Reproducibility of Cephalometric Landmarks and Errors of Measurements of Cephalometric Cranial Distances

JACOB MIDTGÅRD, D.D.S.
GUNNAR BJÖRK, D.D.S.
STEN LINDER-ARONSON, D.D.S.,Ph.D.

INTRODUCTION

The development of cephalometrics has created a need for exactly locating an increased number of landmarks on roentgenograms. These landmarks can be used as registration marks for measurements or for superimposition of films in a series of examinations.

It is known that differences can occur when localizing cephalometric landmarks. By comparing the variation in a number of measurements of different landmarks it is possible to conclude which landmarks are most reproducible.

Some of the landmarks, used clinically, are located on the outlines of the cranium and are comparatively easy to identify due to the sharpness in contrast of the roentgenograms. The structures of the inner cranium are, on the other hand, often indistinct because of summations of superimposed anatomical details. Baumrind and Frantz[1] have found that landmarks nasion and menton which are placed on anatomically-formed edges or creases are easy to identify, whereas landmarks which are placed on curves with wide radii show proportionally greater errors of measurement. Each landmark has its own characteristic pattern of errors.

It is important that nasion be easy to reproduce, since a number of conventionally used square rules in clinical use are registered from here.

From the Department of Orthodontics, Univ. of Örebro, Sweden.

Baumrind and Frantz have statistically concluded that the probability of placing sixteen landmarks correctly is forty-four per cent. Clinical analyses should for this reason be controlled at least twice. A simplification of the process of analyses can be carried out by use of an anatomical data-reducing system.

Björk[2] has described three reasons for error of method in cephalometric measurement studies:

1. Differences between two films of the same individual.
2. Differences caused by variation of the positioning of the landmarks.
3. Errors in the reading process.

His analysis of error of method reveals large differences in precision when localizing different cranial landmarks. Only minor errors in measurement have, however, been established with landmarks which are easily identifiable. Linear errors of measurement in these cases vary between 0.3 and 1.4 mm and angular errors of measurement between 0.3 and 1.6 degrees.

Linder-Aronson[3] has estimated the degree of error in cranial distance measurements by calculating the variance of error for the differences in distance in relation to the variance of the distance in question in the material as a whole. The variety of error has with few exceptions amounted to less than three per cent of the total variation.

Double registrations of size and direction of facial growth have shown minor deviations in a study carried out

by Lundström[4] in which forty-one pairs of twins were registered both initially and after thirteen years. A noticeable difference in the pattern of growth has, however, arisen depending on whether the roentgenograms have been orientated from the anterior or posterior skull base. The conclusion has been drawn that growth analyses based on lateral cephalometric roentgenograms do not give a sufficiently objective picture of the character of the growth changes.

Richardson[5] had two judges register cephalometric landmarks, lines, and angles on ten cephalograms with an interval of one week. He found that ordinary cranial landmarks have a margin of error of less than ±1 mm. Orbitale and Bolton points, however, show a higher count. Vertical deviations rise towards higher counts when anatomical curves in the profile are involved, as is the case with the points subspinale and supramentale. Horizontal deviations have been observed in particular for menton, spina nasalis anterior and pterygomaxillare. All angular measurements have followed the variation tendency of the landmarks.

The purpose of this investigation has been to study the reproducibility of fifteen landmarks which are part of the profile analysis carried out as a standard procedure at the Orthodontic Department at Örebro. Another purpose of the investigation has been to show the margin of error in measurements of seven cranial distances which also form part of this profile analysis.

Material

The research material consisted of lateral cephalometric roentgenograms taken on twenty-five children picked at random. Two consecutive roentgenograms were taken on each child during the same examination. The only criterion for selection was that root development of the upper and lower central incisors should be completed in order that registration of the apices of the upper and lower centrals should be possible. The average age of the examined children has therefore been 11.1 years.

Method

The roentgenograms have been taken with a Philips Rotapractix x-ray apparatus with film-focus distance of 165 cm and rotating anode with a focus 1.2 mm. The roentgenograms have been taken with the child's head held in a wall-mounted cephalostat ad modum Thörne and with the Frankfort plane horizontal.

The central ray passed through two ear rods on the cephalostat. The head was fixed so that the median plane was parallel to the film. The degree of enlargement of the median plane with the above-mentioned arrangement is 6.5 per cent.

Method of analysis

The measurements on the roentgenograms have been carried out on millimetre squared transparent acetate foil on which the landmarks have been marked.

A system for periodical identical mounting of tracing sheets has been constructed. A punching-machine has been used for simultaneous punching of two lateral cephalometric roentgenograms and a number of tracing sheets. When punching, the two roentgenograms were placed on top of each other with both the contours of the anterior and posterior bases of the cranium and the contours of ala magna os sphenoidale as reference plane.

In the process of tracing, the sheets and roentgenograms have been firmly mounted in a tenter (Fig. 1). On the tenter there are two round pegs made to the same dimensions as the punch-holes.

The following cephalometric landmarks and cranial distances have been

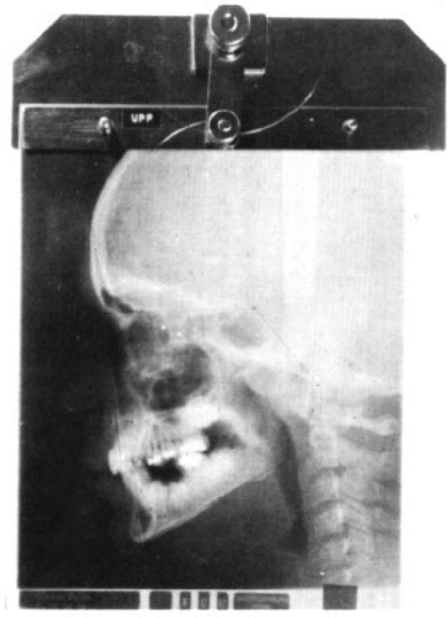

Fig. 1 Tenter with tightening arm for firm assembly of tracing sheet and roentgenogram.

studied: articulare (ar), basion (ba), gonion (go), menton (me), nasion (n), orbitale (or), pogonion (pg), porion (po), pterygomaxillare (pm), apex U-1 — U-1 (ap U-1 — U-1), apex L-1 — L-1 (ap L-1 — L-1), sella (s), spina nasalis anterior (sp), subspinale (ss), supramentale (sm), nasion-sella (n-s), sella-basion (s-ba), nasion-subspinale (n-ss), nasion-supramentale (n-sm), nasion-pogonion (n-pg), nasion-menton (n-me), menton-gonion (me-go).

Measurements

The following measurements have been carried out:

1. A comparison of the location of the landmarks on two consecutively taken roentgenograms of the same individual. This study has been carried out by one observer.
2. A comparison of the location of the landmarks on one and the same roentgenogram on two occasions with an interval of one month. This study has been carried out by each of two observers.
3. A comparison of the findings of two observers of the location of the landmarks on one and the same roentgenogram.
4. A comparison of distance measurements taken at a one month interval.

Statistical methods

With calculation of comparisons of means, significance analyses of t-test have been used.

To assess the importance of the error of method, the error variance $(\overline{E.M.}^2)$ was studied in relation to the variance for all the twenty-five individuals used in the subject material.

Error of method

In the case of variables for which the error variance is less than 3 per cent of the total variance, the error method has been estimated as being of little importance.

Results

Landmarks

When comparing the positions of the landmarks on two consecutively taken lateral cephalometric roentgenograms of the same child, the average differences, shown in Table I, have been obtained by one of the observers.

From Table I can be seen that, in the placing of the different landmarks on two separate roentgenograms, the observer has been so uncertain that the mean differences of all measuring points have differed significantly from 0.

The degree of uncertainty varies, however, from landmark to landmark. The greatest difference between two placings has been found for orbitale, where the mean difference has been greater than 2 mm (2.08). Average differences of about 1 mm have been observed for the landmarks supramentale (1.27), pogonion (1.20), spina nasalis

TABLE I

Mean difference, mean error and t-value of all 15 landmarks, calculated by one observer from two lateral cephalometric roentgenograms taken consecutively on each of 25 individuals.

landmark	$\bar{d} \pm \varepsilon\bar{d}$ mean difference ± mean error	t-value for 1% significance t>2.80
ar	0.52 ± 0.07	7.43
ba	0.70 ± 0.09	7.77
go	0.92 ± 0.17	5.42
me	0.68 ± 0.06	11.33
n	0.88 ± 0.12	7.33
or	2.08 ± 0.24	8.67
pg	1.20 ± 0.15	8.00
po	0.93 ± 0.20	4.65
pt	0.77 ± 0.10	7.70
ap U-1 — U-1	1.12 ± 0.15	7.46
ap L-1 — L-1	1.09 ± 0.15	7.23
s	0.41 ± 0.06	6.83
sp	1.17 ± 0.17	6.88
ss	0.83 ± 0.17	4.88
sm	1.27 ± 0.37	3.43

N = 25

TABLE II

Mean differences, mean error and t-value for the positions of the landmarks estimated by one observer on one and the same lateral cephalometric roentgenogram on two occasions with an interval of one month.

landmark	$\bar{d} \pm \varepsilon\bar{d}$ mean difference ± mean error	t-value for 1% significance t>2.80
ar	0.50 ± 0.13	3.85
ba	0.93 ± 0.18	5.17
go	0.82 ± 0.17	4.83
me	0.52 ± 0.10	5.20
n	0.56 ± 0.15	3.74
or	2.44 ± 0.46	5.30
pg	1.10 ± 0.19	5.80
po	0.63 ± 0.07	9.00
pm	0.81 ± 0.15	5.40
ap U-1 — U-1	0.76 ± 0.12	6.33
ap L-1 — L-1	0.76 ± 0.12	5.56
s	0.46 ± 0.08	5.75
sp	0.72 ± 0.15	4.80
ss	0.81 ± 0.15	5.40
sm	1.17 ± 0.07	6.88

N = 25

anterior (1.17), apex U-1 — U-1 (1.12) and apex L-1 — L-1 (1.09).

The greatest degree of certainty has been found for the landmarks sella turcica (0.41 mm) and articulare (0.52 mm).

When comparing the positions of the landmarks on one and the same lateral cephalometric roentgenogram on two occasions with an interval of one month, carried out by one observer, the mean differences shown in Table II have been found.

Table II shows roughly the same degree of variety in the values as has been shown in Table I. This means that the differences for the most part depend on the uncertainty of the observer when placing the measuring points. The method error with repeated x-raying, when the roentgenograms are taken immediately after each other on one and the same child, can be considered as of minor importance.

No significant differences are found between the landmarks in Table I and those in Table II. This indicates that the interval of one month between the measurements has been of no importance. Neither has any difference been apparent with measurements carried out on lateral cephalometric roentgenograms taken on different occasions.

The values from measurements of two judges of one and the same roentgenogram have also been compared and again there is no evidence of significant differences. The two examiners placed the different landmarks rather similarly. Both examiners have, for example, been of the opinion that the orbitale is the most difficult landmark to place.

Distance measurements

Double determinations must be made to be able to judge the degree of accuracy with distance measurements. In doing this it is of great interest to calculate the size of the error of measurement in relation to the variance in the

TABLE III

Distance measurements, number, error of measurement, variance of error, variance and variance of error in per cent of the variance for the material as a whole.

	E.M.	$\overline{E.M.^2}$	$\overline{S.D.^2}$	$\overline{E.M.^2}$ in % $\overline{S.D.^2}$
distance	error of measurement	the variance of error	variance	the variance of error in % of the variance
n-s	0.55	0.30	10.11	3.0
s-ba	0.50	0.25	9.36	2.7
n-ss	0.87	0.76	12.67	6.0
n-sm	1.10	1.22	21.34	5.7
n-pg	0.98	0.96	29.59	3.2
n-me	0.35	0.12	30.03	0.4
me-go	0.58	0.33	19.18	1.7

N = 25

material as a whole for the variable in question.

In Table III is rendered an account of the various distance measurements, number (n), the error of measurement (E.M.), the variance of error ($\overline{E.M.^2}$), the variance ($\overline{S.D.^2}$) and the variance of error in per cent of the variance for the material as a whole.

From Table III can be seen that the greatest inaccuracy has been found in estimating the distances n-ss and n-sm, since the variance of error in per cent of the variance in the material as a whole for these distances is, respectively, 6.0 and 5.7. Of the remaining distance measurements the greatest accuracy has been obtained for the distances n-me and me-go with the respective percentage values 0.4 and 1.7.

Discussion

Few investigations have been performed of the reproducibility of landmarks. This is surprising since the positions of these landmarks are of primary importance in studies of the interrelationships between the skull bones, as also in growth studies of the facial skeleton. The value of a study of this kind is that one can better evaluate cephalometric analysis if one is sure of the reliability of the landmarks. This is of particular importance if the landmarks are used as a matter of routine in a data-cephalometric system, or if the landmarks are used for scientific reasons.

This investigation shows that those landmarks which have been the subject of analysis have not been exactly reproducible since the mean differences for all landmarks have differed significantly from zero. The low mean errors show, on the other hand, that there has been little deviation between mean differences for the various landmarks. In measuring distances the greatest deviations were shown for the landmarks ss and sm. It is not surprising that comparatively wide deviation has been found for these two landmarks as they are marked on an often unsharp vertically-placed curvature. Of particular interest is that the observed results have shown that the interval of one month between two registrations has not considerably affected the reproducibility. These observations correspond well with the results described by Richardson in 1966 and Baumrind in 1971.

The reproducibility of the landmarks in question does not seem to be affected by the fact that the registrations were made by two judges. From this the conclusion can be drawn that both observers have defined the landmarks similarly and that both have been equally experienced in locating the landmarks.

Most cephalometric researches give account of the size of the errors of distance measurements through calculation of the accidental error between double determinations. This documentation is, however, insufficient if at the same time the variance of error is not put in relation to the variance of the landmarks in question for the material as a whole. The variance of error should not exceed 3 per cent of the variance in the material as a whole. If the variance of error exceeds 10 per cent of the variance in the material as a whole for the landmarks in question, then the applied method of measuring is inappropriate.

In this investigation the 3 per cent limit has been exceeded only when measuring n-sm and n-ss. This corresponds well with what has been shown in Tables I and II, viz., a large difference of the mean value when reproducing the landmarks sm and ss.

Even though most of the landmarks studied in this investigation have not been exactly reproducible, the reproducibility should nevertheless be considered as sufficient for all landmarks with the exception of orbitale. This landmark has on average deviated 2.08 mm and 2.44 mm on two occasions of registration.

The reproducibility of the examined landmarks seems to be acceptable for routine analysis of lateral cephalometric roentgenograms when used as a means of diagnosis. On the other hand it is doubtful whether the exactitude of the reproducibility is satisfactory when the landmarks are to be used as a base for forecasts of growth and treatment in data-cephalometric analysis.

Summary

The purpose of this investigation has been to study the reproducibility of fifteen landmarks which are part of the profile analysis carried out as standard procedure at the Orthodontic Department at Örebro and to show the errors of measurements of seven cranial distances.

From the results it can be seen that in the placing of the different landmarks on two separate roentgenograms the observer has been so uncertain that the mean differences of all measuring points have differed significantly from zero. Roughly the same variance in values has been observed in estimating the positions of the landmarks on one and the same roentgenogram on two occasions with an interval of one month. This means that the differences in measurement accounted for have for the most part depended on the uncertainty of the observer in placing the landmarks, while, on the other hand, the error of method in two consecutively taken roentgenograms can be considered as of minor importance.

No significant differences have been demonstrable in comparison between the measurements of two observers on one and the same film. With calculation of cranial distances, the greatest uncertainty has been found in calculating the distances n-ss and n-sm.

Drottninggatan 46
Örebro S-702 22
Sweden

References

1. Baumrind, S. and Frantz, R. C.: The reliability of head film measurements. 1. Landmark identification. *Amer. J. Orthodont*, 60:111, 1971.
2. Björk, A.: The face in profile. *Svensk tandläk.* -T. 40 suppl. 5 B, 1947.
3. Linder-Aronson, S.: Adenoids. Their effect on mode of breathing and nasal airflow and their relationship to characteristics of the facial skeleton and the dentition. *Acta Otolaryng.* Suppl. 265, 1970.
4. Lundström, A.: *Svenska Tandläkare* Sällskapet sektion för ortodonti, jubileumsskrift 1943 - 1968, p. 57, 1968.
5. Richardson, A.: An investigation into the reproducibility of some points, planes and lines used in cephalometric analysis. *Amer. J. Orthodont.*, 52:637, 1966.