sciendo DOI: 10.2478/aoj-1998-0012

The reproducibility of cephalometric landmarks: an experimental study on skulls

Urban Hägg
DDS, Odont Dr

Michael S. Cooke
PhD, BChD(Hons), DOrth, DDPH(RCS Eng), FDS (RCS Edin), FFD (RCS Ire)

Tommy C.K. Chan
BDS, MDS

Tony T.H. Tng
BDS, MDS

Paul Y.W. Lau
BDS, MDS, DOrth, MOrth

Few previous cephalometric studies have used dry skulls in order to eliminate the effects of the facial soft tissues. The aim of this study were to quantify the intra-examiner reproducibility of the commonly used cephalometric landmarks, angles and distances, using dry skulls and to compare these errors with previous error data obtained from normal cephalograms of living patients.

Thirty skulls were mounted in a purpose-designed holder and repeat cephalograms of each skull were recorded and digitised. All measurements were made relative to an X-Y co-ordinate reference grid. Scattergrams were produced to show the characteristic distribution of the intra-examiner reproducibility errors for the landmarks, and the reproducibility of the angles and distances were tabulated. This data was compared with a related study that included the soft tissue profile.

The standard deviations of the skeletal and dental angles and distances were greater in the presence of the soft tissues. These differences were up to four times larger for measurements including the location of Nasion (e.g. SNA, SNB and SNPg), which was the landmark location most affected by the presence of soft tissues, and for the distance: Incisal tip of Lower Incisor to Pogonion. Anterior Nasal Spine, in the horizontal plane, was also greatly affected by the presence of the soft tissues.

By using dry skulls, the errors observed in this study may serve as standards for the rational interpretation of clinical cephalometric values.

Key Words: Cephalometrics, Craniometry, Reproducibility, Method errors, Soft tissues.

Aust J Orthod 1998; 15: 177-185

Received for publication: April 1996
Accepted: June 1998.

Introduction

The usefulness of cephalometrics both clinically and in research is limited by various sources of inherent errors.[1,2,3] Many studies have investigated the effects of repeat examiners, repeat tracings, and repeat radiographs on the reliability of cephalometric analysis. These studies have primarily used cephalograms from living subjects.[4,5,6,7,8,9,10,11] The major source of variability in landmark determination was related to the observer's opinion of the landmark location.[7] However, it has been reported that some cephalometric landmarks showed significantly less variability on replicated identification.[12]

It has been shown that, to some extent, the level of reliability is affected by the experience and training of the examiners.[6,7,13,14] Gravely and Benzies[7] compared the reliability of landmark location between practising orthodontists and a statistical technician. Kvam and Krogstad[6] compared reliability between dental students and postgraduate orthodontic students. Both studies[6,7] showed that the less experienced and less trained groups were less reliable in locating landmarks. In the study of Stabrun and Danielsen,[9] although the two observers were trained beforehand for the location of landmarks, there were still significant differences between their findings. It was suggested that this was due to disagreements regarding the application of the landmark definitions. Therefore, it is important that the examiners are in agreement, and have an understanding of the definition of the landmarks used. It is also important, if possible, to first complete a methodological study to calibrate the reproducibility of the landmarks.[9] On the other hand, Savage *et al.*[12] compared three groups of orthodontists with different levels of experience; their results showed that the level of observer experience was unrelated to the landmark variability of replicate identifications.

Aim of Study

The aim of this study was to determine the intra-examiner reproducibility of cephalometric landmark identification, using dry skulls. The use of dry skulls eliminates the confusing effects of the soft tissues; the study would therefore establish "absolute" standards against which measurement errors using living subjects could be compared. Data from a related cephalometric study of living subjects is included for comparison.[14,15]

Materials and Methods

Lateral cephalograms were taken of 30 skulls with reference steel-ball markers placed in the cranium. Lateral cephalograms were taken with the aid of a purpose-designed skull holder.[16,17] The radiographic images of two of the steel ball markers (R1 and R2, Figure 1) were used as reference points to construct the reference grid (Figure 1). The two assessments of each cephalogram of the same skull were superimposed along this reference line. The X-axis was defined by using the line joining R1 and R2. The Y-axis was defined as a line from R2, perpendicular to the X-axis (Figure 1).

Fifteen skeletal and dental landmarks (Figure 2) as defined by Björk[18] were digitized.[14,15,16,17] The lateral cephalograms were digitised twice, on two separate occasions, by one examiner and once by a second examiner. The distances along the X- and Y-axes were measured by the computer. The computer compared the differences between the two assessments of the cephalograms and calculated the mean differences, standard deviations and minimum and maximum differences. Student's two-tailed t-test was performed.

The comparative error data from living subjects was taken from a related study by the same research group.[14,15] Thirty adult orthodontic patients had cephalometric radiographs recorded, and repeat measurements made, using the same equipment and measuring methods as used in this current study.[14,15,17]

Results

Intra-examiner cephalometric landmark errors (skull only) (Table I, Study A).

The mean differences between the assessments were small, varying from -0.46 mms to 0.36 mms and there were no statistically significant differences ($p < 0.05$), except for Anterior Nasal Spine, Pogonion, Condylion and Upper Molar Cusp Tip along the X-axis, and for A-point along both the X- and Y-axes (Table I). The scattergrams for the skeletal and dental landmarks are shown in Figure 3. The distribution of errors varied along both the X- and Y- axes.

The standard deviations of the landmarks in the skull base were small in comparison with the other landmarks, being close to 0.20 mms along both the X-axis and the Y-axis. The standard deviations of the skeletal landmarks in the maxilla and mandible were higher, being 0.32 mms to 1.53 mm along the X-axis and 0.24 mms to 1.36 mms along the Y-axis. In the mandible, the standard deviations were large for Gonion (1.53 mms) along the X-axis and B-point (1.36 mms) along the Y-axis. The standard deviations for the dental landmarks Upper Incisor and Lower Incisor Apices (0.74 mms and 0.95 mms) were about four times higher than those of Upper and Lower Incisal Edge (0.20 mms).

Intra-examiner cephalometric angles and distance errors (skull only) (Table II, Study A)

The mean differences for the cephalometric angles and distances were small overall but statistically significant ($p < 0.05$) for SNA, SNB, SNPg, ANB and SN/MxP. The standard deviations for the angles involving skeletal landmarks only (0.23° to 0.62°) were four to six times smaller than those involving skeletal and dental landmarks (1.31° to 1.60°).

Comparison with intra-examiner cephalometric errors with soft tissues present (Tables I and II, Study B)

The original data has been published previously.[14] The mean errors were similar to the present "skull only" values. However, the standard deviations with the soft tissues present were consistently greater than the "skull only" values (Study A); these have been expressed (Table I) as the standard deviations with the soft tissues present (Study B) as a multiple of the standard deviations from Study A (skulls only). The increased standard deviations (with soft tissues present) were four times greater for Nasion on both the X- and Y- axes (X:3.6 Y:4.4) and for Anterior Nasal Spine in the X-axis (4.1). These related to the increased standard deviations for the angles and distances including Nasion location with SNB (4.3) and SNPg (4.6) showing four times the standard deviations with the soft tissues present. SNA (3.3), UI/SN (2.3) and N-Me (3.0) were showing two to three times the standard deviations with the soft tissues present. The distance Lower Incisor Edge to the A-Pogonion line also showed a fourfold increase in relation to the standard deviation with the soft tissues present.

Table I. Intra-examiner errors for the skull sample (A). The linear difference (mm) between repeat recording along the X- and Y-axes (n=30). For comparison, the table also lists the intra-examiner errors from a related study (B) in which living subjects were evaluated (i.e. the soft tissues were present) (Lau P., Cooke M. and Hägg., 1995) In the final column, the standard deviation with the soft tissues present (A) is expressed as a multiple of the standard deviation for the skulls samples only (A). Additionally, the standard deviations along the X-axis are expressed as a multiple of those along the Y-axis.

Variable	SKULL ONLY (A) Intra Examiner Error				WITH SOFT TISSUES (B) Inter Examiner Error			
	Mean		S.D.	S.D. X-Axis / S.D. Y-Axis	Mean	S.D.	S.D. X-Axis / S.D. Y-Axis	S.D.(B) / S.D.(A)
SKELETAL LANDMARKS								
Skull Base								
Sella-X	-0.1		0.24		0.1	0.32		1.3
Sella-Y	-0.1		0.18	1.3	0.1	0.39	0.8	2.2
Nasion-X	0.0		0.16		0.0	0.58		3.6
Nasion Y	0.1		0.21	0.8	0.2	0.92	0.6	4.4
Maxilla								
ANS-X	-0.1		0.32		0.3	1.30		4.1
ANS-Y	0.0		0.31	1.0	0.2	0.42	2.1	1.4
PNS-X	0.0		0.54		0.0	1.69		3.2
PNS-Y	0.1		0.35	1.5	0.0	0.63	2.7	1.8
A-Point-X	-0.2	**	0.33		0.0	1.07		3.2
A-Point-Y	-0.5	**	0.92	0.4	0.0	2.09	0.5	2.3
Mandible								
B-Point-X	0.0		0.40		0.1	0.32		0.8
B-Point-Y	0.2		1.36	0.3	0.1	1.68	0.2	1.2
Pogonion-X	-0.1		0.24		0.1	0.37		1.5
Pogonion-Y	-0.3		0.77	0.3	0.3	1.54	0.2	2.0
Menton-X	0.1		0.46		0.0	0.81		1.8
Menton-Y	0.0		0.23	2.0	0.0	0.38	2.1	1.7
Gonion-X	0.2		1.53		0.1	1.17		0.8
Gonion-Y	-0.2		0.77	2.0	0.0	0.51	2.3	0.7
Condylion-X	0.4		0.78		0.4	1.67		2.1
Condylion-Y	-0.1		0.68	1.1	0.2	1.37	1.2	2.0
DENTAL LANDMARKS								
Maxilla								
Upper Molar Cusp Tip-X	-0.4	*	0.89		0.0	0.95		1.1
Upper Molar Cusp Tip-Y	0.2		0.52	1.5	0.3	0.61	1.6	1.2
Upper Incisor Apex-X	0.0		0.82		0.0	1.54		1.9
Upper Incisor Apex-Y	-0.2		0.74	1.7	0.0	1.42	1.1	1.9
Upper Incisor Incisal Tip-X	0.0		0.21		0.1	0.35		1.7
Upper Incisor Incisal Tip-Y	0.0		0.18	1.2	0.0	0.46	0.8	2.6
Mandible								
Lower Incisor Incisal Tip-X	0.0		0.20		0.0	0.25		1.3
Lower Incisor Incisal Tip-Y	0.0		0.21	1.0	0.0	0.36	0.7	1.7
Lower Incisor Apex-X	-0.2		0.95		0.1	0.91		1.0
Lower Incisor Apex-Y	-0.2		0.87	1.1	0.1	1.03	0.9	1.2

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Discussion

The intra-examiner errors for the identification of cephalometric landmarks were all found to be less than 0.5 mms. The statistically significant systematic errors were therefore due to the small standard deviations rather than to large mean differences and, therefore, they were not considered to be of clinical significance.[4,19] However, the 95 per cent confidence interval was significant for the skeletal landmarks (1 to 3 mms) in the maxilla, for PNS-X and A point-Y, and in the mandible for B-Point-Y, Pogonion-Y, Gonion-X-Y and Condylion-X-Y. The 95 per cent confidence interval was significant for dental landmarks except Upper and Lower Incisal Edges (Table I). For the cephalometric angles and distances the intra-examiner errors were all small, although some were statistically significant (Table II). The 95 per cent confidence intervals were significant for all the dental angles, being 2.4 to 5.3 degrees.

In general, the distribution of errors for most landmarks was not random. Each landmark had its own characteristic envelope of errors (Figure 3). The reliability of estimating the

Table II. Intra-examiner errors for the skull sample (A). The angular (?) and linear (mm) difference between repeat recordings (n=30). For comparison, the table also lists the intra-examiner errors from a related study (B) in which living subjects were evaluated (i.e. the soft tissues were present) (Lau P., Cooke M. and Hägg U., 1995). In the final column, the standard deviation with the soft tissues present (B) is expressed as a multiple of the standard deviation for the skull sample only (A)

Variable	SKULL ONLY (A) Intra-Examiner Error		WITH SOFT TISSUES (B) Inter-Examiner Error			
	Mean	S.D.	Mean	S.D.	S.D.(B) / S.D.(A)	
SKELETAL ANGLES						
SNA	-0.2	**	0.36	0.1	1.17	3.3
SNB	-0.1		0.24	0.0	1.03	4.3
SNPg	0.1		0.23	0.1	1.06	4.6
ANB	0.1		0.29	0.0	0.94	3.2
SN/MxP	0.2		0.44	0.1	1.03	2.3
SN//MnP	0.0		0.43	0.1	0.90	2.1
MxP-MnP	-0.2		0.62	0.2	0.93	1.5
SKELETAL DISTANCES						
N-Me	0.0		0.32	0.1	0.99	3.0
N-MxP	0.1		0.33	0.0	0.97	3.0
MxP-Me	0.1		0.27	0.1	0.56	2.1
DENTAL ANGLES						
UI/SN	0.1		1.34	0.1	3.11	2.3
UI/MxP	0.3		1.34	0.4	2.55	1.9
LI/MnP	0.1		1.31	0.1	1.43	1.1
UI/LI	0.0		1.60	0.2	2.99	1.9
DENTAL DISTANCES						
Ii-Apg	0.1		0.21	0.0	0.88	4.2

(lower incisor incisal tip to A-Pg)
* $p < 0.05$ ** $p < 0.01$

landmarks was a function of how distinctly the contour appeared in the region of the point being estimated.[2] In this study, landmarks with good edge definition and less affected by superimposed structures, such as Sella and Nasion, were more reliable than those within the confines of the skulls, where definition was blurred by superimposed structures. Similar findings were reported in previous studies.[2,9,20] The error in estimating Posterior Nasal Spine tended to be distributed primarily along the X-axis. Due to the overlapping images of the adjacent structures, there was a lack of direct physical image of the tip of the Posterior Nasal Spine on the cephalograms, thereby increasing the difficulty in estimating this landmark. As a result, the estimation tended to be along the palatal plane. These findings confirm other studies.[2,20,21]

Reproducibility errors, expressed by the 95 per cent confidence interval, for Upper and Lower Incisor Apices were about four times greater than those for the Upper and Incisal tip of Lower Incisor. This was due to the superimposed images of the surrounding structures over these landmarks. The lack of physical evidence of the apices forced the projection of the position of the apices based on a general knowledge of the usual length of a tooth, and on the expected rate of taper perceived from the crown and visible portion of the root.[9] The superimposed images of the surrounding teeth of Upper Molar Cusp Tip accounted for the lack of physical images of the left and right Molar Cusp, which made the estimation of this landmark difficult, especially along the X-axis.[9,20]

Where the outline was a gradual curve, the distribution of errors tended to be proportionately larger; as along the surface of the skull.[4,9,20,21] This was found in this study for A-point, B-point and Pogonion, whose errors tended to be distributed along the Y-axis, and for Menton and Gonion along the X-axis, especially in those skulls where the curve outline was gradual or flat.

The errors of landmark location on cephalograms found in this study confirmed many of the findings of previous studies[4,9,20,21,22] that have also reported an envelope of error as characteristic for each landmark. The errors

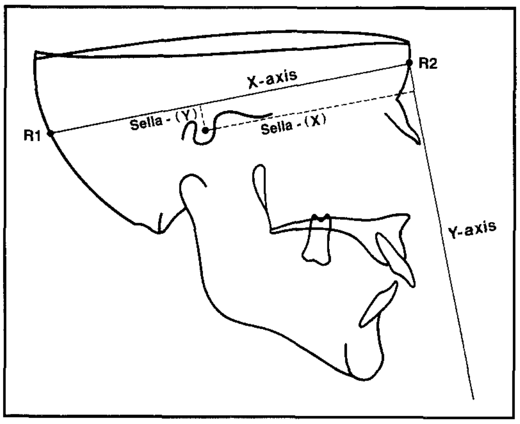
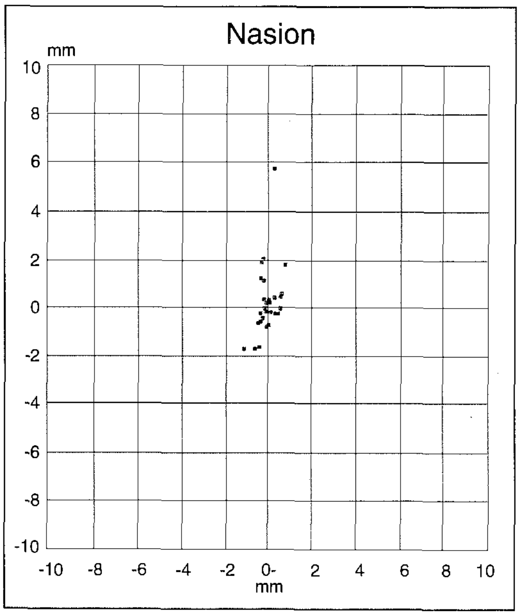


Figure 1. The co-ordinate (X-Y) reference system based on the arbitrary reference points R1 and R2. All the landmarks were measured with respect to the X and Y axes, as illustrated by the landmark Sella.



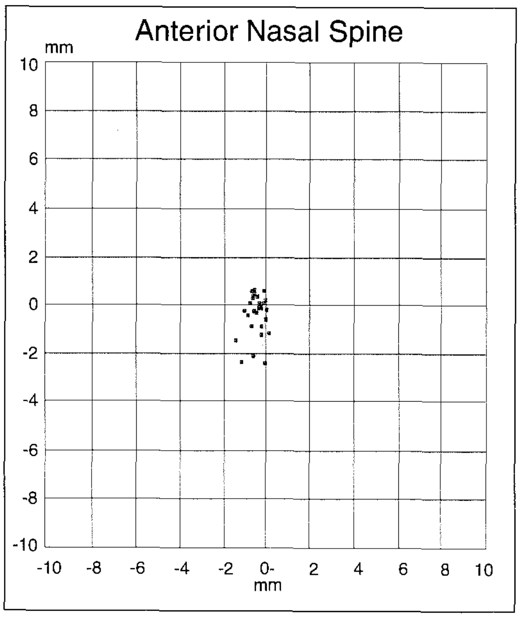


Figure 3a. Scattergrams showing the reproducibility (intra-examiner double registrations) of Nasion, Sella, ANS and PNS. The distribution of errors for Sella, Nasion and ANS tended to spread equally along both the X- and Y- axes, whereas for PNS, the errors were mainly along the X-axis.

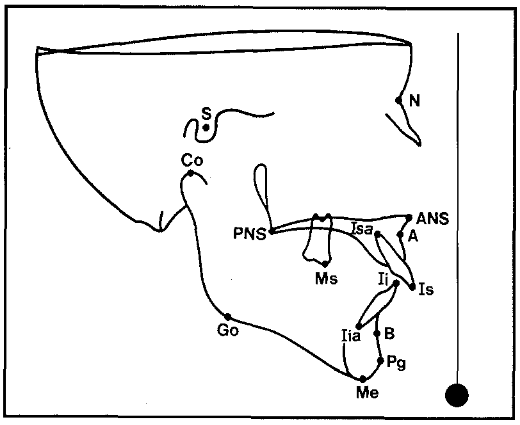
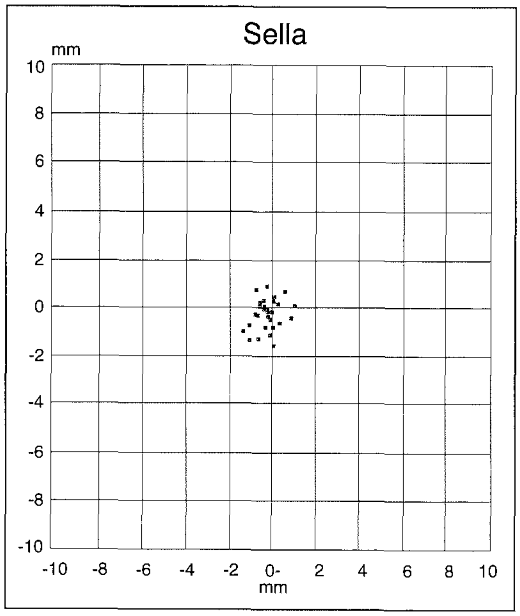
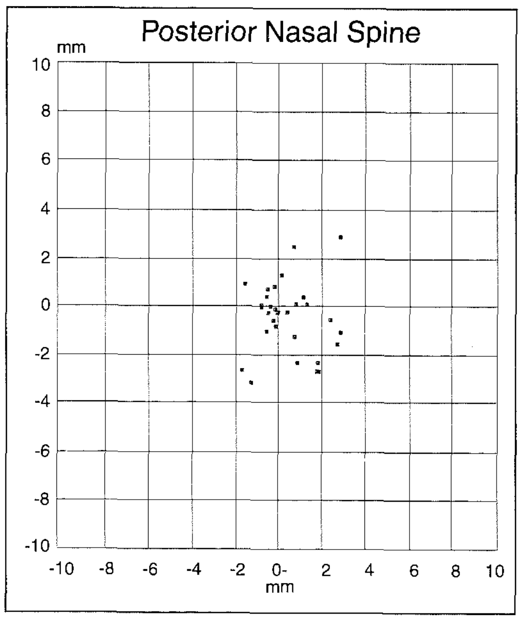


Figure 2. The dental and skeletal landmarks (Björk, 1947) used in this study. The maxillary plane (MxP), mandibular plane (MxP), upper incisor long axis (UI) and lower incisor long axis (LI) were defined by their respective landmarks PNS-ANS, Go-Me, Isa-Is and Iia-Ii. The upper and lower face height (N-MxP and Me-MxP respectively) and the distance lower incisor edge to the A-pogonion plane (Ii-APg) were defined as the shortest distance between the landmarks and the plane. Landmarks defined as "deepest", "most anterior" for example, were located relative to the plumb line.





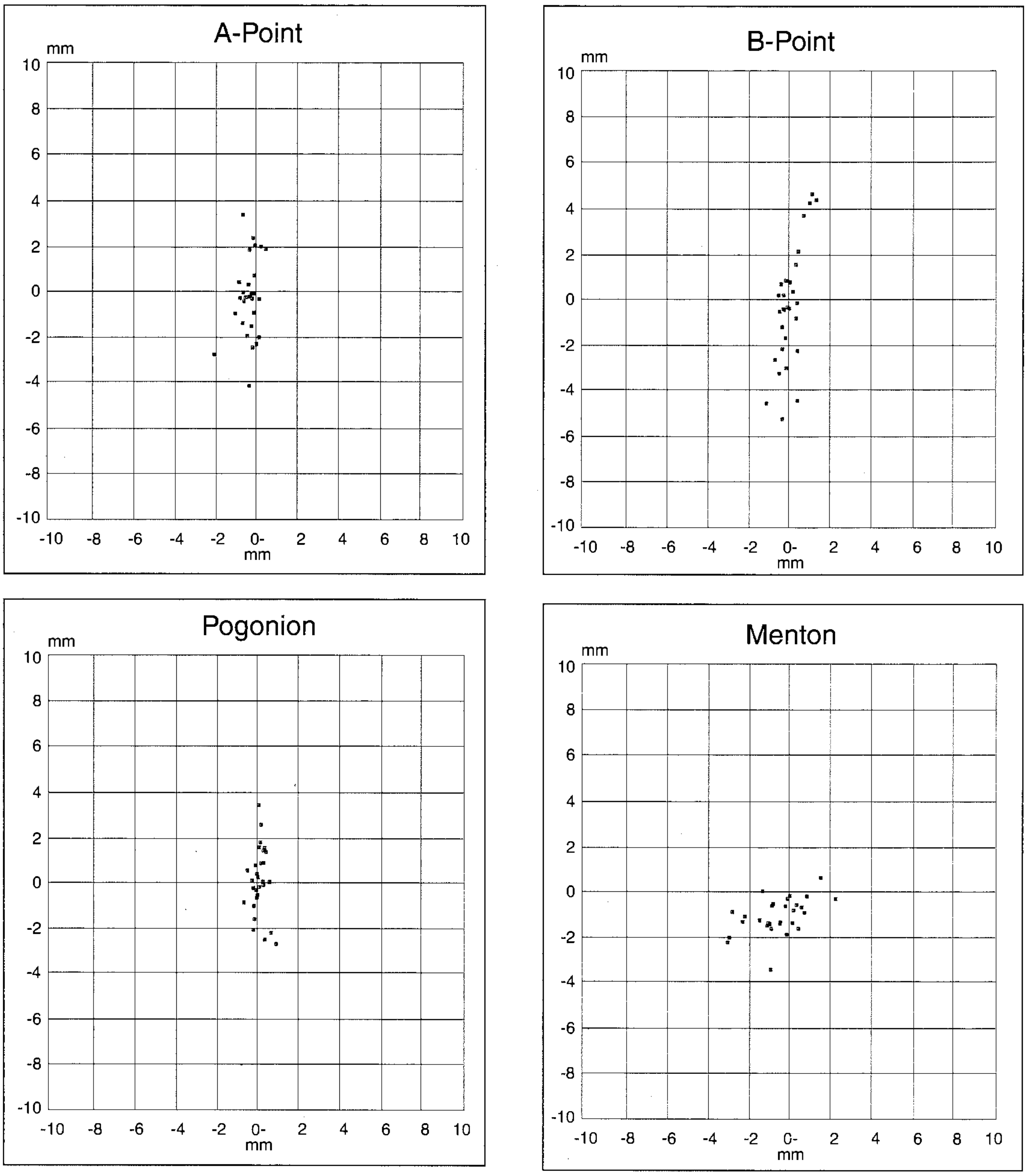


Figure 3b. Scattergrams showing the reproducibility (intra-examiner double registrations) of A-Point, B-Point, Pogonion and Menton. The distribution of errors for A-Point, B-Point and Pogonion were mainly along the Y-axis, whereas for Menton the errors were mainly along the X-axis.

(reproducibility) showed a characteristic distribution for each landmark in respect to the X- and to the Y- axes. The characteristic error distribution will have corresponding effects on the measurement and interpretation of cephalometric angles and distances.

It was useful to be able to compare the measurement errors "with" and "without" the presence of the soft tissues. This was made possible by the previous completion of a related study in which one examiner used the same landmarks, angles and distances and the same cephalometric techniques.[14] The previous study used living subjects, whereas the present study used skulls only. The comparisons are shown in Tables I and II. The major effect of the presence of soft tissues was to increase the standard deviations of all the variables. The landmarks most affected by the presence of the soft tissues were Nasion (in both the horizontal and vertical) and Anterior Nasal Spine in the horizontal plane. This finding was surprising for Nasion and may relate to the superimposition of the upper eyelid in Chinese subjects, as previously discussed by Tng.[17]

For the angles and distances, the greatest effect of the soft tissues was shown by variables that included Nasion for their assessment; SNB and SNPg being the most affected having a fourfold increase in these standard deviations with the soft tissues present, and SNA, UI/SN and N-Me showing a two- to three-fold increase.

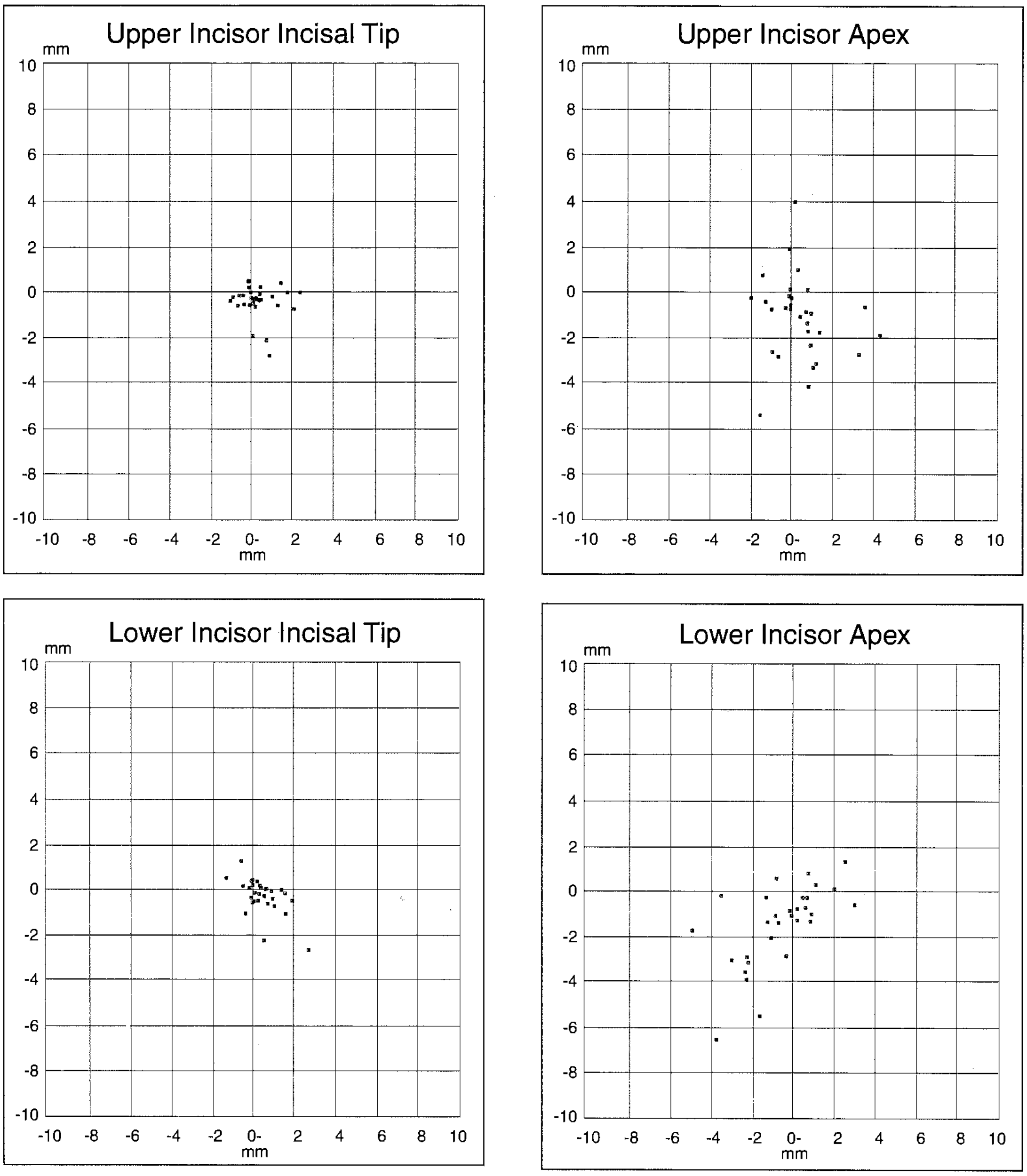
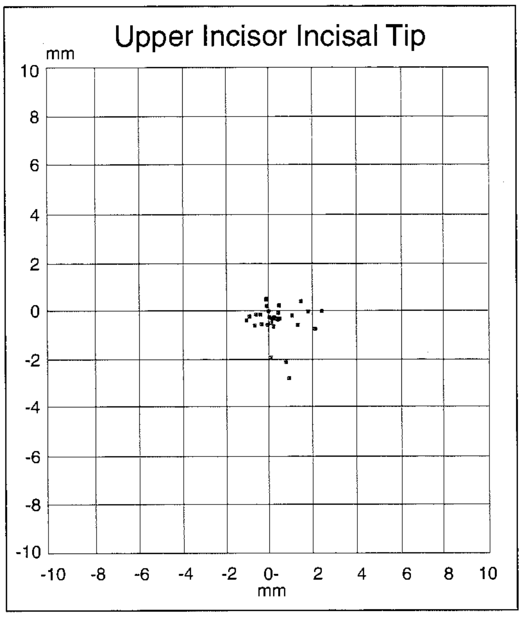


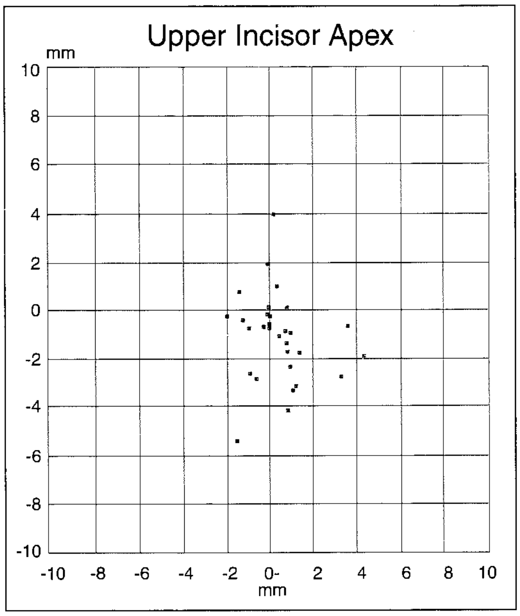
Figure 3c. Scattergrams showing the reproducibility (intra-examiner double registrations) of Upper Incisor Incisal Tip, Upper Incisor Apex, Lower Incisor Incisal Tip and Lower Incisor Apex. The distribution of errors for the four incisal landmarks spread equally along the X- and Y- axes.

The distance from the Incisal tip of the Lower Incisor to the A-Pogonion line also showed a four-fold increase in its standard deviation with the soft tissues present. These findings clearly have clinical relevance and the "errors" in measuring these variables must be understood when making clinical decisions. The present study on skulls only serves to establish "standards" by which to compare cephalometric studies, and error studies, on living subjects.

Conclusions

1. The intra-examiner errors associated with the location of dental and skeletal landmarks were small and were characteristically distributed along the X- and Y- axes.
2. The intra-examiner errors for the cephalometric angles were small and were related to the envelope of errors for the corresponding landmarks.
3. The principal effect of the normal presence of soft tissues (that is, in living subjects) is to increase (up to four times) the standard deviations of the error measurements. The landmarks most affected by the presence of soft tissues are Nasion and Anterior Nasal Spine. The most affected angles are SNA, SNB and SNPg and the most affected distance was Lower Incisor Edge to the A-Pogonion line.





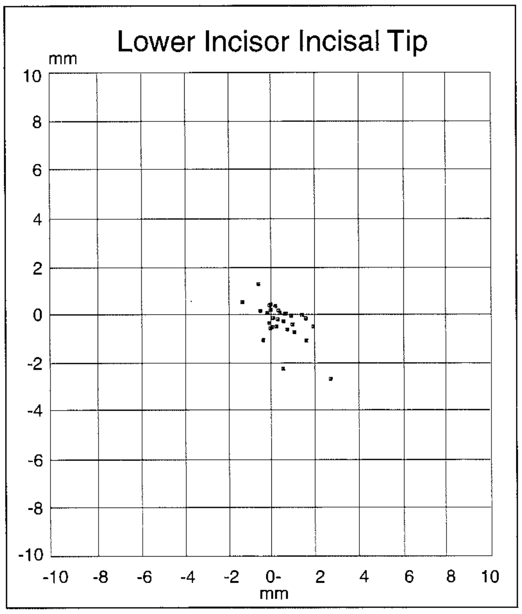


Figure 3d. Scattergrams showing the reproducibility (intra-examiner double registrations) of Gonion, Condylion, and Upper Molar Cusp Tip. The distribution of errors for Gonion and Upper Molar Tip spread mainly along the X-axis, whereas for Condylion the errors were equally along the X- and Y- axes.

Acknowledgements

The authors would like to express their thanks to the Anatomy Department at The University of Hong Kong, for the loan of the skulls; Mr S.L. Lee of the Data Processing Unit for writing the computer program; the Dental Illustration Unit for the illustrations and the University Industrial Centre for constructing the skull holder. Thanks also to Ms Zinnia Pang for preparation of the manuscripts. This research was supported by a grant from the University of Hong Kong.

References

1. Hallett GEM. Inherent errors in cephalometric films and their reduction. Transaction of the Brit Soc of Orthod 1959; 1-10.
2. Baumrind S, Frantz RC. The reliability of head film measurements 2. Conventional angular and linear measures. Am J Orthod 1971; 60(5): 505-17.
3. Houston WJB. The analysis of errors in orthodontics measurements. Am J Orthod 1983; 83(5): 382-9.
4. Richardson A (1966). An investigation into the reproducibility of some points, planes, and lines used in cephalometric analysis. Am J Orthod 1966; 52(9): 637-51.
5. Broadway ES, Healy MJR, Poyton HG. The accuracy of tracing from cephalometric lateral skull radiographs. Dent Pract 1962; 12(12): 455-8.
6. Kvam E, Krogstad O. Variability in tracings of lateral head plates for diagnostic orthodontic purposes. A methodologic study. Acta Odont Scand 1969; 27: 360-9.
7. Gravely JF, Benzies PM. The clinical significance of tracing error in cephalometry. Brit J Orthod 1974; 1:95-101.
8. Midtgard J, Björk G, Linder-Aronson S. Reproducibility of cephalometric landmarks and errors of measurements of cephalometric cranial distances. Angle Orthod 1974; 44: 56-61.
9. Stabrun AE, Danielsen K. Precision in cephalometric landmark identification. Eur J Orthod 1982; 4: 185-96.
10. El-Mangoury NH, Shaheen SI, Mostafa YA. Landmark identification in computerized posteroanterior cephalometrics. Am J Orthod 1987; 91(1): 57-61.
11. Cooke MS and Wei SHY. Cephalometric errors: A comparison between repeat measurements and retaken radiographs. Aust Orthod J 1991; 36:38-43.
12. Savage AW, Showfety KJ, Yancey J. Repeated measures analysis of geometrically constructed and directly determined cephalometric points. Am Orthod 1987; 91(4): 295-9.
13. Krogman WM. Validation of the roentgenographic cephalometric technique. Am J Orthod 1958; 44(12): 933-9.
14. Lau YW, Cooke MS, Hägg U. Effects of training and experience on cephalometric measurement errors on surgical patients. Int J Adult Orthod Orthog Surg 1997; 12: 204-13.
15. Lau YW, Cooke MS, Hägg U. Comparative cephalometric errors for orthodontic and surgical patients. Int J Adult Orthod Orthog Surg 1995; 10: 119-26.
16. Chan T, Tgn TH, Hägg U, Cooke MS. A skull holding devise for experimental cephalometric research. Eur J Orthod 1996; 18: 193-7.
17. Tng TH, Chan CK, Hägg U and Cooke MS. Validity of cephalometric landmarks. An experimental study on human skulls. Eur J Orthod 1994; 16: 110-20.
18. Björk A. The face in profile: An anthropological X-ray investigation on Swedish children and conscripts. Svensk Tandläkare Tidskrift 40 (supple. no. 5B): 1947; 1-180.
19. Cook PA, Gravely JF. Tracing errors with Björk's mandibular structures. Angle Orthod 1988; 169-78.
20. Vincent AM, West VC. Cephalometric landmark identification error. Aust Orthod J 1987; 10(2): 98-104.
21. Broch J, Slagsvold O, Rosler M. Error in landmark identification in lateral radiographic headplates. Eur J Orthod 1981; 3: 9-13.
22. Cohen AM. Uncertainty in cephalometrics. Brit J Orthod 1984; 11: 44-8