A comparison of the accuracy of cephalometric landmark location between two screen/film combinations

By D.R. Stirrups, BA, BDS, FDS, DORTH, FDS, DDORTH

The use of high-speed intensifying screen/film combinations and reduced exposure will affect the clarity of the radiographic image produced. It is not known how far this loss of distinction can be tolerated before it adversely affects the accuracy of cephalometric landmark location and hence the reliability of cephalometric analysis.

Halse and Hedin,[1] when comparing five fast screen/film combinations and a conventional system using a phantom head, found that the five systems examined were inferior to the conventional system because of loss of definition. However, this loss of definition was thought to have little influence on the reliability of cephalometric measurements. Kaugars and Fatouros[2] and Fatouros et al.[3] also found that the use of rare earth screen/film combinations did not adversely affect film quality at the level required for cephalometric analysis. However all these reports were qualitative rather than quantitative and a previous study by McNicol and Stirrups[4] showed that subjective methods of assessment of landmark clarity were not reproducible and that a quantitative method was desirable.

One way of assessing the clarity of a landmark and its effect on the reproducibility of cephalometric measurements is to measure the difference between the localization of a landmark on two separate occasions. A method of doing this by recording landmarks using a digitizer linked to a microcomputer was developed by McNicol.[5] Using a small series of films he calculated location differences for a number of cephalometric landmarks with two observers each locating the landmarks on two occasions. He showed that the method had good inter- and intra-observer reliability. He suggested that the mean location errors for a number of landmarks can be found for each of two screen/film combinations and these mean values used to assess the effects of altering screen/film types on the accuracy of landmark location.

The purpose of this investigation was to use

Abstract

The introduction of faster screen/film combinations allows a reduction in radiation dose but at the expense of film quality. This study investigated the accuracy of landmark location on cephalometric radiographs made using two screen/film combinations. No significant differences in reproducibility of landmark location were found. There was a 20 percent difference in the radiation dose between the two film types studied.

This manuscript was submitted in 1986.

Key Words

Cephalometry • Cephalometric landmarks • Intensifying screens

Table 1

	Standard Films		Fast Films	
	Mean (mm.)	Standard Deviation	Mean (mm.)	Standard Deviation
Superimposition Error	0.012	0.009	0.014	0.010
Nasion	0.855	0.644	0.908	0.754
Sella	0.853	0.732	0.871	0.758
A point	0.997	0.555	1.026	0.721
Menton	0.939	0.512	0.961	0.672
Upper Incisor Tip	0.748	0.485	0.751	0.580
Lower Incisor Apex	1.103	0.919	1.222	0.904

$T^2 = 6.18$ $F = 0.9957$ Probability $= 0.431$

Table 1
Summary statistics.

McNicol's method to compare the location accuracy of six landmarks (Table 1) on radiographs made using either a rare earth screen/film combination or a conventional Barium/Strontium sulphate screen/film combination, the former requiring 20 percent less radiation for adequate exposure.

Material and method

Lateral cephalograms of 152 children, between eight and 15 years old, were taken in a standard way. After a given date, cephalograms were taken using 3M Trimax 8* rare earth intensifying screen in combination with 3M XUD film on the first 76 children requiring them for routine orthodontic reasons. These radiographs will be referred to as the fast screen films. The remaining 76 were taken immediately prior to the change using a barium/strontium sulphate intensifying screen (Kodak X-Omatic Regular**) in combination with Kodak XRP film. These will be referred to as the standard screen films. The choice of fast screen/film combination and exposure values was made by assessing a range of films exposed using a radiological phantom head.

The fast screen films were made using a setting of 75 kV and mAs of 16. The standard screen films were made using the same kV setting but with a setting of 20 for mAs. These exposures produced radiographs of similar image density. This was assessed using a step wedge as the object and a Leitz ASBA image analyzer*** to measure the differences between the resultant images. Figure 1 shows the resultant images; the figures are the means of ten measurements of arbitrary grey levels. The Leitz instrument is capable of distinguishing 264 grey levels while the human eye can detect only about 50; the small differences between the images are not visually detectable. The radiographs were all developed in the same automatic processor.

On each of the radiographs two pin holes were made a constant distance apart using a jig. The two marks, hereafter referred to as the fiducial points, were orientated so that half the landmarks to be identified lay above the line joining the two fiducial points. This minimized the effect of superimposition error, which is described later. The line between the fiducial points was used as the reference line for the recorded positions of the landmarks. By using a back illuminated electronic digitizer (GTCO DIGI-PAD 5**** with a resolution of 0.001 inch/0.0254 millimeters) under standard conditions the positions of the landmarks and fiducial points were recorded on a microcomputer as X and Y coordinates. The fiducial points were recorded twice, before and after recording the landmarks. The recording method automatically rejected both the fiducial points if their coordinates differed by more than 0.2 millimeters between the two occasions and the digitization was repeated. Each film was digitized on two occasions two weeks apart.

The coordinates of all points and landmarks recorded on the second occasion were numerically adjusted so that the reference line used for the landmarks located on the second occasion was superimposed on the reference line used on the first occasion with the mid-points of the lines coincident. For each landmark, the straight line distance between the first recorded coordinates and the second adjusted coordinates was calculated. This distance, measured in millimeters, will be referred to as the location difference.

For each radiograph, location differences were calculated for each of the six landmarks. The location difference for either fiducial point gives a measure of the superimposition error. The mean and standard deviation of these location differences were calculated for each landmark and the second fiducial point for each screen/film type.

For this study, an increase in mean location difference for any of the chosen landmarks between the standard and fast screen/film types of more than 0.5 millimeters was unacceptable.

* *3M Health Care, Loughbrough, England*
** *Eastman Kodak Company, Rochester, New York*
*** *E. Leitz Instruments Ltd., Luton, England*
****GTCO, Rockville, Maryland*

To avoid the possibility of one single landmark being adversely affected by the type of screen/film combination, six landmarks were used to assess the possible differences between the film/screen combinations. The statistical test chosen was Hotelling's T^2 which is a generalization of the Students '*t*'-test to a situation involving more than one variable. A 95 percent confidence that this was a true difference was required as was a 90 percent confidence that if such a difference existed it would be detected by the study. Means and standard deviations of location errors for the same landmarks for the standard films from the study by McNicol[5] was used in the calculation of the number of radiographs required to fulfill these requirements.

Results

The data are summarized in Table 1 with the calculated value of T^2, the value of the corresponding F statistic and its significance.

The significance of T^2 value is determined using the F statistic where

$$F=\{[N_1+N_2-p-1]/[(N_1+N_2-2)p]\}T^2$$

where p is the number of variables (the six landmarks), N_1 and N_2 are the number in each sample (the 76 radiographs in each group) and the significance of the value of F is determined for p and (N_1+N_2-p-1) degrees of freedom.

The null hypothesis of no difference in mean landmark location differences between the two screen/film combinations cannot be rejected.

A posterori calculation of the power of the test[o] showed that there was at least a 90 percent probability of detecting 0.5 millimeter differences in the means for landmark location difference between the film types with a five percent chance of erroneously rejecting the null hypothesis.

The frequency distributions of the location differences for the landmarks were plotted as histograms with Figure 2 as an example. This shows a skewed distribution; similar distributions were found for all landmarks.

Discussion

Although the mean location differences were greater for all landmarks on the fast screen films, the magnitude of these differences was small and did not approach the 0.5 millimeters considered clinically unacceptable. It is therefore not surprising that no statistically significant difference was found between the two film/screen combinations since the experimental method was not designed to detect differences of the level found.

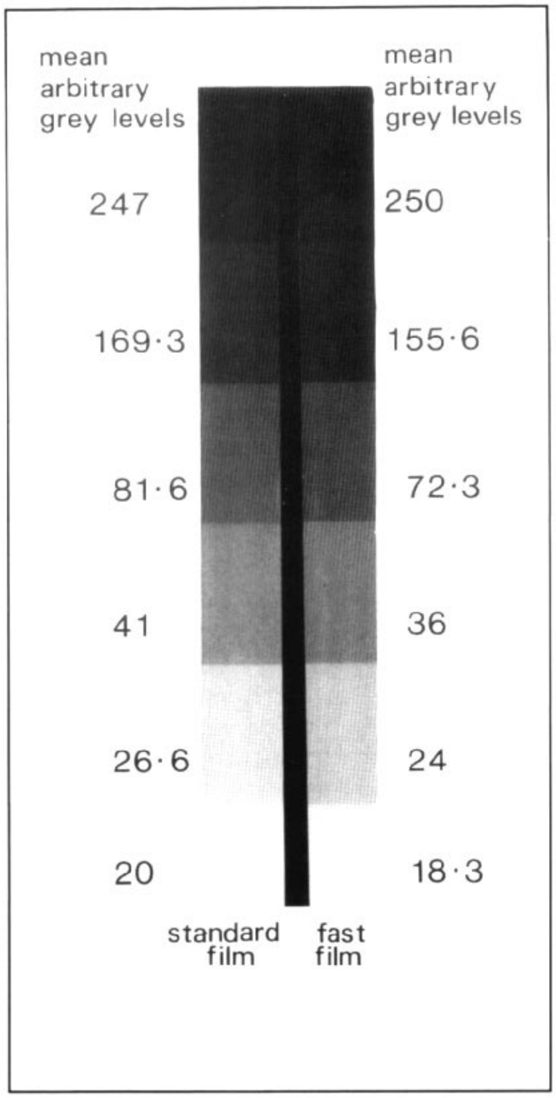
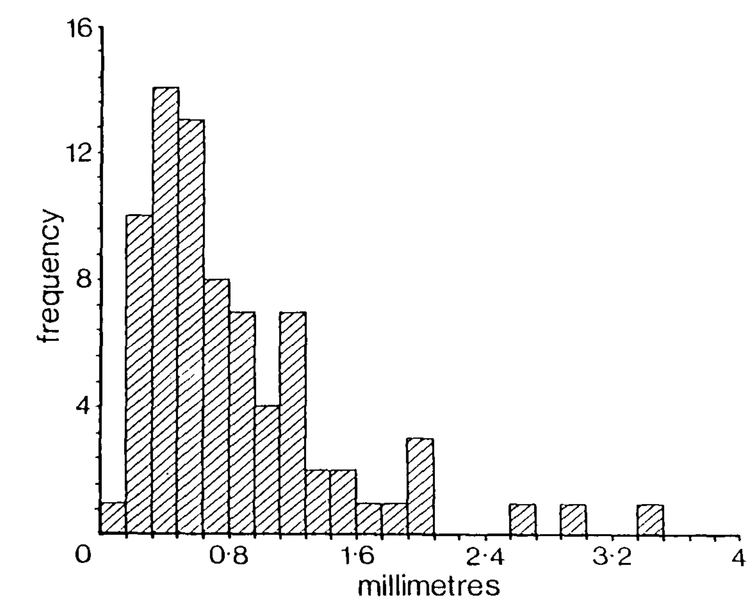


Figure 1
Comparison of image densities for the two film/screen combinations.

Figure 2
Histogram showing absolute differences in locating nasion on the standard films (in millimeters).

Figure 2



Factors which may have contributed to the size of the location differences are:
- the skill of the observer;
- errors in the recording process;
- the clarity of the landmark.

The last of these would have depended upon:
- the size of the focal spot;
- the characteristics of the landmark;
- blurring of the structures caused by secondary radiation or movement during exposure;
- screen sharpness and radiologic mottle (which are directly related to the screen/film combination used);
- the kilovoltage used as this affects contrast and thus image definition;
- image density which was dependent on the contrast and latitude of the film/screen combinations (since the kV was constant). As Figure 1 shows, this was acceptably close at the exposures used for this investigation. Increasing the kV to 90 for the rare earth screen/film combination may have allowed use to be made of the steeper response of the gadolinium phosphors to higher kV's. A further reduction in radiation dosage would result from the reduction in mAs, while comparable image density would be maintained. With extreme reductions in exposure, the effect of quantum mottle may become important.

These factors, except those related to the screen/film combinations, will have affected both sets of radiographs equally. By using two groups of 76, the anatomic variation between the two groups should be random and the net differences minimal. Two radiographs, one of each screen/film combination of the same 76 patients would have ensured no anatomical difference between the two groups; this was unacceptable because of the additional radiation involved. Thus the only significant factor which would have affected the size of the location differences and which was not common to both sets of films was the screen/film combination used.

Both sets of location differences would be affected equally by the skill of the observer and this will influence the size of the mean values of landmark location error and possibly also the difference in mean values between screen/film type. This "skill" effect is probably not significant since McNicol[5] showed that the method used in this study had good intra- and interobserver reproducibility.

Figure 1 was an attempt to investigate directly the influence of the rare earth screens on the contrast and latitude of the radiographs. These two factors will have affected the final radiographic image. The effects of any differences were also assessed indirectly by the measurement method.

As well as being of benefit to the patient, reduced exposure allows either a reduction in the current or the exposure time. The first allows the use of a smaller focal spot, hence a sharper image whereas the second reduces the risk of blurring from subject movement. Rare-earth screens also allow the use of high definition films without an increase in patient exposure. The amount of exposure reduction that is possible with the use of rare-earth screens, before the image definition becomes significantly affected, remains to be determined. Such an investigation would be difficult since there is no agreement about the degree of landmark location error which is considered acceptable. The arbitrary approach used in this study was to assume that the location errors on the currently used (standard) films were acceptable and that an unacceptable increase would be 0.5 millimeters in the mean value for any landmark. Other magnitudes of difference, power of the test or choices of landmarks can be used with the method reported in this study. Increasing the number of landmarks used will greatly increase the numbers of radiographs required. Tables used for calculating the power of the test are only available for up to a maximum of eight landmarks.

The choice of Hotelling's T^2 as the statistical test requires some justification and explanation. Using separate 't' tests for each of the six landmarks leads to a multiple comparison problem. Using the five percent level of significance for 't' tests on each of the variables independently gives the probability of falsely identifying at least one of the means as significantly different of between 0.05 (if all the means were perfectly correlated) and 0.23 (the error rate if all six are independent). Hotelling's T^2 statistic is a method of testing the null hypothesis that the two populations from which two groups are sampled do not differ in their means on any of the measures. The comparison of the two groups can be made with a known maximum probability of a false rejection of the null hypothesis which is chosen in advance (in this study 0.05) and forms part of the null hypothesis. For up to eight variables the power of the test (i.e. the probability of not rejecting the null hypothesis when a true difference of the means actually exists) can be determined and in this study there was a 90 percent chance of detecting differences in the means of 0.5 millimeters.

Hotelling's T^2 involves an assumption of multivariate normality for the variables and as Figure 2 shows, even univariate normality did not apply. However, for sufficiently large sample sizes computed T^2 values conform to the F distribution no matter what the shape of the parent population.[7] In calculating T^2 the assumption is made that both samples have covariance matrices that are random variations of a common population covariance matrix. Rejection of the null hypothesis could be due to inequalities in the covariance matrices of the samples rather than, or in addition to, the differences in the means. Fortunately, Ito and Schull[8] have shown that the true significance level of T^2 is unaffected by discrepancies between the covariance matrices provided reasonably large equal sized samples are used. Since this study had 76 in each sample, the lack of multivariate normality or possible inequalities of the covariance matrices should not affect the validity of the results.

Conclusions

The results show that the use of the selected rare earth screen/film combination did not significantly affect the reproducibility of the chosen cephalometric landmark locations. The exposure required by the rare earth screen/film combination was 20 percent less than that required by the conventional barium/strontium sulphate screen/film combination although this could probably be reduced further by using 90 rather than 75 kV.

This study confirms that rare earth screen/film combinations can be used to reduce radiation exposure during cephalometric radiography without significant adverse effects on the accuracy of landmark location.

The method described can be adapted to study the effect of changing screen/film combinations on any method of measurement using landmarks on cepahlometric radiographs.

Acknowledgements

Dr. R. Strang, Senior Medical Physicist, for the measurements in Figure 1, and the Departments of Radiology and Medical Illustration for their assistance.

Author Address

Dr. D.R. Stirrups
Consultant Orthodontist
Orthodontic Department
Glasgow Dental Hospital and School
378 Sauchiehall Street
Glasgow G2 3JZ
Scotland

Dr. Stirrups is a consultant orthodontist at the Glasgow Dental Hospital and School in Glasgow, Scotland.

References

1. Halse, A. and Hedin, M.: May fast screen/film combinations be used for cephalometry. Brit. J. Orthod. 5:189-192, 1978.
2. Kaugars, G.E. and Fatouros, P.P.: Clinical comparison of conventional and rare earth screen/film systems for cephalometric radiographs. Oral Surgery 53:322-325, 1982.
3. Fatouros, P.P., Gibbs, S.J., Skubic, S.E. and Rao, G.U.V.: Imaging characteristics of new screen/film systems for cephalometric radiography. Angle Orthod. 54:36-54, 1984.
4. McNicol, T. and Stirrups, D.R.: Assessment of screen/film combinations for cepahlometric radiography. Brit. J. Orthod. 12:117-121, 1985.
5. McNicol, T.: Reduction in Radiation Hazard during Orthodontic Radiography. MSc Thesis University of Glasgow, 1984.
6. Morrison, D.F.: Multivariate Statistical Methods, 2nd. edition. New York McGraw-Hill, pp160-164, 1976.
7. Harris, R.J.: A Primer of Multivariate Statistics. Academic Press London, 1975.
8. Ito, K. and Schull, W.J.: On the robustness of the T^2 test in multivariate analysis of variance when variance-covariance matrices are not equal. Biometrika 51:71-82, 1964.