (º)	28.76 (4.92)	24.52 (6.04)	4.24	<0.0001*
FMIA (º)	53.94 (8.80)	61.15 (8.55)	7.21	<0.0001*
IMPA (º)	96.32 (6.66)	94.33 (7.42)	1.98	0.0270*
1-NA (mm)	4.28 (2.24)	5.87 (1.95)	1.60	<0.0001*
1-NB (mm)	5.72 (2.60)	6.37 (2.70)	0.64	0.0008*
Pg-NB (mm)	1.54 (0.85)	1.20 (1.42)	0.35	0.0069*

*$p<0.05$

nificantly higher (p<0.05) with the automatic method in relation to the semi-automatic method, except for 1-NA, 1-NB, SN.GoGn and FMIA.

DISCUSSION

Considering the importance of radiographic-based cephalometric analysis in the diagnostic process and treatment plan in orthodontics, this study proposed to compare two different tracing methods: automatic and semi-automatic. The most frequent angular and linear measurements used in cephalometric analyses were obtained from Steiner and Tweed analyses and, interestingly, significant statistical differences were found in all measurements for both tracing methods, which demonstrates that automatic and semi-automatic methods led to different outcomes. Previous scientific studies on this topic and with similar objectives to the present study revealed the lack of criteria regarding the choice of the ideal cephalometric landmark to be tested[7-12]. In the current methodology, most of the chosen landmarks were of easy locating to avoid the bias of possible interference of inaccurate values when comparing both automatic and semi-automatic methods.

Some authors state that the automatic method presents good reliability because it behaved similarly to the semi-automatic mode performed by two dentists; however, it was not disclosed on those studies whether the observers could subjectively adjust the location of the landmarks generated using the automatic method[2,5,13]. The current methodology did not include any possibility of subjective adjustment in the automatic tracing method for considering that this would exclude the benefit it presents in relation to the semi-automatic method. This can justify all the significant differences we found between both methods. It is important to emphasize that the automatic method allows the observer to have a good estimation of the position of the anatomical landmarks necessary for radiographic-based cephalometric analysis, which can be refined by the clinician, possibly reducing working time and improving the reproducibility[1,2,5,14]. However, this statement is controversial, since, for some professionals, adjusting the location of automatic-identified landmarks can be more laborious than marking it correctly once using the semi-automatic method. Figure 1B clearly shows that the Gonion (the most inferior and posterior point on the mandibular angle) is not in its most appropriate location.

Otherwise, this study revealed good and excellent intra-observer agreement for most of the measurements according to Cichette[6], which demonstrates that the observers produced reliable measurements, even of some landmarks of intermediate complexity of identification, such as nasion, porion, orbitale, and gonion[3,15]. The inter-observer evaluation showed that most of the measurements presented an excellent correlation, except for the 1.NA and Pg-NB measurements. Nevertheless, out of a total of ten measurements, only two of them did not show good correlation. This can possibly indicate that the location of the N (nasion) landmark is more difficult[15].

Ferreira e Telles[15] demonstrated that the angular measurement SN.GoGn did not change significantly between manual and semi-automatic methods, even considering that it makes use of the N (nasion) landmark, which has a difficult location according to the authors. The same measurement presented significant difference in the present
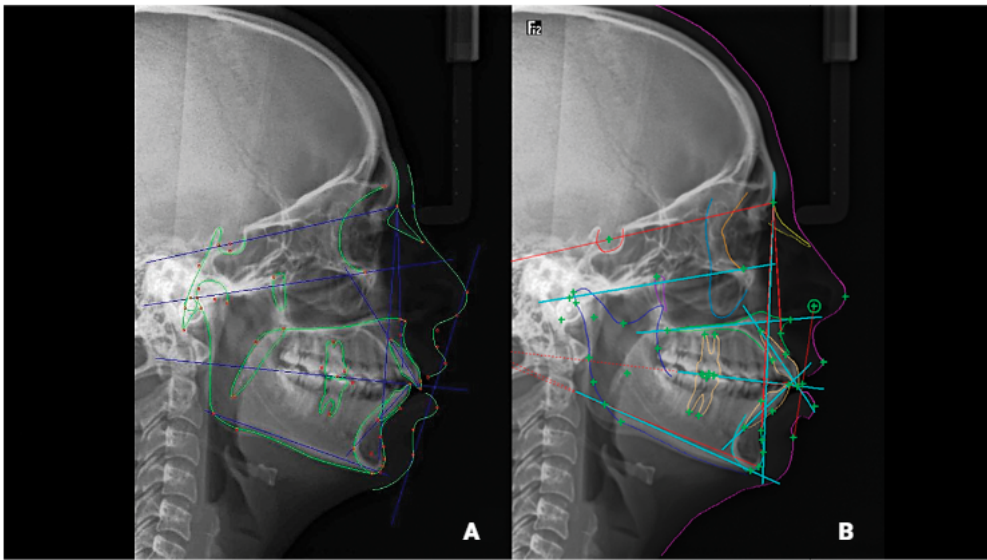
Figure 1. Radiographic-based cephalometric analysis: A, semi-automatic method; B, automatic method.

study. This can be possibly explained by the fact that in the first study, the researchers compared the manual method with the semi-automatic method, both performed by the same observers and, in this study, the observers could only directly interfere with the semi-automatic method. In the study conducted by Codari et al.[2], the Go (Gonion) landmark was considered to be of hard identification and this was attributed to a possibly greater anatomical variability of the mandibular angle.

The Frankfurt horizontal plane, used to obtain the FMA and FMIA measurements, originates from the upper margin of the auditory meatus (Porio) to the inferior border of the orbit (Orbital), which, on the lateral cephalometric radiograph, has the disadvantage of being more vulnerable to distortions and facial asymmetries[3,16]. In the study of Hutton et al.[17], the Porio landmark was located with reduced accuracy. This information can explain the greatest difference of means observed for FMA and FMIA between both tracing methods (Table 3).

Observing the difference of means found for SNA, SNB, 1-NA, 1.NB, FMA, and FMIA measurements, it can be noted that the cephalometric method can have serious clinical implications[16]. For instance, the value found for FMA using the automatic and semi-automatic methods were 28.76° and 24.52°, respectively. This variation could clearly alter the orthodontic treatment plan, especially with respect to the evaluation of the total discrepancy that combines the cephalometric discrepancy to the model discrepancy, and helps the orthodontist to decide whether tooth extraction is needed or not. It is worth mentioning that for this therapeutic decision, in addition to obtaining the total discrepancy resulting from the sum of the cephalometric discrepancy and the model discrepancy, there is a need for an evaluation of the facial profile and type, stability of the triangle of Tweed, and also the depth of the Spee's curve[18].

It is well known that diagnosis in orthodontics for appropriate elaboration of treatment plan is composed of anamnesis, clinical examination, radiographic and pho-

tographic assessment, and study of models. Thus, any technological approach that can be developed and employed to facilitate and systematize this process is extremely relevant.

Currently, some studies have developed CBCT software applications for cephalometric analysis because this technique presents images without overlaps, distortions and magnifications[19,20]. In 2018, Neelapu et al.[19] revealed that CBCT-based cephalometric analysis is reliable and with an acceptable mean error. In the same year, Montúfar et al.[20] demonstrated high accuracy of new software for 3D automatic cephalometric analysis. They further stated that this was a significant advance from 2D analysis. However, since 2009, Chien et al.[3] concluded that both 2D and 3D methods are clinically comparable. Therefore, considering radiological protection aspects for both adult and pediatric patients, the most appropriate technique is the 2D[21-23].

Considering the limitations of the current methodology, semi-automatic cephalometric analysis cannot be replaced with a completely automatic method. Further studies to assess the impact of subjective adjustment of automatic methods are recommended.

REFERENCES

1. Shahidi S, Shahidi S, Oshagh M, Gozin F, Salehi P, Danaei SM. Accuracy of computerized automatic identification of cephalometric landmarks by a designed software. Dentomaxillofac Radiol. 2013;42(1):20110187. doi: 10.1259/dmfr.20110187.

2. Codari M, Caffini M, Tartaglia GM, Sforza C, Baselli G. Computer-aided cephalometric landmark annotation for CBCT data. Int J Comput Assist Radiol Surg. 2017 Jan;12(1):113-121. doi: 10.1007/s11548-016-1453-9.

3. Chien PC, Parks ET, Eraso F, Hartsfield JK, Roberts WE, Ofner S. Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography in vivo. Dentomaxillofac Radiol. 2009 Jul;38(5):262-73. doi: 10.1259/dmfr/81889955.

4. Kafieh R, Sadri S, Mehri A, Raji H. Discrimination of bony structures in cephalograms for automatic landmark detection. Commun Comput Inf Sci. 2008;6 CCIS:609-20. doi: 10.1007/978-3-540-89985-3_75.

5. Lindner C, Wang CW, Huang CT, Li CH, Chang SW, Cootes TF. Fully Automatic System for Accurate Localisation and Analysis of Cephalometric Landmarks in Lateral Cephalograms. S Sci Rep. 2016 Sep 20;6:33581. doi: 10.1038/srep33581.

6. Cicchetti D V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instrument in psychology. Psychol Assess. 1994 Dec;6(4):284-90. doi: 10.1037/1040-3590.6.4.284.

7. Chen Y-J, Chen S-K, Yao JC-C, Chang H-F. The effects of differences in landmark identification on the cephalometric measurements in traditional versus digitized cephalometry. Angle Orthod . 2004 Apr;74(2):155–61.

8. Chen S-K, Chen Y-J, Yao C-CJ, Chang H-F. Enhanced speed and precision of measurement in a computer-assisted digital cephalometric analysis system. Angle Orthod. 2004 Aug;74(4):501-7.

9. Davis DN, Mackay F. Reliability of cephalometric analysis using manual and interactive computer methods. Br J Orthod. 1991 May;18(2):105-9.

10. Dana JM, Goldstein M, Burch JG, Hardigan PC. Comparative study of manual and computerized cephalometric analyses. J Clin Orthod. 2004 May;38(5):293-6.

11. Nimkarn Y, Miles PG. Reliability of computer-generated cephalometrics. Int J Adult Orthodon Orthognath Surg. 1995;10(1):43-52.

12. Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmarks identification and reproducibility: A Meta analysis. Am J Orthod Dentofac Orthop. 1997 Aug;112(2):165-70.

13. Mosleh MA, Baba MS, Malek S, Almaktari RA. Ceph-X: development and evaluation of 2D cephalometric system. BMC Bioinformatics. 2016 Dec 22;17(Suppl 19):499. doi: 10.1186/s12859-016-1370-5.

14. Mosleh MAA, Baba MS, Malek S, Almaktari RA. Ceph-X: development and evaluation of 2D cephalometric system. BMC Bioinformatics. 2016 Dec 22;17(Suppl 19):499. doi: 10.1186/s12859-016-1370-5.

15. Ferreira JTL, Telles C de S. Evaluation of the reliability of computerized profile cephalometric analysis. Braz Dent J . 2002;13(3):2014.

16. Ahmed M, Shaikh A, Fida M. Diagnostic performance of various cephalometric parameters for the assessment of vertical growth pattern. Dental Press J Orthod. 2016 Jul-Aug;21(4):41-9. doi: 10.1590/2177-6709.21.4.041-049.oar.

17. Hutton TJ, Cunningham S, Hammond P. An evaluation of active shape models for the automatic identification of cephalometric landmarks. Eur J Orthod . 2000 Oct;22(5):499-508.

18. Li Q, Zheng Z, Bai D, Pang G. [A retrospective study of morphologic basis for the extraction decision in Class II, division 1 malocclusion]. Hua Xi Kou Qiang Yi Xue Za Zhi. 1999 Nov;17(4):341-3. Chinese.

19. Neelapu BC, Kharbanda OP, Sardana V, Gupta A, Vasamsetti S, Balachandran R, et al. Automatic localization of three-dimensional cephalometric landmarks on CBCT images by extracting symmetry features of the skull. Dentomaxillofac Radiol. 2018 Feb;47(2):20170054. doi: 10.1259/dmfr.20170054.

20. Montúfar J, Romero M, Scougall-Vilchis RJ. Automatic 3-dimensional cephalometric landmarking based on active shape models in related projections. Am J Orthod Dentofacial Orthop. 2018 Mar;153(3):449-458. doi: 10.1016/j.ajodo.2017.06.028.

21. Hedesiu M, Marcu M, Salmon B, Pauwels R, Oenning AC, Almasan O, et al. Irradiation provided by dental radiological procedures in a pediatric population. Eur J Radiol. 2018 Jun;103:112-117. doi: 10.1016/j.ejrad.2018.04.021.

22. Abdelkarim AA. Appropriate use of ionizing radiation in orthodontic practice and research. Am J Orthod Dentofacial Orthop. 2015 Feb;147(2):166-8. doi: 10.1016/j.ajodo.2014.11.010.

23. European Commission. Radiation Protection n.172. Cone beam CT for dental and maxillofacial radiology (Evidence-based guidelines). Luxembourg; 2012 [cited 2018 Jul 15]. Available from: http://www.sedentexct.eu/files/radiation_protection_172.pdf.