Seminars in
ORTHODONTICS

Reliability of Digital Versus Conventional Cephalometric Radiology: A Comparative Evaluation of Landmark Identification Error

Scott R. McClure, P. Lionel Sadowsky, André Ferreira, and Alex Jacobson

The purpose of this study was to determine whether direct digital lateral cephalometric radiographs are of equal value in diagnosis and treatment planning as conventional cephalometric radiographs by investigating differences in landmark identification on direct digital and conventional cephalometric radiographs. An evaluation of precision and the distribution of landmark identification error at each cephalometric landmark was undertaken. Ten observers, all orthodontists or postgraduate orthodontic residents, identified 19 landmarks twice on 6 digital images and 6 conventional cephalometric films obtained from the records of 6 patients at the University of Alabama School of Dentistry Graduate Orthodontic Clinic. Patient records selected were of adults with existing pretreatment conventional cephalometric films and posttreatment direct digital cephalometric images on file. Landmark identification recordings were transferred into a standardized coordinate system, adjusted for magnification differences, and evaluated separately along both the x- and the y-coordinates. Statistically significant differences in landmark identification error were found along the x-coordinate for A point (Apt) and along the y-coordinate for anterior nasal spine (ANS) and condylion (Co). These statistically significant differences, as well as those found to be not statistically significantly different, were all below 1 mm, indicating that even the statistically significant differences between the two methods of image acquisition were unlikely to be of clinical significance. Each landmark exhibited its own unique pattern of landmark identification error, the magnitude and distribution of which must be taken into consideration when selecting landmarks for use in cephalometric analysis or when interpreting these analyses for diagnostic or treatment planning purposes. The results of the present study indicate similar precision and reproducibility in landmark identification using both direct digital images and conventional lateral cephalometric head films.
Semin Orthod 11:98-110

In 1922, Pacini[1] described a rather primitive method for the standardization of radiographic imaging of the head. He recommended the positioning of a subject at a fixed distance of 2 m from the x-ray source with a film cassette fixed to the head with a wrapping of gauze bandages. Almost a decade later, Hofrath[2] of Germany and Broadbent[3] of the United States simultaneously published their own methods of obtaining standardized head radiographs. Their methods, published in the journals *Fortschritte der Orthodontie* and the *Angle Orthodontist* in 1931, introduced the field of cephalometry to the orthodontic community.

Cephalometric radiography is a vital clinical tool in orthodontics for evaluation of the craniofacial complex, determination of morphology and growth, diagnosis of anomalies, forecasting future relationships, planning treatment, and evaluating the results of growth and the effects of treatment.[4] Cephalometrics remains the only practical quantitative method that permits the investigation and examination of the spatial relationships between both cranial and dental structures. The lateral cephalogram provides information regarding skeletal, dental, and soft tissue morphology as well as relationships between these structures.

As a research tool, cephalometrics has been the most widely used imaging modality reported in the orthodontic literature. The range of possible uses of the lateral cephalo-

From the Department of Orthodontics-Graduate Clinic, University of Alabama at Birmingham, Birmingham, AL.
Address correspondence to Dr. P Lionel Sadowsky, University of Alabama at Birmingham, Department of Orthodontics-Graduate Clinic, 305 School of Dentistry Building, 1919 Seventh Avenue South, Birmingham, AL 35294-0007; E-mail: sadowsky@uab.edu

1073-8746/05/$-see front matter
doi:10.1053/j.sodo.2005.04.002

Table 1 Descriptive Statistics: Sample Demographics

Sample	N	Mean Age at T_1	Mean Age at T_2	Mean Treatment Time (T_2−T_1)
Males	3	28 y 8 m	30 y 8 m	24.0 m
Females	3	23 y 7 m	25 y 4 m	20.7 m
TOTAL	6	26 y 2 m	28 y 0 m	22.3 m

N, number of patients; T_1, pretreatment; T_2, posttreatment.

metric radiograph in research includes: quantifying craniofacial parameters in individuals as well as populations of individuals, distinguishing normal from abnormal, comparing treated samples to untreated controls, differentiating populations as homogenous or mixed, and assessing patterns of change over time.[4] With such a multitude of useful applications, it is no wonder that cephalometrics, despite its various limitations, has withstood the test of time in its usefulness in orthodontics.

Since the advent of modern cephalometric radiography in 1931, the conventional process of cephalometric image capture has remained largely unchanged. Digital radiography has been widely accepted for use in the field of medicine; however, it was not until the1980s that the first intraoral sensors were developed for use in dentistry.[5] The introduction of extraoral digital radiography was initially delayed due to the high cost of extraoral systems. Recently, the development of cost-effective extraoral digital technology, coupled with an increased utilization of computers in orthodontic practice, has made direct digital cephalometric imaging a valid option.[5] As a result, an increasing number of conventional film-based radiographic units are being replaced by direct digital machines. Direct digital images can be acquired through the use of photostimulable phosphor plates, or charge coupled device receptors, both of which offer a number of advantages over film. These advantages include instantaneous image acquisition, reduction of radiation dose, facilitated image enhancement and archiving, elimination of technique sensitive developing process and its associated costs, and facilitated image sharing.[4,5]

Before a complete shift from the widely accepted film-based to digital cephalometric radiology can occur, the accuracy of landmark identification utilizing these two different image acquisition methods must be compared. Studies have reported comparisons of digitized (scanned) and conventional (film) cephalometric radiographs regarding accuracy of landmark identification and the linear and angular measurements.[6-9] Comparisons of film-based and storage phosphor plate digital images have also been reported.[10-13] Landmark identification using charge coupled device direct digital cephalometric images has been compared with the landmark identification using the film-based counterpart in a sample of dried skulls.[14]

The purpose of this study was to determine whether direct digital lateral cephalometric radiographs are of equal value in diagnosis and treatment planning as contemporary film-based cephalometric radiographs by investigating any differences in landmark identification between these two acquisition methods. An evaluation of precision as well as the distribution of landmark identification error at each cephalometric landmark was performed. A sample of patients having undergone treatment at the University of Alabama Postgraduate Orthodontic Program was utilized. Corresponding film and digital cephalometric units of the same manufacturer were used for image acquisition. The null hypothesis was that there would be no differences in landmark identification when comparing the two modalities.

Materials and Methods

A parent sample encompassing the patient record files in the Department of Orthodontics at the University of Alabama was investigated for cephalometric radiographs to be utilized in the present study. Inclusion criteria for the present study were: 1) adult patients with no anticipated remaining growth, 2) patients must have had a pretreatment cephalometric radiograph produced by the Orthopantomograph OP 100 conventional film-based cephalometric machine (Instrumentarium, Munich, Germany), 3) patients must have had a posttreatment cephalometric radiograph produced by the Orthopantomograph OP 100-D digital cephalometric machine (Instrumentarium), and 4) all landmarks to be examined must have been within the confines of the image and available for identification within both the pretreatment head films, and posttreatment images. Exclusion criteria used for radiograph selection were: 1) obvious malposition of the head within the cephalometer, 2) unerupted or missing incisors, or 3) unerupted teeth overlying the incisor apices. Twenty-one (21) sets of patient records satisfying the inclusion criteria were identified and reviewed by three of the authors.

After applying the exclusion criteria to these 21 pairs of radiographs, 6 patient records remained, all of which were selected for use in the present study. The resulting sample consisted of six adult patients, with an average age of 26 years 2 months at the time of acquisition of the pretreatment cephalometric radiograph (range: 21 years 1 month to 29 years 8 months), and the average age at the posttreatment digital cephalometric image acquisition was 28 years 0 months (range: 23 years 3 months to 31 years 4 months). The sample consisted of three males and three female patient images (Table 1).

Nineteen (19) commonly used cephalometric landmarks were included in this analysis. An agreement was reached on the definitions of landmarks before carrying out this study, and these written definitions for each landmark were reviewed with and provided to the 10 evaluating participants (Fig 1, Table 2).

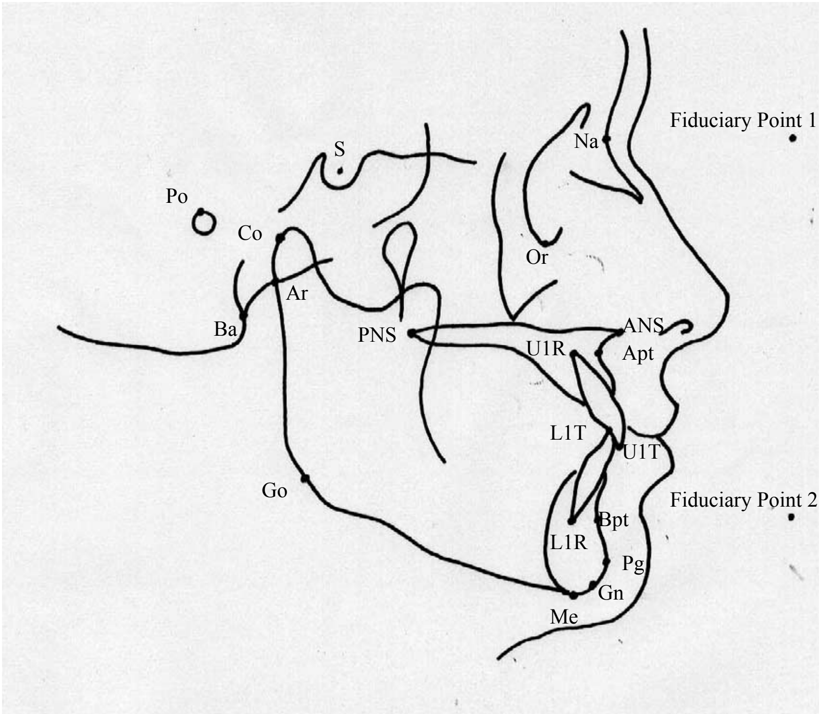

Figure 1 Lateral cephalometric landmarks and fiduciary points.

Five (5) orthodontic residents and five (5) members of the orthodontic faculty at the University of Alabama Department of Orthodontics were asked to identify the landmarks on the 6 posttreatment digital images and their corresponding 6 pretreatment radiographic films. There was an intervening time interval of at least 1 week between each manual tracing and/or digital tracing session. This entire process was repeated a second time for both the pretreatment film and the posttreatment digital image with at least 1 week of time between each landmark identification session. All manual and digitized landmark identification sessions took place in a dimly lit room without interruption for as long as each investigator needed to complete each landmark identification session.

For the manual tracing of the conventional lateral cephalometric films, landmark identification was performed on a light box designed specifically for this purpose, using acetate paper and the same mechanical pencil with 0.5-mm graphite for all investigators. Investigators were allowed to change any landmark positions during each session for landmarks identified during that session only. The tracings of the conventional film-based cephalometric radiographs were scanned into a digital image JPEG format using an Epson 3001 flatbed scanner at settings recognized as acceptable in recent literature.[7,8]

For the digital images, landmark identification was performed directly on a flat panel high resolution monitor-displayed image with a mouse-controlled cursor in connection with the computerized program for imaging and cephalometric analysis (Dolphin Imaging 10.0). After recording a landmark with the mouse, a dot on the monitor-displayed image indicated its position. The landmark position could be corrected until the operator was satisfied at that session. The end result included 2 tracings of each film/digital image pair for each of 6 patients, by each of the 10 participating orthodontists (5 faculty and 5 residents).

Following the landmark identification by the 10 faculty/resident participants, the principal investigator (SRM) standardized the position of the two fiduciary points on each of the digitized digital images and each of the scanned film images to allow for more precise superimposition and measurement using these points. The digital cephalometric unit used in this study was outfitted with a stainless steel wire with bends placed at a known distance of 100 mm within the field of radiographic exposure. This wire projected a radio-opaque image onto the digital image created by the digital cephalostat. The locations at which the wire was bent, measuring 100 mm apart, were identified and marked digitally as the two fiduciary points for these digital images. For the conventional cephalometric films, two pin holes, also at a known distance of 100 mm, were punched through the original six cephalometric films themselves before any landmark identification began. These pinholes were then transferred through each acetate before each landmark identification session to establish two fiduciary markers consistent between manual tracings. Following the digitization (scanning) of the acetate film tracings, these fiduciary points were also standardized by the principal investigator.

The positions of the identified landmarks were recorded in the format of x, y-coordinates. The origin of x, y-coordinates was oriented with the x-axis constructed by vertically connecting the two fiduciary points and a calculated y-axis perpendicular to this x-axis through the more inferior of these two fiduciary points (Fig 2). The origin of the x, y-coordinate

Table 2 Lateral Cephalometric Landmark Definitions*

Landmark	Abbreviation	Definition
A point	Apt	Deepest point of the curve of the maxilla, between anterior nasal spine (ANS) and the dental alveolus.
Anterior nasal spine	ANS	The tip of the anterior nasal spine.
Articulare	Ar	Posterior border of the neck of the condyle.
B point	Bpt	Most posterior point in the concavity along the anterior border of the symphysis.
Basion	Ba	Most inferior posterior point of the occipital bone at the anterior margin of the occipital foramen.
Condylion	Co	Most posterior superior point of the condyle.
Gnathion	Gn	Midpoint between the most anterior and inferior point on the bony chin.
Gonion	Go	Most convex point where the posterior inferior curve of the ramus meet.
Mandibular central Incisor root apex Incisor root apex	L1R	Root apex of the mandibular central incisor.
Mandibular central Incisor tip Incisor tip	L1T	Incisal tip of the mandibular central incisor.
Menton	Me	Most inferior point of the symphysis.
Nasion	Na	Intersection of the internasal suture with the nasofrontal suture in the midsagittal plane.
Orbitale	Or	Lowest point of the roof of orbit; most inferior point of the external border of the orbital cavity.
Pogonion	Pg	Most anterior point on the mid-sagittal symphysis.
Porion	Po	Highest point of the ear canal; most superior point of the external auditory meatus.
Posterior nasal spine	PNS	Tip of the posterior nasal spine.
Sella	S	Center of the pituitary fossa of the sphenoid bone.
Upper central incisor Root apex	U1R	Root apex of the maxillary central incisor.
Upper central incisor Tip	U1T	Incisal tip of the maxillary central incisor.

*This table lists the static lateral cephalometric landmarks and the definitions agreed upon by the investigators in this study. For an illustration of these landmarks, see Figure 1.

grid was therefore arbitrarily selected as the lowermost fiduciary point. The landmarks' locations on the digital images and the transparent acetate films could then be described by using x, y-coordinates with the aid of the computerized program described previously. This procedure was performed by the principal investigator only, to prevent introducing additional random error.

For each landmark, placement differences between tracings of identical radiographs or digital images were assessed by transforming the two sets of x, y-coordinates with identical fiduciary reference points. The x-coordinate and y-coordinates were further analyzed to evaluate the pattern of recording differences in horizontal and vertical directions. Moreover, the mean position for each of the 19 landmarks identified by the 10 observers was defined as the "gold standard" in this study. The gold standard was used to determine the average error at each landmark in both modalities of original film and digital image. The mean distance in millimeters between the gold standard and the 20 locations identified by the 10 observers at two different time points was defined as the mean interobserver error, which was used as the variable determining reliability for each landmark. Consequently, the reliability of landmark identification in each of the two modalities (original radiograph and digital image) could be compared as the differences in magnitude of these distances from the mean, between imaging modalities.

The gold standard average position for each landmark was also used to facilitate accurate superimposition in the creation of scattergrams for each landmark. By superimposing an individual landmark's scattergrams from six different patients on these average positions for a landmark, a composite scattergram was constructed. These 19 composite scattergrams (one for each landmark) incorporated identification error from six different patients' images/films into one graphical representation of the landmark identification error at each of the 19 landmarks examined in the present study.

Results

The mean landmarking errors in the x and y directions and the distribution of landmark registration for film and digital landmark identification are shown in Tables 3 and 4, respectively. All the mean errors made by film tracing were larger than those made by digital identification except for L1R, Pg,

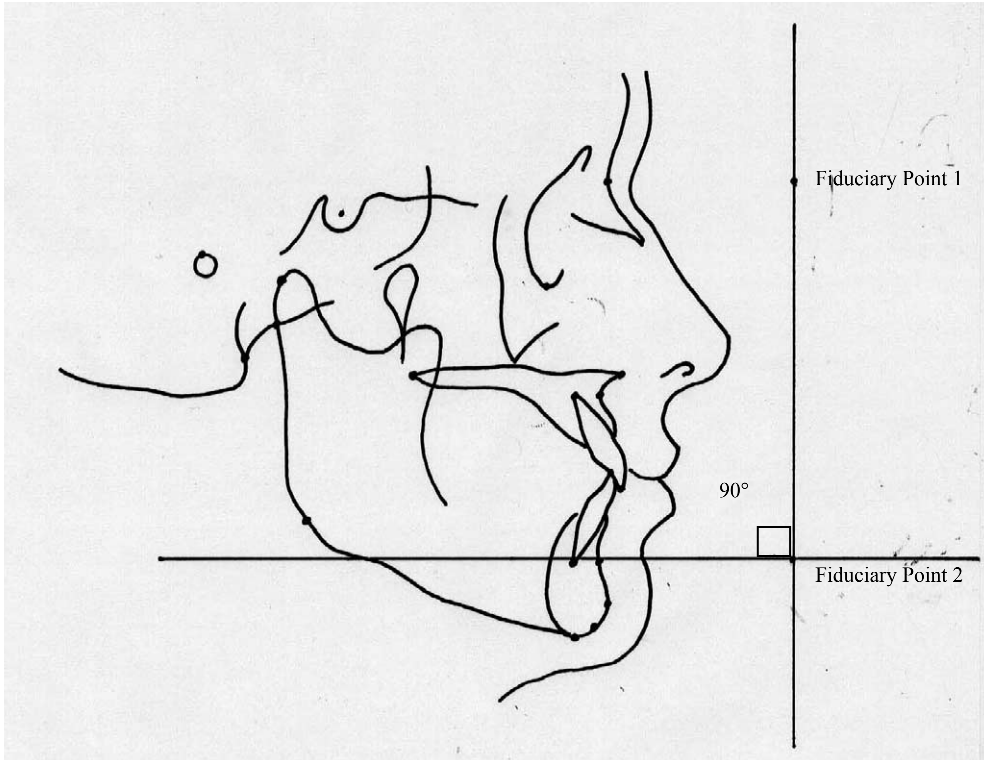

Figure 2 Construction of (x, y) coordinate axis. The origin of the x, y-coordinate grid was therefore arbitrarily selected as fiduciary point 2.

and PNS along the x-axis and Ba, L1R, Or, Pg, and Po along the y-axis. Mean differences were all below 1 mm. In both film and digital methods, U1T was the most accurately identified landmark, and Ba was the least. Of the skeletal landmarks, Sella was the most accurately defined landmark in both film and digital methods.

The overall differences between errors made by film and digital identification of all landmarks considered as a whole were not statistically significant ($P < 0.05$). However, according to paired *t*-tests comparing the scatter for each individual film landmark to the corresponding scatter for the same digital landmark, precision differed significantly between the digital and the film-based images along the y-coordinate for the following landmarks: ANS ($P = 0.001$) and Co ($P = 0.002$). The x-coordinate of A point ($P = 0.040$) also differed significantly (Table 5). The difference between mean film and mean digital error along both the x- and y-coordinates were all below 1 mm. Differences in precision are graphically illustrated as scattergrams in Fig 3. These scattergrams depict the raw data points plotted with the origin of each scattergram set at the mean of all points composing the scatter for each landmark.

Table 3 Mean Film Landmark Identification Errors and Standard Deviations (mm)

Landmarks	Type of Distribution	Mean $Error_x$ (SDx)	Mean $Error_y$ (SDy)
A	O	0.86 (0.57)	1.13 (0.94)
ANS	—	1.79 (1.33)	0.56 (0.43)
Ar	O	0.99 (0.99)	2.18 (2.59)
B	I	0.22 (0.14)	0.92 (0.82)
Ba	O	2.95 (3.30)	2.59 (1.83)
Co	O	1.66 (0.84)	1.84 (1.21)
Gn	⁄	0.73 (0.41)	0.71 (0.42)
Go	\	2.76 (1.86)	2.23 (1.40)
LIR	⁄	0.91 (0.69)	1.82 (1.38)
LIT	O	0.35 (0.27)	0.41 (0.27)
Me	—	1.31 (0.72)	0.46 (0.28)
Na	O	0.48 (0.33)	0.75 (0.70)
Or	O	2.10 (1.54)	1.23 (1.08)
Pg	I	0.25 (0.16)	1.00 (0.76)
PNS	—	1.85 (1.64)	0.55 (0.33)
Po	O	1.48 (1.49)	2.17 (1.75)
S	O	0.46 (0.24)	0.86 (0.65)
UIR	\	1.05 (0.65)	1.40 (1.10)
UIT	O	0.30 (0.26)	0.28 (0.14)

SDx, Standard deviation for error in horizontal direction; SDy, standard deviation for error in vertical direction; O, circular pattern; —, primarily along horizontal axis; I, primarily along vertical axis; ⁄, diagonal scatter descending right to left; \, Diagonal scatter descending left to right.

Discussion

Up to the present time, no studies have reported on the possible differences between the accuracy of landmark identification using hand-traced films compared with landmark identification utilizing direct digital cephalometric images using a live patient

Table 4 Mean Digital Landmark Identification Errors and Standard Deviations (mm)

Landmarks	Distribution	Mean $Error_x$ (SDx)	Mean $Error_y$ (SDy)
A	O	0.81 (0.76)	0.79 (0.53)
ANS	—	1.44 (1.28)	0.50 (0.49)
Ar	O	0.56 (0.67)	1.42 (1.92)
B	|	0.17 (0.11)	0.67 (0.74)
Ba	O	2.32 (2.54)	2.89 (2.31)
Co	O	1.41 (1.05)	1.10 (0.72)
Gn	⁄	0.54 (0.30)	0.55 (0.34)
Go	\	2.45 (1.64)	1.99 (1.30)
LIR	⁄	0.94 (0.74)	1.83 (1.71)
LIT	O	0.18 (0.11)	0.27 (0.15)
Me	—	1.02 (0.56)	0.39 (0.25)
Na	O	0.26 (0.22)	0.62 (0.51)
Or	O	1.78 (1.03)	1.61 (1.05)
Pg	|	0.27 (0.24)	1.01 (0.78)
PNS	—	2.07 (1.72)	0.48 (0.37)
Po	O	1.25 (1.06)	2.85 (2.11)
S	O	0.30 (0.20)	0.64 (0.72)
UIR	\	0.83 (0.64)	1.39 (1.34)
UIT	O	0.15 (0.09)	0.23 (0.14)

SDx, Standard deviation for error in horizontal direction; SDy, standard deviation for error in vertical direction; O, circular pattern; —, primarily along horizontal axis; |, primarily along vertical axis; ⁄, diagonal scatter descending right to left; \, Diagonal scatter descending left to right.

sample. Numerous studies have been undertaken to explore differences between landmark points selected from traditional radiographic film compared with the same selected landmarks from the scanned images of the same radiographic films. Direct comparison of the results of the present study with those of previously reported similar studies was considered inappropriate because of differing methods and variables. Comparisons were made between previous studies and the present study regarding the magnitude and distribution of landmark error for individual landmarks. The clinical implications of landmark error and distribution are discussed.

Initially a comparison of landmark identification variability between orthodontic faculty compared with orthodontic residents was planned. The variability between these two groups, however, indicated no significant differences in landmark identification error utilizing either the film or the digital images. Landmark identification variability is reportedly high regardless of the experience of the clinicians.[15] For these reasons, it was decided that further comparison between faculty and residents was not warranted.

It has been reported that landmark identification with an error below 1 mm is considered precise.[16] Based on this criterion only six of the landmarks evaluated in the present investigation (B, Gn, L1T, Na, S, U1T) could be considered accurate along both x- and y-axes, by both digital and hand-tracing identification techniques. U1T was the most accurately identified landmark for both digital and film-based methods [digital error x-axis (dE_x) = 0.15 digital error y-axis (dE_y) = 0.23, film error x-axis (fE_x) = 0.30 film error y-axis (fE_y) = 0.28] falling well within this definition of precision.[16]

A previous study had defined the accepted normal range of most cephalometric measurements as ±2 mm.[17] Using this definition of precision (error <2 mm = precise), all landmarks would be considered precise in the present study with the exception of: Ar in the vertical (fE_y = 2.18 mm), Or in the horizontal (fE_x = 2.10 mm) for film identification, PNS in the horizontal (dE_x = 2.07 mm) for digital identification, Ba (dE_x = 2.32, dE_y = 2.89; fE_x = 2.95, fE_y = 2.59), Go (dE_x = 2.45, dE_y = 1.99; fE_x = 2.76, fE_y = 2.23) in both x- and y-coordinates, and Po (dE_y = 2.85, fE_y = 2.17) in the y-coordinate, for both methods. Of these imprecise landmarks, based on the 2-mm range, Ba proved to be the least accurately identified landmark for both film and digital methods (dE_x = 2.32, dE_y = 2.89; fE_x = 2.95, fE_y = 2.59). The difficulty in locating the landmarks Ba, Or, Po, and Ar may be the result of a blurred image due to the superimposition of adjacent or bilateral structures. The landmark Or was more inaccurate in the horizontal plane, most likely the result of the left and right images of the orbits being more closely aligned vertically than anteroposteriorly. Alternatively, Ar was more imprecise vertically since this landmark is defined as the most posterior point on the neck of the vertically oriented condyle. The tortuous route of the ear canals create multiple vertically overlapping radiolucent structures, which was likely a contributory factor in the inaccuracy of identification of Po in the vertical direction. The uncertainty in identification of Go may result from the difficulty of establishing this landmark's position along a curved anatomical structure. PNS proved to be more reliably identified in the vertical than in the horizontal dimension as this landmark is identified as the posterior limit of the horizontally oriented hard palate. As the hard palate extends posteriorly, this radio-opaque structure fades toward its end as the more radiolucent soft palate becomes evident. As a result, the exact point at which to locate the landmark PNS is obscured in the horizontal dimension, more than in the vertical dimension. The participating orthodontic clinicians were most likely aware of the greater importance of the location of PNS in the vertical and this may have contributed to their desire to be more attentive to the accuracy of this point in this dimension.

Geelen and coworkers[12] reported similar results when examining the precision of certain landmarks, finding both Or and Po to be outside the (<2 mm error) range of precision. These same authors found Ba, Go, and PNS to be approaching imprecision (0.75-1.75 mm error), while Ar was found to be rather precise (<0.75 mm error). Geelen and coworkers[12] did not, however, separate their results into error along the vertical and horizontal axes; therefore, direct comparison between their study and the present study was considered inappropriate. Baumrind and Frantz[18] did delineate differences in error along the x- and y-axes and reported findings similar to those of the present study. In Baumrind and Franz's study, Go was found to be imprecise along both axes, and Or neared imprecision on the horizontal but was rather precise along the vertical. Ba, Ar, and PNS were not examined in their study, and machine Po was utilized. This mechanical landmark was predictably precise.

Trpkova and colleagues[19] and associates performed a

Table 5 Comparison of the Mean Errors for Film and Digital Landmark Identification along X and Y Axes

Landmark	Coordinate	Mean Film Error (mm)	Mean Digital Error (mm)	Difference (Film-Digital)	*P* Value
A	X	0.86 (0.57)	0.81 (0.76)	0.05	0.040*
	Y	1.13 (0.94)	0.79 (0.53)	0.34	0.126
ANS	X	1.79 (1.33)	1.44 (1.28)	0.35	0.996
	Y	0.56 (0.43)	0.50 (0.49)	0.06	0.001*
Ar	X	0.99 (0.99)	0.56 (0.67)	0.43	0.213
	Y	2.18 (2.59)	1.42 (1.92)	0.76	0.456
B	X	0.22 (0.14)	0.17 (0.11)	0.05	0.104
	Y	0.92 (0.82)	0.67 (0.74)	0.25	0.982
Ba	X	2.95 (3.30)	2.32 (2.54)	0.63	0.156
	Y	2.59 (1.83)	2.89 (2.31)	−0.3	0.740
Co	X	1.66 (0.84)	1.41 (1.05)	0.25	0.123
	Y	1.84 (1.21)	1.10 (0.72)	0.74	0.002*
Gn	X	0.73 (0.41)	0.54 (0.30)	0.19	0.234
	Y	0.71 (0.42)	0.55 (0.34)	0.16	0.152
Go	X	2.76 (1.86)	2.45 (1.64)	0.31	0.099
	Y	2.23 (1.40)	1.99 (1.30)	0.24	0.086
L1R	X	0.91 (0.69)	0.94 (0.74)	−0.03	0.870
	Y	1.82 (1.38)	1.83 (1.71)	−0.01	0.485
L1T	X	0.35 (0.27)	0.18 (0.11)	0.17	0.537
	Y	0.41 (0.27)	0.27 (0.15)	0.14	0.125
Me	X	1.31 (0.72)	1.02 (0.56)	0.29	0.842
	Y	0.46 (0.28)	0.39 (0.25)	0.07	0.222
Na	X	0.48 (0.33)	0.26 (0.22)	0.22	0.132
	Y	0.75 (0.70)	0.62 (0.51)	0.13	0.753
Or	X	2.10 (1.54)	1.78 (1.03)	0.32	0.791
	Y	1.23 (1.08)	1.61 (1.05)	−0.38	0.777
Pg	X	0.25 (0.16)	0.27 (0.24)	−0.02	0.124
	Y	1.00 (0.76)	1.01 (0.78)	−0.01	0.051
PNS	X	1.85 (1.64)	2.07 (1.72)	−0.22	0.599
	Y	0.55 (0.33)	0.48 (0.37)	0.07	0.178
Po	X	1.48 (1.49)	1.25 (1.06)	0.23	0.805
	Y	2.17 (1.75)	2.85 (2.11)	−0.68	0.274
S	X	0.46 (0.24)	0.30 (0.20)	0.16	0.727
	Y	0.86 (0.65)	0.64 (0.72)	0.22	0.203
U1R	X	1.05 (0.65)	0.83 (0.64)	0.22	0.381
	Y	1.40 (1.10)	1.39 (1.34)	0.01	0.456
U1T	X	0.30 (0.26)	0.15 (0.09)	0.15	0.823
	Y	0.28 (0.14)	0.23 (0.14)	0.05	0.192

*Denotes statistical significance ($P < 0.05$).

meta-analysis using six pertinent studies to assess the magnitude of cephalometric analysis error along the x- and y-axes, respectively. Their study examined identification error of 15 skeletal landmarks, 14 of which were evaluated in the present study. The results of the meta-analysis performed by Trpkova and colleagues[19] led these authors to recommend that 0.59 mm of total error for the x-coordinate and 0.56 mm for the y-coordinate are acceptable levels of accuracy. Applying these values, they found the landmarks B, A, Ptm, S, and Go to be accurate along the horizontal. Employing the same parameters in the present study, the landmarks Ar, B, Gn, L1T, Na, Pg, S, and U1T (dE_x = 0.56, 0.17, 0.54, 0.18, 0.26, 0.27, 0.30, 0.15, respectively) would be considered accurate for digital identification, and B, L1T, Na, Pg, S, and U1T (fE_x = 0.22, 0.35, 0.48, 0.25, 0.46, 0.30, respectively) accurate when using film-based landmark identification. Trpkova and colleagues[19] found Ptm, A, and S to be accurate vertically. In the present study the landmarks ANS, Gn, L1T, Me, PNS, and U1T (dE_y = 0.50, 0.55, 0.27, 0.39, 0.48, 0.23, respectively) would be considered accurate with digital identification methods, and these same landmarks excluding Gn (fE_y = 0.56, 0.41, 0.46, 0.55, 0.28, respectively) would be considered accurate for film-based identification when applying Trpkova's parameter for accuracy in the vertical dimension. In the present study, when employing Trpkova's standards, landmark identification using digital images had more landmarks that proved to be precise in both x- and y-dimensions than the conventional film-based landmark identification. The latter seems to indicate that digital images allow for as clinically acceptable landmark identification as that obtained by traditional film-based methods.

All mean errors made by film tracing were larger than those

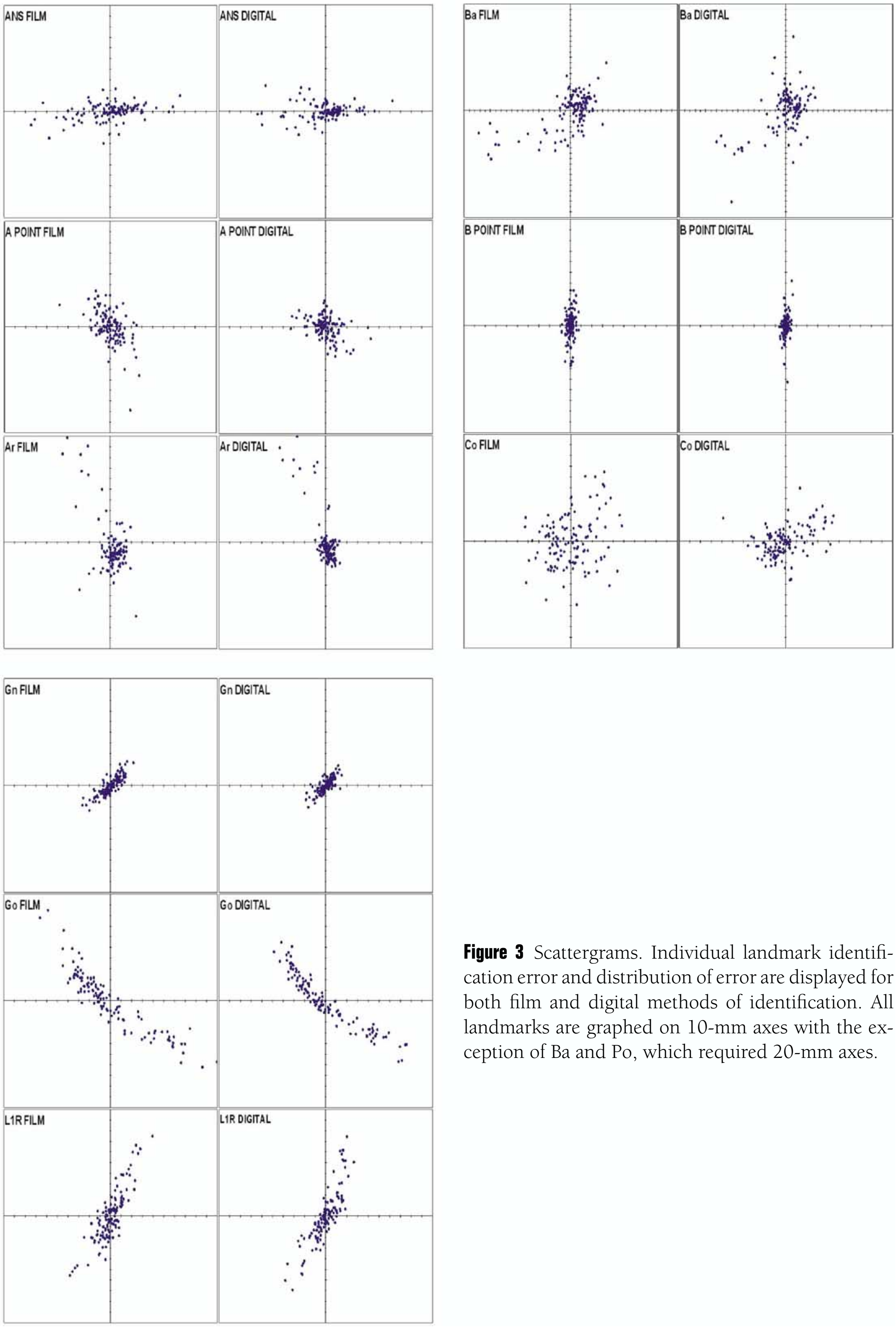

Figure 3 Scattergrams. Individual landmark identification error and distribution of error are displayed for both film and digital methods of identification. All landmarks are graphed on 10-mm axes with the exception of Ba and Po, which required 20-mm axes.

made by digital identification except for L1R, Pg, and PNS along the x-axis and Ba, L1R, Or, Pg, and Po along the y-axis. However, statistically significant differences ($P < 0.05$) between film and digital modalities were revealed only between the digital and the film-based images along the y-coordinate for the following landmarks: ANS ($P = 0.001$) and Co ($P = 0.002$), and along the x-coordinate for A point ($P = 0.040$). While the differences in mean landmark identification error found at these landmarks were statistically significant, none of the differences at any of the 19 landmarks were more than 1 mm. Differences of such small magnitude, though of statistical significance, would not prove to be of equal clinical

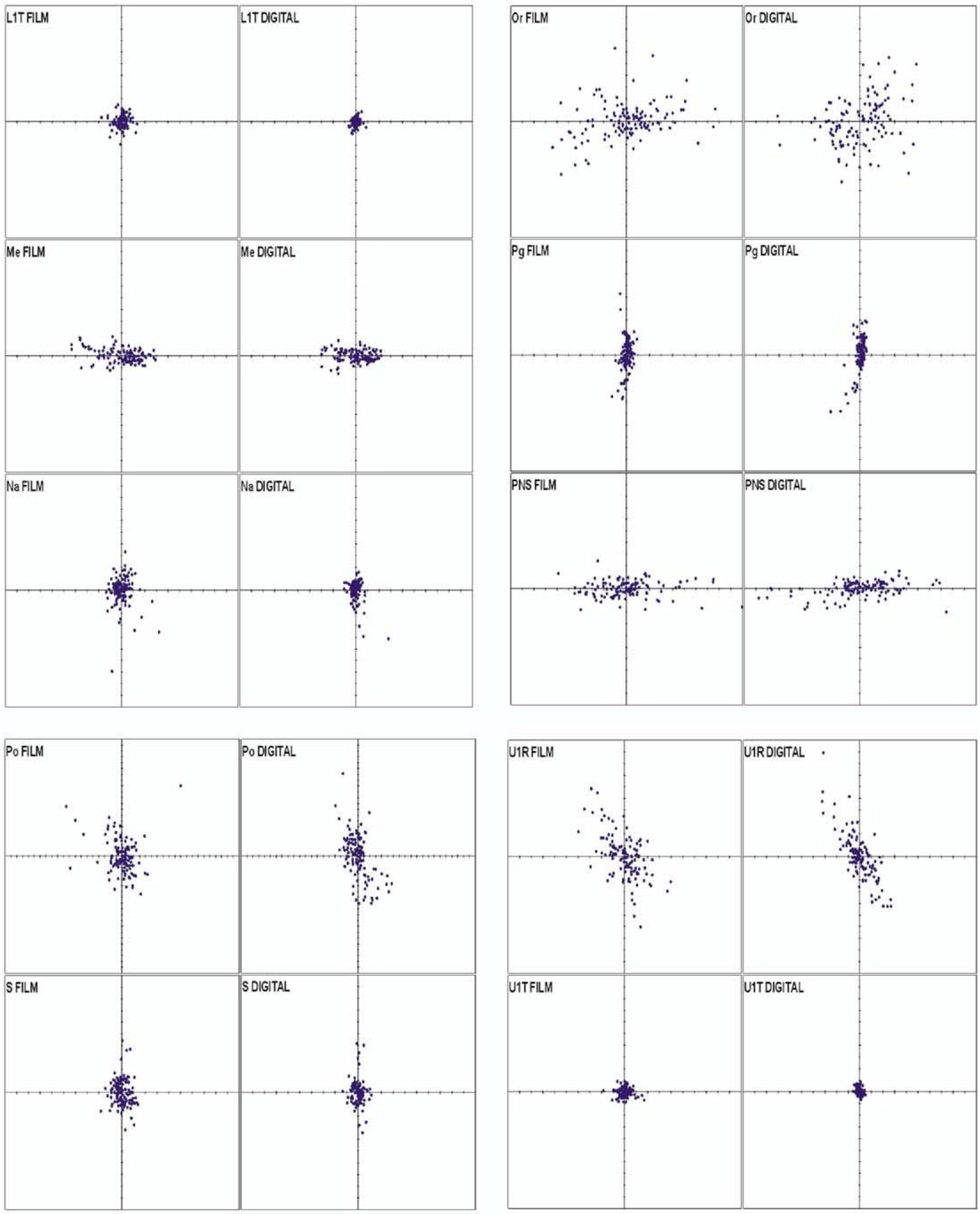

Figure 3 (continued)

significance. These results indicate that the precision of landmark identification with direct digital images is generally comparable to that obtained through film-based identification methods.

Certain cephalometric landmarks have been reported to be more reliable in the horizontal dimension while others are more reliable in the vertical dimension. This is clearly illustrated in the distribution of error seen in the scattergram of each landmark (Fig 3). The reasons for these differences in distribution of landmark identification error are often related to the definition of the landmarks themselves. This often relates to anatomical variability of the landmark location. A sharp incisal edge would likely have less error associated with its identification than a landmark location associated with a more gradual curve as in Apt, Bpt, Go, Gn, Pg, Me, and so on. Errors in the latter would be influenced by the vertical or horizontal orientation of the curve. For example, landmarks such as Me, Go, ANS, and PNS are likely to have more x-axis error than y-axis error. In addition, points A and B, among others would have more error in the y-axis than in the x-axis envelopes of error for the same reasons listed above. As a result envelopes of error vary for each specific landmark.

Another factor entering into the error of cephalometric landmark identification is the degree to which the edge or point to be located contrasts with the surrounding cephalometric structures. For this reason, cephalometric landmarks that lie within the confines of the skull, where numerous radio-opaque structures tend to overlap, will have a greater margin of error in identification than landmarks located on the surfaces of the skull. In the latter there will be a sharp contrast between the radio-opaque skeletal structures and the surrounding radiolucent soft tissue. It should be noted that the superimposition of dental structures such as incisor and often cuspid teeth adds to the difficulty in accurately locating a single incisor root apex. The source of error due to contrast will certainly play a greater role in the landmark identification error evident in the location of landmarks such as Po, Or, Ar, and Co, and root apices (internal), compared with landmarks such as, Na, Bpt, Pg, Gn, Me, and Go and incisal edges (surface).

Finally, the complexity of landmark definition may exhibit a direct adverse effect on the certainty of the landmark identification. A definition for landmark identification that includes terms such as "the tip of" would prove much easier to identify than a landmark with a definition involving the "most inferior posterior point" or the "midpoint between the most anterior and inferior point."

A discussion of the importance of the landmark identification error, and the distribution of this error reported in the present study, would be incomplete without addressing the clinical implications of the scatter at each landmark. The relative importance of the distribution of error for a given landmark is determined by the use of that landmark in various cephalometric analyses. If a landmark is to be used to determine the magnitude of a horizontal discrepancy of the jaws relative to one another in an angular measure such as SNA, SNB, and ANB, then the error of the landmarks A point and B point along the horizontal axis would be of greater significance than the error of these landmarks along the vertical axis. Any change in the horizontal position of A point or B point would result in a significant change in the resulting angular measures SNA, SNB, and ANB. The error of point S, however, would be of greater significance along the vertical axis than the error along the horizontal. The relative importance of error in one direction as opposed to the other varies for each landmark depending on the use of the landmark the cephalometric analysis.

Since cephalometric landmarks are used in combinations with others to assess linear or angular measure, error at each landmark site is of significance. Both the magnitude as well as the distribution of the landmark identification error is of importance when selecting a landmark for use in a cephalometric analysis, and when selecting a cephalometric analysis to arrive at diagnostic conclusions and treatment planning decisions.

In the present study, S has been reported to be a very precisely identified landmark, with error in all directions well below 1 mm for both digital and film methods ($dE_x = 0.30$, $dE_y = 0.64$; $fE_x = 0.46$, $fE_y = 0.86$). The envelope distribution of this landmark identification error for both methods is essentially circular with occasional more extreme values along the vertical axis. The circular pattern of identification error is supported by the work of Richardson[20] as well as Liu and colleagues,[17] both of whom examined the identification error by using manual film-based identification. When S, as a landmark, is being used to construct the angles SNA and SNB, the landmark's error is of greater significance along its y-axis. More extreme values along the vertical axis were found for the landmark S, and these would most certainly increase diagnostic concerns. This concept of outliers and their effects on clinical interpretation of cephalometric analyses has been referred to by numerous authors including Baumrind and Frantz,[21] Stabrun and Danielsen,[22] and Perillo and coworkers.[23]

Nasion (Na) was also very precisely identified in this study ($dE_x = 0.26$, $dE_y = 0.62$; $fE_x = 0.48$, $fE_y = 0/75$), ranking second only to U1T ($dE_x = 0.15$, $dE_y = 0.23$; $fE_x = 0.30$, $fE_y = 0.28$) and closely followed by S ($dE_x = 0.30$, $dE_y = 0.64$; $fE_x = 0.46$, $fE_y = 0.86$). The distribution of Na's identification error for both methods was circular, contrary to that of Richardson[26] and Liu and coworkers.[22] When Na, as a landmark, is being used to construct the angles SNA and SNB, the error is of importance along both the x- and y-axes. A horizontal change in the position of Na has the most obvious effect on the angular values of SNA, SNB, and ANB. The effect of a vertical change in point Na would have an equally significant effect on the angular measures SNA, SNB, and ANB. Therefore, landmark identification error in both the vertical and horizontal directions would be of consequence when determining the reliability of Na in cephalometric analyses. The present study is in agreement with the reports of others[17,20] indicating that Na is a relatively reliable landmark, the use of which introduces little error into cephalometric analyses.

The final two points making up the angles SNA, SNB, and ANB are A point and B point. As both A and B make up the inferior points of the angles SNA, SNB, respectively, as well as the angle ANB, the error recorded for these points along the horizontal axis would be of greater concern than the error in the vertical axis. In the present study, A point displayed a circular distribution of error that was within the range of acceptable error along both the x- and y-axes when identified by digital means ($dE_x = 0.81$, $dE_y = 0.79$). Conventional film-based identification of A point proved to be within the acceptable range along the x-axis only ($fE_x = 0.86$) but beyond the acceptable range for the y-axis (error <1.0 mm; $fE_y = 1.13$). This vertical error, while statistically significant, was not of a magnitude that would prove clinically significant. A study by Liu and coworkers[17] examining film-based landmark identification found a circular envelope of error at A point, supporting the results of the present study. Stabrun and Danielsen,[22] however, found a distribution of error primarily in the vertical direction in contrast with the findings for A point in the present study.

The distribution of the identification error of B point has been determined by the present study to be primarily along the vertical axis and remained within the range of acceptability for both vertical and horizontal dimensions. This accuracy and distribution of B point is in agreement with the findings of both Stabrun and Danielsen[22] and Liu and coworkers.[17] B point has proven to be a reliable landmark because of the accuracy of identification and because variations in the vertical dimension would not pose serious clinical concerns.

The landmark S is often used to construct a line representing the floor of the anterior cranial base (S-Na), from which perpendicular linear as well as angular measurements are made. Ba–Na is another commonly used line to represent the floor of the cranial base. There continues to be much debate as to which of these two cephalometric lines is a more accurate and reproducible reference baseline from which to measure. With Na being common to both of these planes, any error at Na would be common to both of these planes (S-Na and Ba-Na), and therefore the true difference in the reliability of these reference planes lies in any error in S compared with any error at Ba. The distribution of both Ba and S has been shown by the present study to be circular. However, the error reported at Ba ($dE_x = 2.32$, $dE_y = 2.89$; $fE_x = 2.95$, $fE_y = 2.59$) established Ba as the landmark with the greatest identification error of all of the landmarks studied. This magnitude of error at Ba far exceeded that reported at S ($dE_x = 0.30$, $dE_y = 0.64$; $fE_x = 0.46$, $fE_y = 0.86$), which was found to be the most accurate of all the skeletal landmarks in both film and digital landmark identification. As a result, S-Na, as well as any measurements made from or utilizing S-Na, would prove to be more accurate than similar measurements made by using Ba-Na as the baseline.

Error of landmark identification at ANS and PNS are of little consequence because their patterns of error were primarily within the x-axis. The palatal plane is a cephalometric structure constructed by connecting the landmarks ANS and PNS. Palatal plane is primarily used as a cephalometric measure of the vertical location and horizontal orientation of the maxilla in relation to the rest of the craniofacial complex. As a result, error in identification of either ANS or PNS in the vertical dimension would be of greater consequence in the variability of the resulting palatal plane. The present study has found the distribution of error for both ANS and PNS to be essentially along the horizontal axis for both digital and film-based identification methods, a distribution of error that has been repeatedly reported for these landmarks.[17,20,22] The accuracy of identification of ANS and PNS along the horizontal (ANS: $dE_y = 0.50$, $fE_y = 0.56$; PNS: $dE_y = 0.48$, $fE_y = 0.55$) was far from optimal. It is the precision in the vertical direction (ANS: $dE_x = 1.44$, $fE_x = 1.79$; PNS: $dE_x = 2.07$, $fE_x = 1.85$), however, that is of greater concern when attempting to construct a consistent palatal plane using ANS and PNS.

Frankfort horizontal is a reference plane commonly used as an aid in the assessment of the vertical dimension, as well as a plane from which linear and angular measurements are made. The validity of Frankfort horizontal as a cephalometric structure has been investigated by numerous authors[24-26] and the landmark identification error inherent in the landmarks used to construct this plane are of importance in this assessment. Frankfort horizontal is drawn as a line connecting the cephalometric landmarks Po and Or, the locations of which were shown in the present study to be identified with significant imprecision. Po, for both film-based and digital identification methods, displayed a circular pattern of error that was of a magnitude that fell significantly outside the accepted range of precision ($dE_x = 1.25$, $dE_y = 2.85$; $fE_x = 1.48$, $fE_y = 2.17$). Orbitale (Or) also displayed a wide scatter in a circular distribution of error that has likewise been deemed imprecise ($dE_x = 1.78$, $dE_y = 1.61$; $fE_x = 2.10$, $fE_y = 1.23$). In fact, Po and Or are among the least precise landmarks as far as identification error is concerned, being only second and third to last behind Ba, respectively, in the present study. The findings of Liu and coworkers[17] are in agreement with the findings of the present study as it relates to Or; however, these authors found a primarily vertical envelope of error for Po in contrast to the present study.

Another reference plane frequently used in cephalometric analysis is the mandibular plane, constructed as a line connecting the cephalometric landmarks Go and Gn. The landmark identification errors inherent at both of these landmarks (Go and Gn) are in patterns of distribution that result in limited reliability of the mandibular plane as a cephalometric structure. While the error at landmark Gn is within the acceptable range to be considered precise ($dE_x = 0.54$, $dE_y = 0.55$; $fE_x = 0.73$, $fE_y = 0.71$), Gn displays a diagonal scatter descending from right to left (/). At Go there is both a greater magnitude of error, as well as a distribution of this error that likewise increases the amount of this error that is incorporated into the construction of mandibular plane. The error at Go as determined by the present study for both film and digital identification methods is well beyond that which could be considered precise ($dE_x = 2.45$, $dE_y = 1.99$; $fE_x = 2.76$, $fE_y = 2.23$). The distribution of this error is in a diagonal direction descending from left to right (\). A vertical distribution of error at Go was found by Liu and coworkers,[17] along with a circular envelope of error at Gn. It should be mentioned, however, that Liu and coworkers[19] did not differentiate vertical or horizontal error into diagonal distributions. The present study describes this diagonal distribution due to the particular characteristics of these scattergrams. Richardson[20] found a circular pattern of landmark identification error at Go. The magnitude of error seen at Go, as well as the distribution of error at both Go and Gn, call into question the reliability of mandibular plane as an infinitely reliable reference plane commonly used to determine the vertical orientation of the mandible in relation to the remainder of the craniofacial complex.

Lateral cephalometric radiographs are also used to assess the position of dental structures, most importantly, the position of the maxillary and mandibular incisors. The cephalometric position of the incisors often determines whether the overall treatment plan would involve the extraction of teeth. As a result, the reliability of the landmarks that determine the position of the incisors is of great importance to the treatment planning process itself.

Of the four landmarks that define the long axis of the maxillary and mandibular incisors, U1T was reported as the most reliable; in fact, this landmark was identified with the greatest precision of all landmarks, skeletal or dental, for both digital and film-based methods ($dE_x = 0.15$, $dE_y = 0.23$; $fE_x = 0.30$, $fE_y = 0.28$). The distribution of the error at U1T was circular. Therefore, both the precision as well as distribution of the error found at U1T make this landmark an extremely reliable approximation of the incisal tip for measurement purposes and for the construction of the long axis of the maxillary incisor. This determination of U1T as a reliable landmark based on both the precision and distribution of U1T's error is supported by the results reported by Liu and coworkers.[17] UIR, on the other hand, was shown in this study to have identification error beyond the range of precise for the x- and y-axes in film-based identification ($fE_x = 1.05$, $fE_y = 1.40$) and for the y-axis alone for the digital identification ($dE_y = 1.39$). The x-axis for the digital identification was, however, shown to be approaching imprecision ($dE_x = 0.83$). The distribution of error that would have the greatest effect on the construction of the long axis of the maxillary incisors would be one along a diagonal descending from right to left, potentially incorporating the greatest amount of error in the determination of the maxillary incisor's inclination. The distribution found at U1R in the present study was along a diagonal in the opposite direction (descending left to right), a pattern that would result in very little variation in the maxillary incisor inclination. The only effect that an envelope of error in this direction might have is in the estimated length of the maxillary incisor, a variable that is of little concern in cephalometric analysis. The participants in the present study were most likely more concerned with accurately identifying a point along the long axis of the maxillary incisor, rather than identifying the actual root apex of this tooth, resulting in the pattern of error displayed at U1R. This distribution of error will not significantly affect the clinical reliability of the long axis of the maxillary incisor as a cephalometric reference.

Similar concerns are obvious in assessing the mandibular incisors. The landmark L1T is second among all dental and skeletal landmarks in precision ($dE_x = 0.18$, $dE_y = 0.27$; $fE_x = 0.35$, $fE_y = 0.41$) and displays a similar circular distribution to that of U1T. Again, Liu and coworkers[17] have reached similar conclusions as to the precision, distribution, and therefore the reliability of L1T as a cephalometric landmark. The error at L1R is beyond the accepted range of precision for both film and digital along the y-axis ($fE_y = 1.82$, $dE_y = 1.83$) and just within this range for the x-axis in both methods ($fE_x = 0.91$, $dE_x = 0.94$). This finding is supported by the findings of Stabrun and Danielsen,[22] who likewise reported the lower incisor apex to "show a wide scatter." These authors concluded that the lack of certainty in locating apex inferior (L1R) should be taken in to account when using the axial inclination of the lower incisor in diagnosis and treatment planning. The distribution of error at L1R is in a diagonal pattern descending right to left, a distribution that, as in the upper incisor, would affect the length of the long axis of the lower incisor much more than it would the buccolingual inclination of the lower incisor. It is recommended, based on the variability in landmark identification of LIR, that while caution should be exercised in relying too heavily on the axial inclination of the lower incisor alone in diagnosis and treatment planning decisions, the landmarks involved offer a relatively reliable approximation of the incisor's position.

This discussion of the interaction of separate cephalometric landmarks and their associated identification error should be continued through further examination of the results of the present study. Utilization of digital cephalometric technology can be justified only through the repetitive examination of landmark identification showing at least comparable, or preferably improved, accuracy when compared with that of the contemporary film-based cephalometrics.

Conclusions

The purpose of this study was to determine whether digital lateral cephalometric radiographs are of equal value in diagnosis and treatment planning as contemporary film-based cephalometric radiographs. This was determined by investigating differences in landmark identification error between contemporary film and direct digital lateral cephalometric images. An evaluation of precision as well as distribution of landmark identification error at each of 19 commonly used cephalometric landmarks was undertaken and the following conclusions were drawn:

- The 19 landmarks evaluated show various degrees of precision.
- Each landmark has a specific scatter pattern, the magnitude and distribution of which must be taken into consideration when choosing landmarks for use in cephalometric analysis, or when interpreting these analyses for diagnostic or treatment planning purposes.
- A wide landmark identification error distribution along the vertical axis makes a point unsuitable for construction of a horizontal reference plane.
- A wide landmark identification error distribution along the horizontal axis makes a point unsuitable for measurement within the anteroposterior dimension.
- Three of the 19 landmarks indicated statistically significantly higher landmark identification error for film-based identification methods than for digital image based identification.
- None of the differences in landmark identification error between the film-based and digital methods, including the three statistically significant differences, was greater than 1 mm. This indicates that even the statistically significant differences between the two methods of image acquisition were unlikely to be of clinical significance.
- The results of the present study indicate similar precision and reproducibility in landmark identification using both direct digital images as well as conventional lateral cephalometric head films.

Digital radiography offers a number of important advantages over film, including reduced radiation exposure to the patient, instantaneous acquisition of the radiographic image, elimination of the darkroom facilities and development processing time and expense, simplified storage, handling, and sharing of images with appropriate professionals, and the ability to enhance the image to suit the orthodontist's needs. These advantages coupled with proven clinical performance equal to that of film may lead to a shift in what is considered the standard for cephalometric radiography in the future.

References

1. Pacini AJ: Roentgen ray anthropometry of the skull. J Radiol 3:230-231, 1922
2. Hofrath H: Bedeutung der Rontgenfern und Abstands Aufnahme fur die Diagnostik der Kiefereanomalien. Fortschr Orthod 1:231-258, 1931
3. Broadbent BH: A new x ray technique and its application to orthodontics. Angle Orthod1 1:45-66, 1931
4. Quintero JC, Trosien A, Hatcher D, Kapila S: Craniofacial imaging in orthodontics: historical perspective, current status, and future developments. Angle Orthod 69:491-506, 1999
5. Brennan J: An introduction to digital radiography in dentistry. J Orthod 29:66-69, 2002
6. Schulze R, Krummenauer F, Schalldach F, d'Hoedt B: Precision and accuracy of measurements in digital panoramic radiography. Dentomaxillofac Radiol 29:52-56, 2000
7. Held CL, Ferguson DJ, Gallo MW: Cephalometric digitization: a determination of the minimum scanner settings necessary for precise landmark identification. Am J Orthod Dentofacial Orthop 119:472-481, 2001
8. Halazonetis DJ: At what resolution should I scan cephalometric radiographs? Am J Orthod Dentofacial Orthop 125:118-119, 2004
9. Chen YJ, Chen SK, Chang HF, Chen KC: Comparison of landmark identification in traditional versus computer-aided digital cephalometry. Angle Orthod 70:387-392, 2000
10. Hildebolt C, Couture RA, Whiting BR: Dental photostimulable phosphor radiography. Dent Clin North Am 44:273-297, 2000
11. Gregston MD, Kula T, Hardman P, Glaros A, Kula K: A comparison of conventional and digital radiographic methods and cephalometric analysis software: I. hard tissue. Semin Orthod 10:204-211, 2004
12. Geelen W, Wenzel A, Gatfredsen MK, Hansson LG: Reproducibility of cephalometric landmarks on conventional film, hardcopy and monitor displayed images obtained by the storage phosphor technique. Eur J Orthod 20:331-340, 1998
13. Sagner, Storr I, Benz C, Rudzki-Janson I: Diagnostic image quality in comparison of conventional and digital cephalometric radiographs [abstract]. Dentomaxillofac Radiol 27:27, 1998
14. Schulze RKW, Burkhardt G, Doll GM: Landmark identification on direct digital versus film-based cephalometric radiographs: a human skull study. Am J Orthod Dentofacial Orthop 122:635-642, 2002
15. Gravely JF, Benzies PM: The clinical significance of tracing error in cephalometry. Br J Orthod 1:95-101, 1974
16. Richardson A: A comparison of traditional and computerized methods of cephalometric analysis. Eur J Orthod 3:15-20, 1981
17. Liu JK, Chen YT, Cheng KS: Accuracy of computerized automatic identification of cephalometric landmarks. Am J Orthod Dentofacial Orthop 118:535-540, 2000
18. Baumrind S, Frantz R: The reliability of head film measurements: 1. Landmark identification. Am J Orthod 60:111-127, 1971
19. Trpkova B, Major P, Prasad N, Nebbe B: Cephalometric landmarks identification and reproducibility: a meta analysis. Am J Orthod Dentofacial Orthop 112:535-540, 1997
20. Richardson A: An investigation into the reproducibility of some points, planes, and lines used in cephalometric analysis. Am J Orthod 52:637-651, 1966
21. Baumrind S, Frantz RC: The reliability of head film measurements: 2. Conventional angular and linear measures. Am J Orthod 60:505-517, 1971
22. Stabrum AE, Danielsen K: Precision in cephalometric landmark identification. Eur J Orthod 4:185-196, 1982
23. Perillo MA, Beideman RW, Shofer FS, Jacobson-Hunt U, Higgins-Barber K, Laster LL: Effect of landmark identification on cephalometric measurements: guidelines for cephalometric analyses. Clin Orthod Res 3:29-36, 2000
24. Ellis E, McNamara J: Cephalometric reference planes—Sella nasion vs. Frankfort horizontal. J Adult Orthod Orthognath Surg 2:81-87, 1988
25. Bjerin R: A comparison between the Frankfort horizontal and the Sella turcica-nasion as reference planes in cephalometric analysis: Acta Odontol Scand 15:1-13, 1957
26. Lundstrom A, Lundstrum F: The Frankfort horizontal as a basis for cephalometric analysis. Am J Orthod Dentofacial Orthop 5:537-540, 1995