

Reproducibility of Cephalometric Measurements Made by Three Radiology Clinics

Heraldo Luis Dias da Silveira^a; Heloisa Emilia Dias Silveira^b

Abstract: The purpose of this study was to assess reproducibility of cephalometric measurements in cephalograms obtained by three dentomaxillofacial radiology clinics. Forty lateral cephalometric radiographs were selected and sent at different times to three different clinics for cephalometric analyses. Each clinic digitized the radiographs with the same resolution, and landmarks were located with the mouse pointer directly on the digitized radiographic image on the screen. Three cephalograms were obtained from each radiograph, totaling 120 analyses. Data were analyzed with analysis of variance. Of the 32 factors studied, reproducibility of results was satisfactory for only four factors: position of maxilla relative to anterior cranial base, inclination of occlusal plane relative to anterior cranial base, position of lower incisor relative to nasion-pogonion line, and soft-tissue profile of face ($P < .05$). Differences in cephalometric measurements were present and such differences were significant for most factors analyzed. The different cephalometric measurements obtained by the three dental radiology clinics were not reproducible. (*Angle Orthod* 2006;76:394–399.)

Key Words: Cephalometry; Reproducibility of results; Landmark identification; Cephalometric analysis

INTRODUCTION

Cephalometric analysis is based on the identification of anatomic landmarks, some of which are difficult to identify. Therefore, some landmarks are more reproducible than others, and absolute accuracy can hardly be achieved because all identifications are subject to some degree of error. The difficulty in identifying cephalometric landmarks is associated with the fact that the images of anatomical structures overlap and that some landmarks are paired with one found on each side of the face. Consequently, they often appear as double, noncoinciding images on lateral radio-

graphs.^{1–4} However, the value of cephalometric analyses depends on the accuracy of measurement techniques because errors in recordings may lead to an incorrect diagnosis.^{5–8}

Dentomaxillofacial radiology clinics in Brazil usually trace the cephalograms and send them, together with the corresponding radiographs, to the orthodontists or oral and maxillofacial surgeons who will use them to establish diagnoses and plan therapeutic interventions. Several studies in the literature have shown the existence of variability in interobserver landmark identification depending on different individual conceptions of landmark definition and perception of landmarks.^{9–11} According to Trpkova et al¹² and Rudolph et al,⁵ the validity of any measurement obtained from cephalometric radiographs depends on the reproducibility of cephalometric landmarks. The quality of the radiograph, the conditions under which measurements are made, and the skill and training of the person who traces the cephalograms affect the magnitude of error in landmark identification.

Lau et al,¹⁰ however, argued that the degree of error probably depends on individual conceptions of landmark definition and perception of landmark location rather than on training or experience. Kamoen et al¹³ studied the reproducibility of interobserver cephalo-

^a Student—Graduate (Ph.D.), Departamento de Cirurgia e Ortopedia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

^b Professor, Departamento de Cirurgia e Ortopedia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

Corresponding author: Dr. Heraldo Luis Dias da Silveira, Departamento de Cirurgia e Ortopedia, Universidade Federal do Rio Grande do Sul, Riva Ramiro Barcelos, 2492, Porto Alegre, Rio Grande do Sul, Brazil 90035-003 (e-mail: heraldods@ig.com.br)

Accepted: June 2005. Submitted: April 2005.

© 2006 by The EH Angle Education and Research Foundation, Inc.

TABLE 1. Description of 32 Cephalometric Factors Analyzed on Cephalograms^a

Cephalometric Factors	Description
S-N.A	Position of maxilla relative to anterior cranial base.
S-N.Occl	Inclination of occlusal plane relative to anterior cranial base.
/1-NPog	Position of lower incisor relative to nasion-pogonion line.
H.(N-B)	Soft-tissue profile of face.
S-N.D	Position of mandible relative to anterior cranial base.
(S-N).(Go-Me)	Inclination of mandibular plane relative to anterior cranial base.
1/.NS	Inclination of upper incisor relative to anterior cranial base.
1/.NA	Inclination of upper incisor relative to maxilla.
FMA	Frankfort mandibular incisor angle.
FMA	Frankfort mandibular angle.
A-(V-T)	Position of maxilla relative to V-T line.
CD (Vigorito)	Vogorito's cephalometric discrepancy.
1/./1	Interincisor angle.
/1-Orb	Distance from long axis of upper incisor to orbitale.
H-Nose	Nose projection relative to Holdaway line.
IMPA	Incisor mandibular plane angle.
(N-Pog).(Po-Orb)	Facial angle.
N-A.Pog	Angle of convexity.
S-N.B	Position of mandible relative to anterior cranial base.
A-N.B	Relative position of maxilla to mandible.
S-N.Gn	Growth axis.
1/.NA	Position of upper incisor relative to maxilla.
/1.NB	Inclination of lower incisor relative to mandible.
(Go-Me).(V-T)	Inclination of mandibular plane relative to V-T line.
F.(V-T)	Inclination of V-T line relative to Frankfort plane.
(Go-Gn).Occl	Inclination of occlusal plane relative to mandibular plane.
Genial tubercle	Projection of genial tubercle.
/1-(V-T)	Inclination of lower incisor relative to V-T line.
H.(V-T)	Angle formed by Holdaway line and V-T line.
/1-NB	Position of lower incisor relative to mandible.
/1-Line I	Position of lower incisor relative to line I.
Pog-NB	Distance from pogonion to NB line.

^a V-T line indicates Long axis of mandibular symphysis Line 1, Interlandi's line.

metric tracings and concluded that there were significant differences in the identification of anatomic landmarks. They also found that tracing accuracy is a limiting factor in cephalometry and that the variation for each landmark is dependent on the quality of the cephalogram.

The purpose of the cross-sectional study presented here was to study the reproducibility of cephalometric measurements obtained by three different dentomaxillofacial radiology clinics.

MATERIALS AND METHODS

Forty radiographs were selected from patient files; selection criterion was the technical quality of radiographs. All radiographs were taken with the same x-ray device (Orthophos 3 Ceph/60–80 kV, 10 mA). Three clinics (A, B, and C) located in the city of Porto Alegre, Brazil, were randomly selected for the analyses.

Each clinic digitized the radiographic image and analyzed one lateral cephalogram for each of the 40 radiographs, and a total of 120 cephalograms was ob-

tained for this study. Each cephalogram had 32 cephalometric measurements (Table 1) corresponding to the analysis of one radiograph, and 3840 measurements were analyzed.

The Radiocef system (Radio Memory Ltda, Belo Horizonte, Brazil) was used to prepare the cephalograms. This system uses an A4 desktop scanner with a transparency adapter to digitize the radiographic images. The digital image resolution was 150 dpi, 8 bits. Landmarks were located with the mouse pointer directly on the digitized radiographic image on the screen.

To ensure that the analyses were performed as usual by the clinics, observers were blinded to the fact that analyses were part of a study. Therefore, a sample of what routinely happens in those clinics was obtained. No criterion standard was established for the accuracy of cephalograms prepared by the different clinics. The purpose of this study was to determine the reproducibility of the results and not to judge which clinic was right or wrong.

Analysis of variance (ANOVA) for three or more paired samples was used to compare each separate

TABLE 2. Results of ANOVA Performed to Compare Cephalometric Measurements Obtained by Clinics A, B, and C in Porto Alegre, Brazil^a

Factors	F ratio	P value	Conclusion	Differences
S-N.A	0.017	.972	NS	No differences between A, B, and C
S-N.Occl	2.573	.083	NS	No differences between A, B, and C
/1-NPog	3.076	.052	NS	No differences between A, B, and C
H.(N-B)	1.947	.156	NS	No differences between A, B, and C
S-N.D	17.844	.000	S	Differences between A, B, and C
(S-N).(Go-Me)	28.030	.000	S	Differences between A, B, and C
1/.NS	25.329	.000	S	Differences between A, B, and C
1/.NA	25.147	.000	S	Differences between A, B, and C
FMIA	37.018	.000	S	Differences between A, B, and C
FMA	25.431	.000	S	Differences between A, B, and C
A-(V-T)	24.538	.000	S	Differences between A, B, and C
CD (Vigorito)	11.243	.001	S	Differences between A, B, and C
1/.1	28.182	.000	S	A different from B and C
/1-Orb	28.826	.000	S	A different from B and C
H-Nose	7.175	.001	S	A different from B and C
IMPA	29.485	.000	S	A different from B and C
(N-Pog).(Po-Orb)	14.043	.000	S	B different from A and C
N-A.Pog	10.305	.000	S	B different from A and C
S-N.B	9.419	.001	S	B different from A and C
A-N.B	12.969	.000	S	B different from A and C
S-N.Gn	12.019	.000	S	B different from A and C
1/-NA	11.964	.000	S	B different from A and C
/1.NB	18.743	.000	S	B different from A and C
(Go-Me).(V-T)	14.751	.000	S	B different from A and C
F.(V-T)	19.778	.000	S	B different from A and C
(Go-Gn).Occl	10.410	.000	S	C different from A and B
Genial tubercle	23.540	.000	S	C different from A and B
lii-(V-T)	8.726	.000	S	C different from A and B
H.(V-T)	23.254	.000	S	C different from A and B
/1-NB	4.810	.011	S	A different from B
/1-Line I	7.429	.002	S	A different from B
Pog-NB	5.153	.008	S	A different from C

^a NS indicates not significant; S, significant.

factor for the three clinics, according to the SPSS:GLM routine (general linear model repeated measures). For the multiple comparison procedures, a significance level of 5% was established.

RESULTS

The results of ANOVA analyses for each factor under study are shown in Table 2. The most important results are described below.

- Only four factors did not show any differences between the three clinics: S-N.A; S-N.Occl; /1-NPog; H.(N-B).
- A significant difference was found between the three clinics (A and B, A and C, B and C) in eight factors: S-N.D; (S-N).(Go-Me); 1/.NS; 1/.NA; FMIA; FMA; A-(V-T); CD (Vigorito cephalometric discrepancy).
- A significant difference was found between clinics (A and B, A and C) in four factors: 1/.1; /1-Orb; H-Nose; IMPA.
- A significant difference was found between clinics (A and B, B and C) in nine factors: (N-Pog).(Po-Orb);

N-A.Pog; S-N.B; A-N.B; S-N.Gn; 1/-NA; /1.NB; (Go-Me).(V-T); F.(V-T).

- There was a significant difference between clinics (A and C, B and C) in four factors: (Go-Gn).Occl; genial tubercle; lii-(V-T); H.(V-T).
- A significant difference was found between clinics A and B in two factors: /1-NB; /1-Line I.
- Results for the Pog-NB factor were significantly different between clinics A and C.

The maximum differences between the clinics were, in general, very high (Tables 3 through 5). Differences of up to 11.33 mm were found for linear measurements. Differences of up to 20.13° were found in angular measurements for the FMIA factor. The greatest difference found for Vigorito's CD was 17.70 mm (Table 4).

DISCUSSION

In computer-aided cephalometry, which uses digitized radiographs and specific software for tracings and measurements, the chance of error lies in locating

TABLE 3. Mean, Standard Deviation, and Minimum and Maximum Differences Between Clinics A and B for Cephalometric Measurements (in Millimeters or Degrees), Porto Alegre, Brazil

Variable	Differences Between Clinics			
	Mean	SD	Minimum	Maximum
(N-Pog).(Po-Orb)	3.01	2.21	0.43	10.22
N-A.Pog	2.21	2.68	0.09	16.02
S-N.A	1.51	1.50	0.00	6.93
S-N.B	1.31	1.12	0.04	5.03
A-N.B	0.92	1.01	0.06	6.11
S-N.D	1.41	1.15	0.00	4.84
S-N.Gn	1.25	0.94	0.03	4.07
S-N.Ocl	1.98	1.54	0.04	6.65
(S-N).(Go-Me)	2.48	1.54	0.07	6.95
(Go-Gn).Ocl	2.50	2.36	0.14	11.66
/1./1	4.37	3.25	0.17	13.85
1/.NS	2.64	2.06	0.01	8.68
/1-Orbita	3.28	2.40	0.09	9.18
1/.NA	2.67	2.27	0.05	11.02
1/-NA	1.44	1.84	0.01	11.33
/1.NB	3.58	2.84	0.02	16.28
/1-NB	0.84	1.12	0.04	6.81
/1-NPog	0.69	1.13	0.01	7.00
H.(N-B)	1.41	1.42	0.04	5.45
H-Nose	1.30	0.86	0.11	3.81
Pog-NB	0.80	0.53	0.01	2.46
Genial tubercle	0.88	0.76	0.02	3.36
FMIA	5.28	4.05	0.16	20.13
FMA	2.66	2.01	0.13	8.31
IMPA	4.30	3.09	0.19	17.24
/1-Linha I	0.92	1.32	0.01	6.73
(Go-Me).(V-T)	3.01	1.88	0.14	7.82
F.(V-T)	3.58	2.99	0.06	10.22
A-(V-T)	2.35	1.76	0.03	6.56
/1.(V-T)	1.33	1.09	0.05	4.18
H.(V-T)	3.15	2.12	0.24	8.62
CD(Vigorito)	1.05	1.46	0.00	6.92

TABLE 4. Mean, Standard Deviation, and Minimum and Maximum Differences Between Clinics A and C for Cephalometric Measurements (in Millimeters or Degrees), Porto Alegre, Brazil

Variable	Differences Between Clinics			
	Mean	SD	Minimum	Maximum
(N-Pog).(Po-Orb)	2.52	1.81	0.13	8.58
N-A.Pog	2.13	2.32	0.04	12.57
S-N.A	1.36	1.32	0.07	6.48
S-N.B	0.92	0.92	0.01	3.66
A-N.B	0.92	0.90	0.01	4.69
S-N.D	1.05	0.88	0.03	3.63
S-N.Gn	0.93	0.76	0.08	3.00
S-N.Ocl	1.98	1.66	0.17	7.86
(S-N).(Go-Me)	1.98	1.32	0.16	5.58
(Go-Gn).Ocl	2.75	2.26	0.05	9.57
/1./1	5.05	4.00	0.21	18.96
1/.NS	3.92	3.09	0.04	13.63
/1-Orbita	3.05	2.00	0.04	7.80
1/.NA	3.96	3.13	0.21	12.90
1/-NA	1.55	1.79	0.04	10.82
/1.NB	2.24	2.54	0.04	14.43
/1-NB	0.94	0.99	0.09	5.41
/1-NPog	0.81	1.01	0.04	6.10
H.(N-B)	1.24	1.34	0.04	7.17
H-Nose	1.28	1.04	0.05	4.25
Pog-NB	0.70	0.57	0.00	2.43
Genial tubercle	1.07	1.02	0.03	4.34
FMIA	3.59	2.97	0.25	16.03
FMA	2.61	2.00	0.07	7.20
IMPA	4.04	3.21	0.06	18.06
/1-Linha I	0.82	1.09	0.02	6.15
(Go-Me).(V-T)	3.28	2.38	0.08	8.89
F.(V-T)	3.95	2.31	0.28	9.71
A-(V-T)	2.80	2.00	0.22	7.47
/1.(V-T)	1.74	1.39	0.17	4.60
H.(V-T)	4.70	3.59	0.08	15.46
CD(Vigorito)	6.66	4.90	0.41	17.70

and marking cephalometric landmarks. When the issue of identifying these landmarks is under discussion, questions about the accuracy of cephalometry are raised.^{5,9}

Several authors have investigated the difficulty in landmark identification in studies that compare, for example, cephalograms traced from two consecutive radiographs obtained from the same patient;^{14,15} computerized and manual cephalometry;^{6,16} radiographic cephalometry determined for two samples, one of dry skulls and one of patients;¹⁷ radiographic cephalometry with or without steel ball markers;^{7,18} cephalometry on digitized and conventional digital radiographs;¹⁹ and accuracy on digital radiographs with varying image resolutions.¹¹

Few authors have focused their studies on interobserver variation in landmark identification.¹⁰ Results have shown a high rate of interobserver errors in the identification of landmarks. In this study, we found interobserver variation in the identification of landmarks

in cephalometric analyses conducted by different radiology clinics.

In Brazil, the current practice is for dental radiology clinics to prepare cephalograms. It has been noticed that the dentists who require such cephalograms hardly ever question the values of cephalometric measurements and base their decisions on measurements that may be wrong. The idea for this study arose from this observation.

Cephalometric errors may be classified as errors of acquisition, identification, or measurement.¹¹ In this study, acquisition and measurement errors were controlled by the use of radiographs of good technical quality and by the fact that all analyses were computer-aided digitally. The results of this study revealed a significant difference in most of the measurements made by the three different clinics for the same radiograph. Such results strongly suggest that landmarks must have been identified in different sites. These findings are in agreement with those reported by several authors who have studied interobserver variation in

TABLE 5. Mean, Standard Deviation, and Minimum and Maximum Differences Between Clinics B and C for Cephalometric Measurements (in Millimeters or Degrees), Porto Alegre, Brazil

Variable	Differences Between Clinics			
	Mean	SD	Minimum	Maximum
(N-Pog).(Po-Orb)	1.47	1.11	0.08	4.58
N-A.Pog	2.07	1.79	0.06	6.43
S-N.A	1.08	0.98	0.01	3.96
S-N.B	1.15	0.81	0.05	3.20
A-N.B	0.88	0.74	0.02	2.64
S-N.D	1.01	0.61	0.03	2.25
S-N.Gn	0.99	0.74	0.03	3.04
S-N.Ocl	2.02	1.60	0.07	6.26
(S-N).(Go-Me)	1.22	0.91	0.03	3.32
(Go-Gn).Ocl	2.11	1.81	0.16	8.53
1/.1	2.44	1.90	0.05	8.63
1/.NS	2.52	1.94	0.07	9.98
/1-Orbita	2.27	1.59	0.09	6.47
1/.NA	2.59	1.76	0.02	8.02
1/-NA	1.52	0.99	0.10	3.76
/1.NB	2.56	1.60	0.18	6.32
/1-NB	0.78	0.64	0.02	2.45
/1-NPog	0.68	0.63	0.02	2.70
H.(N-B)	0.98	0.95	0.00	5.12
H-Nose	0.86	0.67	0.00	2.74
Pog-NB	0.63	0.42	0.03	1.65
Genial tubercle	1.22	0.86	0.04	3.90
FMIA	3.67	1.89	0.02	7.88
FMA	3.00	1.64	0.02	5.79
IMPA	1.81	1.47	0.05	5.94
/1-Linha I	0.71	0.69	0.01	3.09
(Go-Me).(V-T)	3.33	2.56	0.00	10.54
F.(V-T)	4.35	3.08	0.34	12.48
A-(V-T)	3.30	2.31	0.04	9.52
/1.(V-T)	1.57	1.17	0.00	3.98
H.(V-T)	4.10	3.28	0.05	13.08
CD(Vigorito)	6.52	4.70	0.05	16.82

the preparation of cephalograms.^{10,13,17,20} Significant differences were found between all clinics in eight factors: S-N.D; (S-N).(Go-Me); 1/.NS; 1/.NA; FMIA; FMA; A-(V-T); and CD (Vigorito's cephalometric discrepancy).

In some cases, the result from one clinic differed from those from the two other clinics, which had equal results. Clinics A and C showed the most agreement, with equal results for 46.9% of the factors. Clinics A and B had the least agreement (28.1%). The analysis of maximum differences in interincisal angle (1/.1) revealed differences of 18.96° (between clinics A and C), 13.85° (between A and B), and 8.63° (between B and C). For other measurements involving the long axis of the incisors, significant differences were also found: FMIA = 20.13°; IMPA = 18.06°; /1.NB = 16.28°; and 1/NA = 12.9°.

Chan et al⁹ also found difficulties in the identification of upper and lower incisor angulation. These differences, although present in single cases, are significant and may affect therapeutic decisions. A difference of

up to 17.7 mm was found for CD (Vigorito). Cephalometric discrepancies, together with model discrepancies, may lead to misguided decisions about whether teeth should be extracted. For the factors SNA, S-N.Occl, /1-Npog, and H.(N-B), there was agreement in the location of landmarks, as shown by the fact that there were no statistically significant differences between the results for the three radiology clinics. This shows that the landmarks were located in similar sites by the three clinics.

Further studies to investigate cephalometry should be conducted not only because of the importance of this diagnostic resource but also because such studies may develop new methods with a lower probability of error.

CONCLUSIONS

- Our results indicate a very low reproducibility in the identification of cephalometric landmark points and angles for the majority of cephalometric measurements investigated in the study.
- Of all the cephalometric measurements evaluated, the ones that showed the least reproducibility were S-N.D; (S-N).(Go-Me); 1/.NS; 1/.NA; FMIA; FMA; A-(V-T); and CD (Vigorito).
- The aspects discussed above underscore the need to carefully evaluate the measurements provided by cephalometric analyses. They also show that observer calibration is fundamental for scientific purposes because of the high probability of error.

REFERENCES

1. Langlade M. *Cefalometria Ortodôntica*. São Paulo, Brazil: Editora Santos; 1993:269.
2. Moyers RE. *Handbook of orthodontics*. 4th ed. Chicago, Ill: Year Book Medical Publishers; 1988:577.
3. Richardson A. An investigation into the reproducibility of some points, planes and lines used in cephalometric analysis. *Am J Orthod*. 1966;52:637-651.
4. Vion PE. *Anatomia Cefalométrica*. 2nd ed. São Paulo, Brazil: Editora Santos; 2002:173.
5. Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop*. 1998;113:173-179.
6. Chen SK, Chen YJ, Yao JC, Chang HF. Enhanced speed and precision of measurement in a computer-assisted digital cephalometric analysis system. *Angle Orthod*. 2004;74:501-507.
7. Tng TT, Chan TC, Hägg U, Cooke MS. Validity of cephalometric landmarks. An experimental study on human skulls. *Eur J Orthod*. 1994;16:110-120.
8. Trenouth MJ, Laitung G, Naftel AJ. Differences in cephalometric reference values between five centers: relevance to the Eurocleft Study. *Br J Oral Maxillofac Surg*. 1999;37:19-24.
9. Chan CK, Tng TH, Hägg U, Cooke MS. Effects of cepha-

- lometric landmark validity on incisor angulation. *Am J Orthod Dentofacial Orthop.* 1994;106:487–495.
10. Lau PY, Cooke MS, Hägg U. Effect of training and experience on cephalometric measurement errors on surgical patients. *Int J Adult Orthodont Orthognath Surg.* 1997;12:204–213.
 11. Ongkosuwito EM, Katsaros C, van't Hof MA, Bodegom JC, Kuijpers-Jagtman AM. The reproducibility of cephalometric measurements: a comparison of analogue and digital methods. *Eur J Orthod.* 2002;24:655–665.
 12. Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmarks identification and reproducibility: a meta analysis. *Am J Orthod Dentofacial Orthop.* 1997;112:165–170.
 13. Kamoen A, Dermaut L, Verbeeck R. The clinical significance of error measurement in the interpretation of treatment results. *Eur J Orthod.* 2001;23:569–578.
 14. Cooke MS, Wei SH. Cephalometric errors: a comparison between repeat measurements and retaken radiographs. *Aust Dent J.* 1991;36:38–43.
 15. Midtgard J, Bjork G, Linder-Aronson S. Reproducibility of cephalometric landmarks and errors of measurements of cephalometric cranial distances. *Angle Orthod.* 1974;44:56–61.
 16. 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;74:155–161.
 17. Major PW, Johnson DE, Hesse KL, Glover KE. Landmark identification error in posterior anterior cephalometrics. *Angle Orthod.* 1994;64:447–454.
 18. Hägg U, Cooke MS, Chan TC, Tng TT, Lau PY. The reproducibility of cephalometric landmarks: an experimental study on skulls. *Aust Orthod J.* 1998;15:177–185.
 19. Schulze RK, Gloede MB, Doll GM. Landmark identification on direct digital versus film-based cephalometric radiographs: a human skull study. *Am J Orthod Dentofacial Orthop.* 2002;122:635–642.
 20. Chen YJ, Chen SK, Chang HF, Chen KC. Comparison of landmark identification in traditional versus computer-aided digital cephalometry. *Angle Orthod.* 2000;70:387–392.