J Forensic Sci, July 2009, Vol. 54, No. 4
doi: 10.1111/j.1556-4029.2009.01075.x
Available online at: www.blackwell-synergy.com

Mary Cox,[1] *B.Sc.(Hons.); Matthew Malcolm,*[1] *B.Sc.(Hons.); and Scott I. Fairgrieve,*[1] *Ph.D.*

A New Digital Method for the Objective Comparison of Frontal Sinuses for Identification*

ABSTRACT: Use of the frontal sinuses for identification requires an objective method of comparison to meet Daubert standards. Christensen's application of Elliptical Fourier Analysis and Likelihood Ratios seems to be a viable solution for this problem. The proposed method draws upon this work and attempts to simplify its application. Variation between pairs of digitized sinus tracings was quantified by summing the difference between corresponding measurements taken from a fixed origin to the outer edge of the sinus outlines using Adobe Photoshop® CS2. Same-skull and different-skull pairs were used to develop reference distributions from which the probability of unknown pairs coming from the same or a different individual was estimated. Error rates of 0% were achieved. Resulting correlation coefficients demonstrated inter-rater and test–retest reliability. Further refinement of the reference distributions and more rigorous testing of error rates should make this technique applicable to casework.

KEYWORDS: forensic science, forensic anthropology, identification, frontal sinuses, Photoshop® CS2

Radiographic comparison of osteological structures is commonly used to confirm identification of human remains that are highly decomposed, cremated, or otherwise disfigured (1–4). Several structures, often used in conjunction, have been utilized for this purpose, including the sella turcica, mastoid air cells, paranasal sinuses, and particularly the frontal sinuses (5). The frontal sinuses (henceforth referred to collectively as the frontal sinus or FS) are contained within the frontal bone, and consist of two air-filled cavities originating at the root of the nose and expanding superiorly into the peri-glabellar region. The FS has long been accepted as an ideal structure for individualization, dating back to the 1920s (6) due to its inherently variable morphology, permanency throughout adulthood, resiliency to damage, and the moderate availability of adequate antemortem (AM) radiographs.

The highly variable nature of FS morphology among individuals is evident even when viewing this three-dimensional (3D) structure projected onto a two-dimensional (2D) plain film radiograph. As a 2D image, the outer borders remain clearly defined and show a highly irregular, scalloped outline. This variability has been attributed to the complexity of the naso-facial area and the impact that environmental and anatomical factors have during development (7–9). The uniqueness of FS morphology is supported by observed variation of the structure in monozygotic twins (7,10). Although this uniqueness is widely accepted in the scientific community, most of these claims have been based on subjective observation with little empirical testing (7,10–14). This drawback was recently addressed by measuring Euclidean distances between 1000 combinations of sinus outlines generated by Elliptical Fourier Analysis (EFA), which empirically demonstrated adequate variability in FS morphology to allow for reliable identification (15).

A second feature making the FS an ideal structure for identification is its permanency throughout life after the age of 20 (9). Few pathologies have the ability to alter the FS shape, and those that can, do so in a manner readily noticeable to an experienced examiner (16). Additionally, in old age, resorption of the bone walls of the FS has been observed to lead to enlargement of these cavities (7,17), potentially affecting its use in positive identification. However, the ability to identify remains using visual inspection of the FS is reported as being unaffected by the time elapsed between AM and postmortem (PM) radiographs (18).

Thirdly, the resiliency of the FS makes it useful in many forensic contexts. It is readily recovered intact as its internal bony structure and arched nature protect it from damage and decomposition. As it is posterior to the thick outer table of the frontal bone in the glabellar region, its stability is further enhanced (19). It has been noted that 800–1600 foot-pounds of force are required to fracture the walls of the FS, as with victims of high impact accidents and gunshot wounds (19,20). Further, as an internal structure, it is often preserved in human cremains (1–3,21).

Lastly, AM paranasal sinus radiographs are often available from a clinical source as they are commonly taken for diagnostic purposes, particularly when the patient complains of headaches, excessive mucus, or suspected facial fractures (16). Sinus radiographs are usually taken in both the occipitomental and occipitofrontal projection, the latter of which provides the clearest view of the FS, though both have been used for identification (3,22,23).

[1]Forensic Osteology Laboratory, Department of Forensic Science, Laurentian University, Ramsey Lake Road, Sudbury, ON, Canada P3E 2C6.

*Preliminary aspects of this work were presented at the 60th Annual Meeting of the American Academy of Forensic Sciences, February 18–23, 2008, in Washington, DC.

Received 6 June 2008; and in revised form 15 Sept. 2008; accepted 1 Oct. 2008.

Review of Frontal Sinus Comparison Methods

Subjective Comparisons

There are several documented cases in which visual comparison of FS morphology has been used to identify remains (1–5,11,24). In these cases, conclusions were based solely on visual comparison of an AM radiograph of the suspected individual to a PM radiograph that is taken to mimic the orientation of the AM as closely as possible. Sinus patterns are then compared for similarity in 2D shape via side-by-side comparison or superimposition. Reports of this technique show high success rates, provided the quality of the AM record is adequate (18,25). It has been reported that superimposition has allowed for correct identification in 100% of cases, regardless of sex, age, cause of death, and the time elapsed between AM and PM radiographs (18). Although these methods show low error rates, they are criticized for being highly subjective and lacking statistical support as to their reliability. These limitations can be seen as shortcomings to admissibility standards of scientific evidence as laid out in *Daubert v. Merell Dow Pharmaceuticals* (26) in the United States and *R v. Mohan* (27) and *R v. J.-L.J.* (28) in Canada. Hence, a more objective method of FS comparison is required.

Objective Methods

Currently, research on FS comparison has been peer reviewed and gained general acceptance in the forensic community, satisfying two of four criteria laid out in Daubert. Although previous claims of uniqueness have been based mainly on subjective observations, recent quantitative analysis has provided support for these claims (15,29). As it stands, there is currently no agreed upon standard for objective comparison with known error rates. There have been several attempts to develop objective methods of comparison designed to provide statistical substantiation to conclusions of identification. These include coding systems based on morphological and metric traits, as well as purely metric systems (14,15,30–36). Techniques using classification systems have shown varying degrees of success and often require visual comparison as a final confirmatory step. This is likely due to the loss of information when features are grouped and assigned class numbers or when simple linear measurements such as height and breadth are used. The result is a representation of the FS that does not capture the degree of variation needed for reliable identification. Codification and measurement systems provide, at best, a way to perform quick searches to narrow down suspect remains and eliminate nonmatches.

The application of EFA to digitized sinus tracings has been successful in providing quantitative support as to the uniqueness of the FS (15). EFA fits a closed curve to an ordered set of data points (representing the FS outline) using an orthogonal decomposition of the curve into a sum of harmonically related ellipses (15). Euclidean distances are calculated between corresponding points every 2° for 360° around the outlines. These were found to be significantly larger between different individuals than between replicates of the same individual with minimal overlap between the distributions. This technique allowed for quantitative assessment of FS morphology, demonstrating that it is adequately variable to allow for reliable identification.

In a later study, this application of EFA was further developed with a proposed likelihood ratio (LR)-based approach that provided statistical support of identifications (29). EFA generates four shape descriptor coefficients that characterize a given sinus shape. These values were used to calculate the LR for a given pair of sinuses (29). The probability of obtaining those values was tested against two competing hypotheses: (1) the outlines originate from the same individual and (2) the outlines originate from different individuals. The LR was calculated from inter-individual and intra-individual variance in EFA coefficients derived from the statistical properties of constructed reference distributions. For a sample 305 same-skull (SS) replicates, at 20 harmonics, the mean LR was $10^{21.22}$, meaning that it was $10^{21.22}$ times more likely to obtain that difference in coefficients if the pair is from the same individual over different individuals. The EFA–LR approach was concluded to be a reliable method to estimate the probability of correct identification and supported uniqueness of FS shape (29).

As successful as this technique was, it required the use of several software programs and complex calculations, limiting its practicality and accessibility. Also, mathematically derived EFA coefficients are fairly abstract and it may be problematic to explain this concept to a jury. Our proposed method is an attempt to simplify the process, hence, making it easier for examiners to apply and for jurors to comprehend.

To do so, a morphometric approach that quantifies the magnitude of difference in 2D morphology of the superior portion of the FS was used. This was achieved by taking a series of measurements from an objectively established point of origin of a traced sinus outline to the outermost border at specific intervals. Pairs of radiographs were compared by summing the difference in corresponding origin–border measurements, yielding a total difference (TD). By finding TD values for SS replicates and random different-skull (DS) combinations from a Contemporary sample, reference distributions were constructed, representing intra- and inter-individual TD variation, respectively. These reference sets were then used to calculate an odds ratio (OR) which provides a continuous measure of the strength of the evidence against two competing hypotheses (i.e., that they are the same individual [H_0] or they are different individuals [H_1]). To develop the technique and then test its reliability and validity, our study is divided into four areas of examination:

1. Intra- and inter-observer reliability.
2. Develop reference distributions.
3. Test the method on a contemporary subset.
4. Empirically test the predictions of the OR.

The central purpose of our study was to develop a quantitatively reproducible method for objective comparison of FS morphology that will be relatively easy to apply, yield consistent results, and capture sufficient variation to allow for reliable identification.

Materials and Methods

Three adult samples were used in this study: Archeological, Clinical, and Contemporary.

Archeological Sample

The Archeological sample consisted of a radiographic collection of 46 adult skulls obtained from the Laurentian University Forensic Osteology Laboratory. The X-rayed skulls were from a 15th century Peruvian population. The radiographs were taken in the standard Caldwell (occipitofrontal) view for a prior study (36). Of the original 59 radiographs, 11 (18.6%) were excluded due to a lack of sinuses, two (3%) were excluded due to poor clarity, and five (8%) were excluded for unilateral absence. This resulted in a usable sample size of 41, consisting of 25 males and 16 females.

Clinical Sample

The Clinical sample was obtained from the Sudbury Regional Hospital and consisted of 63 documented P–A sinus radiographs (Caldwell Projection) taken in 2007. This sample consisted of adults of mixed ancestry, typical of the Greater Sudbury area (predominantly Caucasian). Among the most common complaints of these patients were sinus headache, excessive mucus production, or a suspected infection. Of the 63 radiographs originally obtained, eight (12.7%) were excluded due to a lack of FS above the determined baseline (see Tracing Procedure), six (9.5%) were excluded due to unilateral absence, and one (1.6%) was excluded due to poor quality. The remaining sample size was 48, with 22 males and 26 females. Access to this information and the radiographs was approved by the Research Ethics Committee of the Sudbury Regional Hospital.

Contemporary Sample

The Contemporary sample consisted of 16 adult skulls of a contemporary origin, part of the Laurentian University Department of Forensic Science human osteology teaching collection. Each skull was X-rayed once at the Sudbury Regional Hospital in the standard Caldwell position to simulate an AM radiograph. Of the 16 skulls, four (25%) were excluded due to lack of FS and three (18.8%) were excluded due to a unilateral absence, resulting in nine usable radiographs. These nine skulls were radiographed again at a later date to simulate the PM radiographs. The skulls were positioned by the X-ray technologist to mimic the orientation of the corresponding AM record. Table 1 provides a summary of the three samples used in our study.

FIG. 1—*A radiograph from the Archeological sample with a corresponding traced acetate.*

Tracing Procedure

The following procedure was performed on each radiograph to obtain an outline of the FS for analysis. A sheet of acetate was secured over the radiograph with tape, aligning the left edge of the radiograph with the left edge of the acetate over an illumination box. A clear ruler was used to delineate the baseline of the sinus tangential to the superior margin of the orbits using a fine-tipped indelible marker, as per previous studies (36–38). The midline of the skull was established using a clear ruler to align mid-sagittal bony landmarks such as nasion, nasospinale, the frontal crest, and crista gali. The intersection of the midline and baseline was termed the origin and served as the point from which all measurements were taken. Following Christensen (29), only the outer margin of the FS is traced onto the acetate, excluding bony septations (Fig. 1).

Image Analysis

The completed acetates were scanned with an HP Scanjet® scanner (Hewlett-Packard [Canada] Co., Mississauga, Ontario, Canada) at 300 dpi and saved as a Tagged Image File (.tif). Each file was duplicated, and these duplicates were used as the working files in Adobe Photoshop® CS2 (39) for measurement analysis. The *Magic Wand* tool, located in the tool palette, was used to delineate an unambiguous, dashed outline on the interior border of the sinus outline to which all of the measurements were taken.

The *Ruler* tool, also located in the tool palette, was used to measure the length of the baseline by zooming in and extending the ruler from one end to the other, overlapping the dashed *Magic Wand* line. The length of the baseline (indicated at the top of the screen as D1:), and the angle (displayed as A:), were recorded. The ruler was then shortened by dragging the left terminus to the origin, ensuring that the angle remained the same. The length of this new line (D1:) was also recorded. These data were recorded to allow for precise repositioning if the placement happened to be lost by accidental double clicking elsewhere in the image or by closing the program.

Pressing <alt> and clicking on the terminus of the baseline at the origin established a second line originating from the origin. The length of this line and the angle between it and the baseline were displayed at the top of the screen as A: and D2: (D1: always indicated the length of the baseline after being shortened). Figure 2 is an example of the Adobe Photoshop® CS2 user window with a scanned acetate image opened and the baseline and measuring line established.

The terminus of the measuring line was then dragged around the dashed outline, established by the *Magic Wand* tool. This origin-to-border distance (D2:) was recorded every 3° from 3° to 177°. A 3° measurement interval was considered sufficient to capture enough variability for the purposes of this study. This produced 59 measurements per radiograph which were recorded. The measurements at 0° and 180° were omitted due to ambiguities occurring at the intersection of the sinus outline and the baseline, caused by the tracing process. The measurement was always recorded at the outermost intersection with the outline (Fig. 3).

TABLE 1—*A summary of the radiographic sources used in this study.*

Sample	Original Sample Size	Excluded	Usable Sample (*n*)	X-Ray Parameters
Archeological	59	18 (31%)	41 (M = 25, F = 16)	90 kV, 20 mA, 8 sec (on photographic paper)
Clinical	63	15 (23.8%)	48 (M = 22, F = 26)	Automatic (on film)
Contemporary	16	7 (43.8%)	9	70 kV, 15 mA, 5.1 msec (on film)

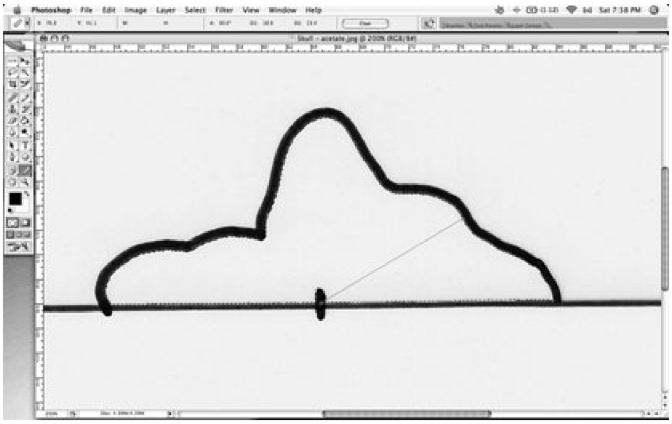

FIG. 2—*A scanned traced acetate opened in Adobe Photoshop CS2®. The baseline and point of origin for measuring have been established. The measurement line extends from the origin laterally to the outline.*

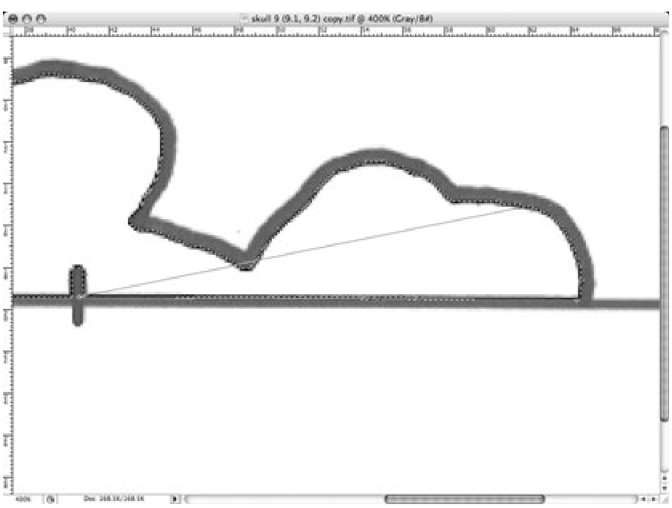

FIG. 3—*A detailed view of a measuring line crossing the sinus outline more than once. The recorded measurement is taken from the outermost intersection.*

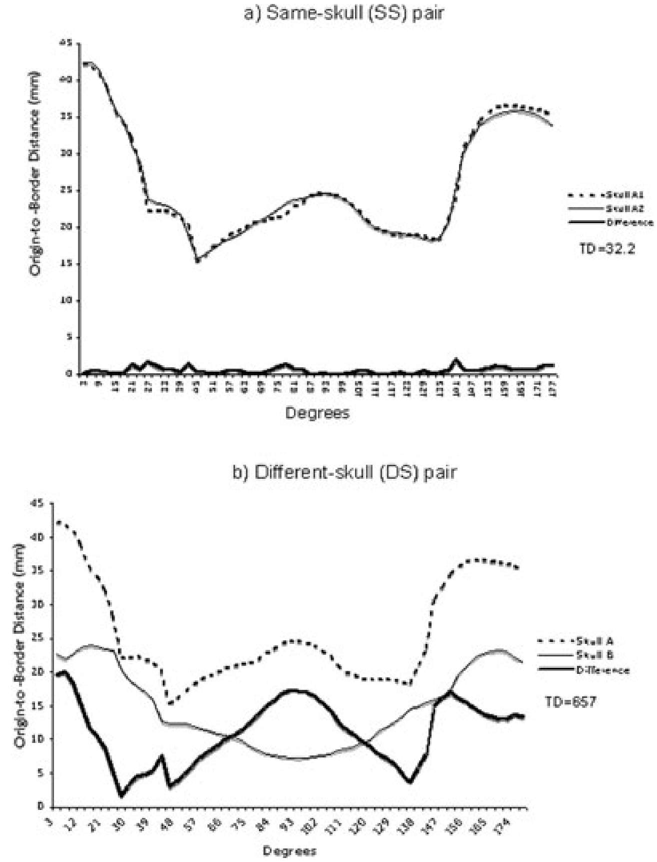

FIG. 4—*A graphical representation of the total difference (TD) for (a) two radiographs of the same-skull and (b) different-skull pairs. The solid bold line represents the magnitude of difference between the two skulls at each corresponding measurement. It is clearly demonstrated that the differences are approaching zero in (a) and highly variable in (b). The sum of the absolute value of these differences is the TD for the pair of skulls.*

Calculating Total Difference

To determine the magnitude of difference in 2D morphology between two radiographed FSs, the difference of the origin-to-border distance (D2, in mm) was found for each corresponding measurement from 3° to 177°. The absolute value of each difference was summed to find the TD, expressed in millimeters, between two sinus outlines. Figure 4 provides a graphical representation of the TD for (1) SS pair and (2) DS pair.

For the Archeological and the Clinical samples, the measurement technique was applied to each radiograph by two experimenters. For the Archeological sample, inter-individual TD scores were found for 110 random DS combinations of radiographs, traced, and measured by the same observer. Additionally, intra-individual TD scores were found between SS replicates of the same radiograph traced and measured by different observers for each of the 41 usable radiographs. This was then repeated for the Clinical sample, where TD values were found for 110 DS combinations and 48 SS replicates.

Intra-Observer and Inter-Observer Error

Intra-observer error was determined by comparing measurement sets of the same radiograph traced and measured twice by the same observer. Ten randomly selected radiographs from the Archeological sample were traced and measured a second time and were compared with their corresponding replicate to which the method had already been applied. Pearson's r correlation coefficients were found between each replicate pair by correlating corresponding measurements at each degree. The correlation coefficients were then used to assess the congruency in patterns achieved by the same observer applying the technique twice.

Inter-observer error was also determined using correlation coefficients. All 41 radiographs of the Archeological sample were traced and measured by two observers. Pearson's r values were then found for each replicate pair to assess consistency between observers.

Frequency Distributions

Both the Archeological and the Clinical sample were assessed for inter-individual and intra-individual variation in TD scores by constructing frequency distributions from samples of SS pairs to represent intra-individual variation, and DS pairs to represent inter-individual variation.

For the Clinical sample, the intra-individual reference distribution was based on the TD found for 48 SS pairs, the number of usable radiographs from the collection traced and measured by different observers. The inter-individual reference distribution was constructed based on the TD values for a sample of 110 DS combinations from the 41 radiographs, traced, and measured by the same experimenter. A sample size of 110 was chosen based on time limitations in applying the technique. Normality was assessed by applying the Shapiro–Wilk test, and log transformations were applied to both of the distributions to smooth these data. An independent *t*-test was used to determine equality of mean logTD between the two distributions. Finally, a one-way ANOVA was used to determine if logTD values exhibited sex dependence by grouping the male–male, female–female, and male–female logTD scores. Because this sample consists of radiographs of modern individuals, taken by standard X-ray procedures, the inter- and intra-individual frequency distributions were chosen as reference sets from which to calculate OR values for unknown pairs when testing the method (see Blind Analysis).

The above procedure and log transformation were repeated for the Archeological sample, where the intra-individual distribution was based on 41 SS pairs, and the inter-individual distribution was based on 130 DS combinations. Sex dependence was also investigated in this sample. This was carried out to evaluate population differences in logTD values when compared with the Clinical sample. Population differences in intra-individual TD scores were assessed by applying an independent *t*-test for the equality of mean intra-individual logTD for the Clinical and Archeological samples. This was repeated for the inter-individual logTDs between these samples.

Blind Analysis

A subset of nine skulls from the Contemporary sample was used to test the method, using the estimated normal curves of the Clinical reference distributions to calculate an OR between simulated AM and PM pairs. The AM radiographs were labeled "A" through "I" and provided to the observers (MC and MM) by the third author (SIF). Each observer independently traced and measured the radiographs using the prescribed method. Approximately 1 week later, the PM radiographs, labeled one through nine, were provided to the observers to be traced and measured. The observers were then given a sheet of 30 combinations of radiographs (one AM and one PM) for comparison. The TD was calculated for each pair and used to calculate the OR, which discriminated the unknown pair as exhibiting SS or DS variation and provided a measure of the strength of the conclusion.

The OR was calculated by dividing the probability of obtaining the given logTD or greater, if the pair is from the same individual (SS), by the probability of obtaining the given logTD or less, if the pair is from different individuals (DS).

$$\mathrm{OR} = \frac{p(\geq \mathrm{logTD}|\mathrm{SS})}{p(\leq \mathrm{logTD}|\mathrm{DS})}$$

These conditional probabilities were derived from the statistical properties of the intra- and inter-individual reference distributions to calculate the numerator and denominator, respectively. This was accomplished by converting the logTD into two *z*-scores according to the formula (score − mean)/SD.

The first *z*-score was calculated using the mean and the standard deviation (SD) of the intra-individual reference distribution (mean = 1.82, SD = 0.142). This standardized score was then used to find the area under the normal curve associated with the given score in the positive tail. This yielded the probability of obtaining that logTD, or higher, given that the pair is from the same individual (the numerator). The second *z*-score was calculated using the mean and the SD of the inter-individual reference distribution (mean = 2.67, SD = 0.236). Again, the resulting score was used to find the area under the normal curve of the inter-individual distribution in the negative tail. This gave the probability of obtaining the given TD, or less, given that the pair is from different individuals (the denominator).

TABLE 2—*Verbal equivalents for LR values used to state confidence ratings for OR values adapted from Rose (40).*

LR/OR	Possible Verbal Equivalent	Exclusion Status
>10,000	Very strong	Failure to exclude
1000–10,000	Strong	
100–1000	Moderately strong	
10–100	Moderate	
1–10	Limited	
1–0.1	Limited	Exclude
0.1–0.01	Moderate	
0.01–0.001	Moderately strong	
0.001–0.0001	Strong	
<0.0001	Very strong	

OR, odds ratio; LR, likelihood ratio.

The resulting OR was calculated by dividing the SS probability by the DS probability. This was used to state conclusions about the unknown pair as more likely demonstrating SS or DS variability. If the OR was greater than one, the radiographs could not be excluded as originating from the same individual, i.e., "Fail to Exclude" (FTE). If the OR was less than one, the opinion was given as an "Exclusion" of the radiographs as having the same source. The further from the value of one, the greater the strength of the evidence. To state a consistent confidence rating, a verbal equivalency chart for LRs was utilized (Table 2).

This process was repeated for all 30 unknown pairs derived from the Contemporary sample, by two different observers. Conclusions and confidence ratings were recorded and the third author (SIF) then used a key to reveal correct matches to establish error rates.

Empirical Testing

As a final stage in this study, we conducted an empirical test of the correctness of the theoretical prediction given by the OR. This was accomplished by finding the OR for each of the 36 combinations and nine replicate pairs from the Contemporary sample. Correctness of predictions was determined by finding the percentage of SS pairs for which the OR was above one, and the percentage of DS pairs for which the OR was below one.

Finally, the TD value which yields an OR = 1 was determined using trial and error. This value was termed the "threshold value." By converting it into a standardized score for each distribution (described above), the probability of a SS pair showing that threshold or less was found, as well as the probability of a DS pair showing that threshold or greater. This was used to assess the degree of overlap of the reference distributions, and the discriminatory power of the OR.

Results

Intra-Observer and Inter-Observer Error

Intra-observer error was assessed by finding Pearson's *r* correlation coefficients between measurement sets of 10 SS radiographs from the Archeological sample, traced, and measured by the same observer (MM). The *r* values were found to be high, ranging from 0.921 to 0.998, with a mean *r* of 0.976 (Table 3).

Inter-observer error was also assessed using the Archeological sample by finding Pearson's *r* correlation coefficients between

TABLE 3—*Archeological sample; descriptive statistics for Pearson's r correlation coefficients used to assess intra-observer and inter-observer reliability.*

Statistic	Intra-Observer	Inter-Observer
Mean	0.976	0.978
Standard deviation	0.028	0.023
Minimum	0.921	0.925
Maximum	0.998	0.999
Sample size	10	41

measurement sets of the same radiographs traced and measured by two different observers (MM and MC). Based on 41 SS replicates, *r* values were also found to be high, ranging from 0.925 to 0.999, with a mean *r* of 0.978 (Table 3).

Archeological Sample

Total difference values were calculated between SS replicates, traced, and measured by two observers and were plotted in a frequency distribution to demonstrate intra-individual variation. This distribution was found to be nonparametric using a Shapiro–Wilk test for normality ($p < 0.0001$). A log transformation was applied and normality was achieved with $p = 0.308$ (Fig. 5).

Total difference values were also calculated for 130 random DS combinations from the Archeological sample, traced, and measured by one observer (MM). These values were plotted in a frequency distribution, which was found to be nonparametric ($p < 0.0001$). Again, a log transformation was applied, and normality was achieved with $p = 0.107$ (Fig. 6). An independent *t*-test for equality of means was applied to the intra- and inter-individual logTD scores, after confirming the assumption of equal variances with the *F*-statistic ($F = 1.457$, $p = 0.229$). This test indicated that the two distributions were significantly different ($p < 0.0001$).

Percent cumulative frequency plots were used to provide a graphical representation with both intra- and inter-individual logTD values plotted on a standardized scale (Fig. 7). Upon visual inspection, it was observed that the two distributions showed a high degree of separation with minimal overlap.

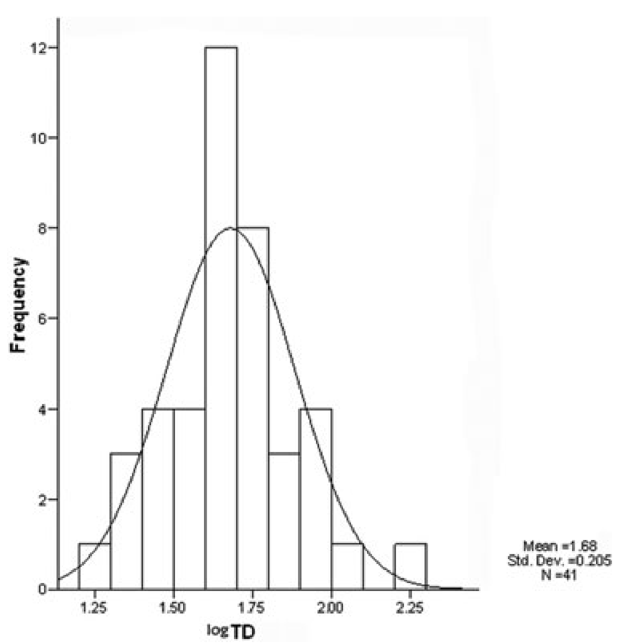

FIG. 5—*Archeological sample—frequency distribution with estimated normal curve of the logTD values of 41 replicate SS pairs, used to demonstrate intra-individual variation.*

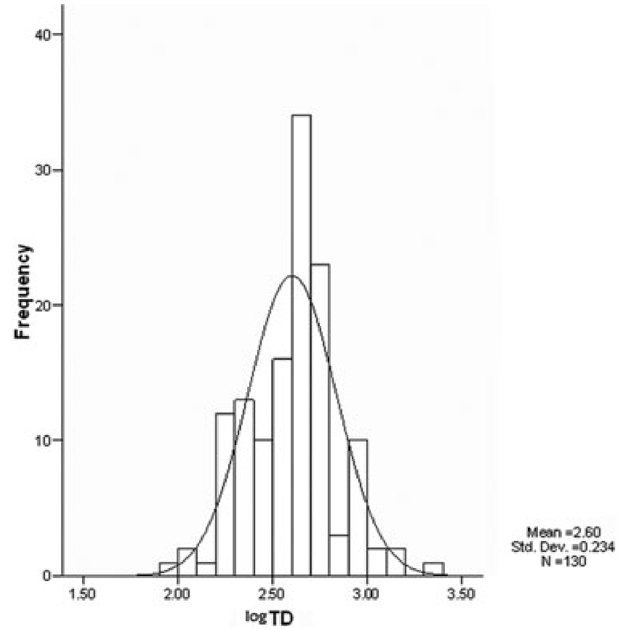

FIG. 6—*Archeological sample—frequency distribution with estimated normal curve of the logTD values of 130 random combinations, used to demonstrate inter-individual variation.*

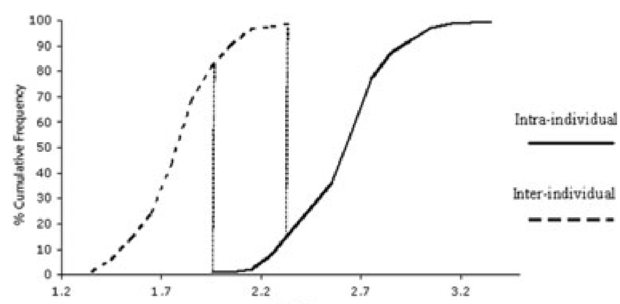

FIG. 7—*Archeological sample—a % cumulative frequency plot of the logTD values for SS (intra-individual) and DS (inter-individual) pairs demonstrating the degree of overlap.*

A one-way ANOVA was used to determine if significant differences in logTD values occurred between same-sex combinations (male–male and female–female) and different-sex combinations (male–female). No significant differences were found between the three groups with $p = 0.531$.

Clinical Sample

Total difference values were found between SS replicates traced and measured by two different observers (MC and MM). These data, representing intra-individual variation, were found to be parametric using a Shapiro–Wilk test for normality ($p = 0.360$). A log transformation was applied to maintain consistency and to improve normality. These transformed data were found to be normal with $p = 0.850$ (Fig. 8). The logTD mean was 1.82 with a SD of 0.142 and $n = 48$. These are the values used to calculate the numerator of the OR, or the probability of obtaining that logTD or greater given that the radiographs are a SS pair.

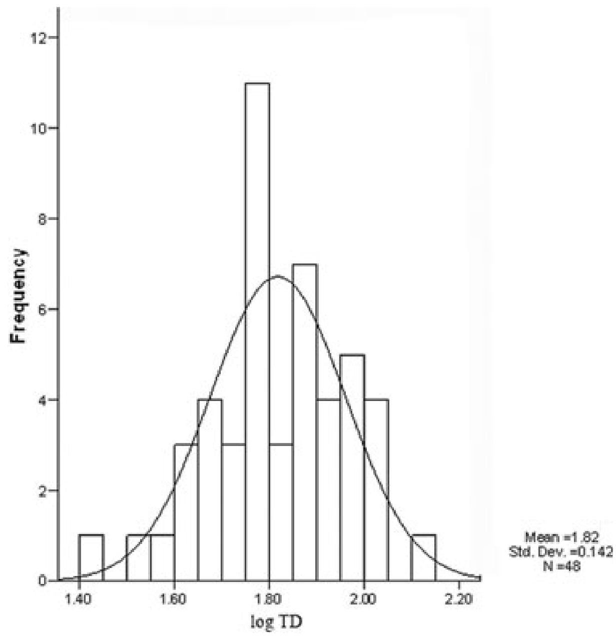

FIG. 8—*Clinical sample—intra-individual reference frequency distribution with estimated normal curve based on logTD values of 48 same-skull replicates.*

Total difference values were also found between 110 random DS combinations, traced and measured by the same observer (MM) to function as the inter-individual reference set. These data were found to be nonparametric ($p < 0.0001$). A log transformation was applied and normality was achieved ($p = 0.086$) (Fig. 9). The mean was 2.67 with a SD of 0.236, which were used to calculate the denominator of the OR, or the probability of obtaining that log-TD value or less given that the radiographs are a DS pair.

The *F*-statistic was calculated to assess variance homogeneity between the intra- and inter-individual distributions. These data sets were found to have significantly different variances ($F = 45.30$, $p < 0.0001$). An independent *t*-test for equality of means, assuming unequal variances, was then applied and the intra- and inter-individual distributions were found to be significantly different ($p < 0.0001$).

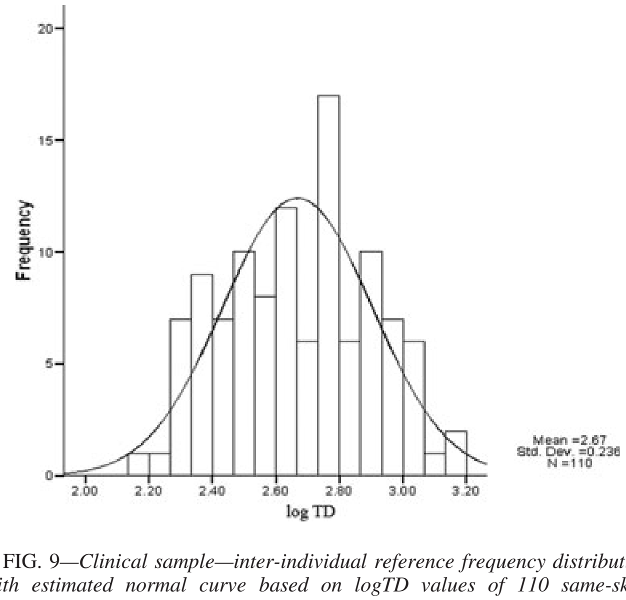

FIG. 9—*Clinical sample—inter-individual reference frequency distribution with estimated normal curve based on logTD values of 110 same-skull replicates.*

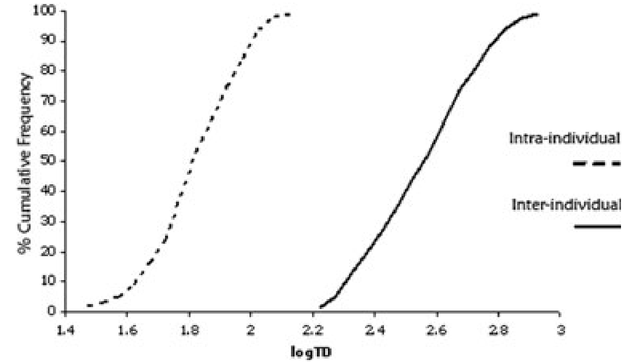

FIG. 10—*Clinical sample—a % cumulative frequency plot of the logTD values for SS (intra-individual) and DS (inter-individual) pairs demonstrating a lack of overlap. The distributions are significantly separated* ($p < 0.0001$).

To visualize the SS and DS distributions on the same scale, percent cumulative frequency plots of the logTD values were constructed (Fig. 10). No overlap occurred between the logTD values of the two distributions.

To determine if sex differences exist in TD values, a one-way ANOVA was applied to the logTD values of three groups: male–male combinations, female–female combinations, and male–female combinations. This test indicated that no significant differences existed between the same-sex and the different-sex combinations; therefore, sex-specific distributions were not deemed necessary.

Population differences between Archeological and Clinical samples were assessed for both the intra- and inter-individual mean logTD using independent *t*-test under the confirmed assumption of equal variance ($F = 2.34$, $p = 0.130$; $F = 0.946$, $p = 0.332$). For intra-individual variation, the *t*-test indicated a significant difference in mean logTD between the two samples ($p < 0.0001$). This was also found for the inter-individual mean logTD ($p = 0.033$). These results indicated that the samples could not be combined.

Blind Analysis

Two observers each performed a blind analysis of the method by finding OR values for 30 unknowns consisting of nine SS pairs and 21 DS pairs. These were compiled from simulated AM and PM radiographs taken of the Contemporary sample. For the nine SS pairs, OR values were found to range from 1.58 to >99,900, classifying all of them in the "FTE" category, with confidence ratings ranging from "limited" to "very strong." For the two observers combined, 12 of 18 (67%) of the "FTE" opinions fell within the "very strong" confidence rating, where the OR exceeded 10,000 (Fig. 11).

For the 21 DS pairs, OR values were found to range from <0.0001 to 0.523 meaning that confidence ratings to "Exclude" the radiographs as coming from the SS ranged from "limited" to "very strong." When combining the results of both observers, 38 of 42 (90%) of "Exclude" opinions fell in the "very strong" confidence rating, where OR < 0.0001 (Fig. 11).

Overall error rates were assessed by determining the percentage of DS pairs where an OR > 1 was obtained (i.e., false positive), and the number of SS pairs where an OR < 1 was obtained (i.e., false negative). For both observers, the error rates for both false

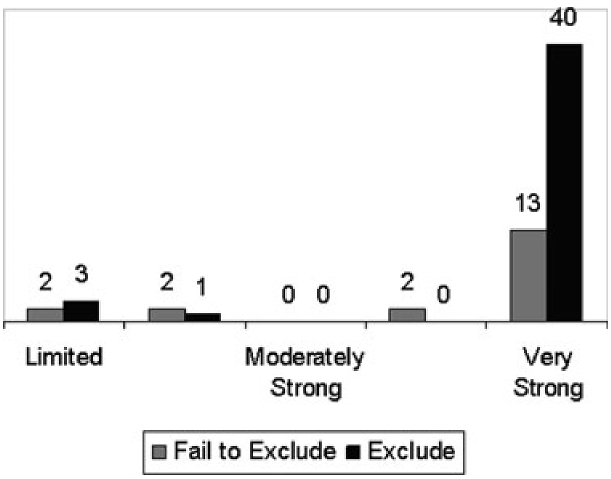

FIG. 11—*A summary of the confidence ratings from our blind analysis for "Exclude" and "Fail to Exclude" opinions, expressed as a percentage.*

positive and false negative were found to be 0%, giving an overall error rate of 0%.

Threshold Value

A threshold TD, yielding an OR = 1, was found to be 137.5 mm. The corresponding *z*-score under the Clinical log intra-individual distribution estimated that 99% of SS pairs with a TD below 137.5 mm will be correctly discriminated. Similarly, based on the Clinical log inter-individual distribution, a TD value above 137.5 mm should theoretically discriminate 99% of DS pairs correctly.

Empirical Testing

For the Contemporary sample, OR values were found between each SS pair and each DS combination by both observers (Table 4). This was performed to empirically test the theoretical predictions of the OR values, where SS pairs with OR > 1 were correctly discriminated, and DS pairs with an OR < 1 were correctly

TABLE 4—*Cross reference charts showing the odds ratio (OR) values found for 36 different-skull (DS) pairs and 9 same-skull (SS) pairs from the Contemporary sample.*

a) Observer 1 (MM)

	A	B	C	D	E	F	G	H	I
A	>70 540								
B	<0.00001	1.557							
C	<0.00001	0.009	38.125						
D	<0.00001	3.62	<0.00001	>99 220					
E	<0.00001	<0.00001	<0.00001	<0.00001	>73 570				
F	<0.00001	<0.00001	<0.00001	0.015	<0.00001	6 517			
G	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	>99 900		
H	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	>99 040	
I	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	>98 540

b) Observer 2 (MC)

	A	B	C	D	E	F	G	H	I
A	63.684								
B	<0.00001	4.758							
C	<0.00001	<0.00001	2 182						
D	<0.00001	3.5	<0.00001	>94 180					
E	<0.00001	<0.00001	<0.00001	<0.00001	>96 780				
F	<0.00001	<0.00001	<0.00001	0.123	<0.00001	>98 540			
G	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	>98 540		
H	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	>96 160	
I	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	>99 990

Strong to very strong Exclude
Limited to moderately strong Exclude
Limited to moderately strong FTE
Strong to very strong FTE

discriminated. For SS pairs, 100% were correctly discriminated (OR > 1; TD < 137.5 mm), and 97% of DS pairs were correctly discriminated (OR < 1; TD > 137.5 mm).

Summary

Pearson's *r* correlation coefficients found between sinuses of the Archeological sample, traced and measured by two observers, were found to be high (mean = 0.978) indicating a high level of congruency in patterns achieved between observers. For both the Clinical and Archeological sample, log transformed distributions were significantly parametric. The Clinical reference distributions were used to test the method, using nine skulls from a Contemporary sample. Combined error rates for both observers showed 0% false positive, and 0% false negatives for 18 SS pairs and 42 DS pairs. The associated confidence ratings were predominantly "very strong." Empirical testing of the ability of OR to discriminate SS from DS logTD values yielded 100% correct discrimination of SS pairs ($n = 9$) and 97% correct discrimination of DS pairs ($n = 36$). A threshold TD of 137.5 mm was found where the OR = 1. Theoretically, 99% of SS pairs with TD less than the threshold should be correctly discriminated, and 99% of DS pairs with a TD greater than the threshold should be correctly discriminated. This closely reflects the results of the empirical testing.

Discussion

The first goal of the study was to determine if the proposed method demonstrated reliability within and between observers in the manual tracing and measurement of FS radiographs. Our results indicated a consistency in the patterns achieved between and within observers, yet seemed to fall short of providing a complete picture of reliability in the Archeological sample. While the correlation coefficients assessed a similarity in pattern, they placed the measurement sets on their own relative scale and, therefore, did not consider consistency in absolute scores. For example, if two observers accurately traced the outline patterns, but one observer consistently drew the baseline lower than the other, these measurements would be consistently larger than those obtained from the outline with the higher baseline. Therefore, the *r*-value would be high regardless of the differences in absolute scores due to the similarity in pattern among the outlines. Further statistical testing could be applied to assess the congruency in absolute scores to yield a more accurate estimate of the reliability of the technique. What may be more relevant is the consistency in discrimination of unknown pairs of radiographs to assess inter-observer reliability. Empirical testing of the accuracy of the OR demonstrated consistent results between observers who found 100% of SS pairs to have OR > 1, 97% of DS pairs to have an OR < 1, and one DS pair to be a false positive (OR > 1). In other words, there was 100% agreement between observers as to which pairs exhibited SS variation, and which skulls exhibited DS variation. Furthermore, confidence ratings were found to be consistent for all but three SS and two DS pairs. Based on these results, it was assumed that inter-observer reliability was achieved as indicated by high Pearson's *r* values, and consistent conclusions when applying the technique. The inconsistencies in confidence ratings were likely related to differences in perception when the outlines were traced.

It must be mentioned that the observed consistencies between observers may be related to the fact that the observers invented the technique in concert. Further investigation into consistency achieved between less experienced examiners would provide a better idea of the overall reliability of our method.

The second goal of our study was to construct appropriate intra- and inter-individual reference distributions. The statistical properties of these reference distributions were used to calculate the probabilities of obtaining the evidence regarding the two competing hypotheses; the skulls are from the same individual (H_0) or the skulls are from different individuals (H_1). The Clinical sample was chosen to develop the reference distributions as it was more representative of a modern population compared with the Archeological sample.

The Clinical intra-individual reference distribution represents within-person variation in TD values for a modern, clinical population. A major limitation of this study is that the TD values used to estimate the Clinical intra-individual normal curve were derived from replicates of the same radiograph, traced and measured by two different observers (MC and MM). This does not account for the inevitable variation introduced by inexact alignment of PM to AM radiographs. One study found that relative to other cranial dimensions, linear measurements of the FS are the most susceptible to within-person variation introduced by differing orientations of the skull (30). For the current technique, it seems as though X-ray misalignment would lead to higher intra-skull TD values because this error would accumulate with each of the 59 recorded measurements. Therefore, an ideal reference distribution should be derived from replicates of the SS radiographed twice, traced, and measured by the same observer. Due to resource limitations, this was not possible in our study, warranting a discussion of the implications of the results.

The misalignment of the specimen during radiography has the potential to introduce more variation in SS pairs. Hence, it was assumed that the intra-individual distribution used in this study demonstrated less overlap with the inter-individual distribution than would occur with the ideal reference set. This meant that with the ideal reference set, the discriminatory power of the TD would be less, as there would be a greater area of overlap and a larger percentage of cases would be assigned an equivocal conclusion. Theoretically, by not accounting for variation introduced by inexact orientation of the radiographs, the intra-individual reference distribution underestimated the probability of obtaining a given TD, describing it as falling closer to the positive tail of the distribution than in actuality. For DS pairs, this would cause OR values to overestimate the strength of the evidence as a "FTE" because the numerator would be underestimated. For SS pairs, OR values would underestimate the strength of the evidence, again by underestimating the numerator. Underestimating the probability of an SS pair as coming from the intra-individual distribution presents a risk of obtaining false negatives. This being said, results of empirical testing revealed that this limitation did not seem to hinder the ability of the OR to discriminate SS and DS pairs from simulated AM and PM radiographs, as 100% of SS and 97% of DS pairs were correctly discriminated. Furthermore, the mean logTD for the SS pairs of the Contemporary sample was 1.67, which is less than that of the intra-individual reference distribution (mean logTD of 1.82). This may have been due to an inadequate sample size ($n = 9$), the skill level of the X-ray technologist, or the experience of the observers at the time the Contemporary sample was analyzed. Nonetheless, the OR values in our study may not provide an accurate measure of the strength of the evidence due to misrepresentation of the intra-individual reference distribution.

The inter-individual reference distribution, derived from the Clinical sample, was constructed to represent variation in logTD values occurring between random combinations of DS pairs from the relevant population. Essentially, this distribution represents the degree of variation between individuals. For example, some random pairs may have similar sinus morphology and yield a lower TD, while others may have quite different morphology, yielding a larger TD. This is likely attributed to overall size rather than a specific scalloping pattern. The normal curve estimated for this reference distribution was used to find the probability of an unknown pair as showing DS variation so OR values could be calculated for discrimination. For this reason, it is vital that this distribution properly represents the relevant population so that the OR provides an accurate measure of the strength of the evidence.

As the name implies, the Clinical sample was derived from a clinical source in which all the individuals were radiographed for diagnostic purposes. In actual casework, AM records would normally be obtained from a clinical source. As our method was tested on a subset of contemporary skulls, not from a clinical source, the reference sets may not have been representative of this particular sample in this sense. Nevertheless, error rates of 0% were found when the method was tested.

The discriminatory power of the TD relies on a high degree of separation between inter- and intra-individual reference distributions. A *t*-test demonstrated that intra- and inter-individual Clinical distributions were significantly different ($p < 0.0001$) with no overlap in logTD values. Minimal overlap did occur between the normal curves estimated from these distributions, however, only in the extreme tails. Because the FS is a fixed structure, variation occurring in the intra-individual distribution could be due to error introduced by the technique, likely from manual tracing (and in the ideal reference distribution from misalignment of the radiographs). Yet, empirical assessment of OR values revealed that 100% of SS pairs and 97% of DS pairs were correctly discriminated for the Contemporary sample. However, a small proportion of conclusions stated were given a "limited" confidence rating, as 11.1% of cases had OR values ranging from 1 to 10 (FTE) and 2.78% of cases had OR values from 0.1 to 1 (Exclude).

A threshold TD value was found to be 137.5 mm when OR = 1 (unity), in which case there was an equal probability of obtaining the TD if the radiographs were of SS or DS origin. Based on associated areas under the normal curves, it was found that 99% of SS pairs fell below this threshold and 99% of DS pairs fell above it. TD values falling at or near the threshold value of 137.5 mm would yield an "Equivocal" conclusion. We suggest that the limited confidence rating (SS OR = 1–10; DS OR = 0.10–1) should be stated as "Equivocal."

No significant differences were found in logTD values between same-sex and different-sex combinations for the Archeological and Clinical samples, indicating that sex-specific reference distributions were not necessary for these samples. Previous studies have demonstrated contrasting results with regards to sex-specific variation in FS size. For example, Yoshino et al. (14) found no significant sex differences in FS area for a Japanese sample, whereas Buckland-Wright (37) found that males have significantly larger sinuses than females with regards to linear measurement of early British crania. However, logTD values are not a direct measure of the size or shape of a sinus; rather, they represent the magnitude of difference that occurs between a given pair of skulls, making meaningful comparison to the results of these past studies difficult. This does not, however, rule out the possibility of other ancestral groups showing significant sex differences in TD scores.

Population differences in logTD values were assessed by comparing the distributions created for the Archeological and Clinical samples. A *t*-test indicated significant differences between the samples for both inter- and intra-individual logTD distributions. While this result indicates a need for population specific reference sets, in a forensic context there is greater interest in whether multiple modern populations are different and thus require separate distributions. Additionally, the Clinical sample was of a mixed ancestry and presumed to be predominantly Caucasian due to the hospital's location. Other studies have noted population differences in linear measurements of the FS between contemporaneous Inuit populations, possibly related to cold adaptations (38). Clearly, further investigation in this area is required.

The third goal was to test our method using a subset of nine contemporary skulls for which AM and PM radiographs were simulated to determine error rates of the technique. Results of the blind analysis demonstrated no false positive or false negative results. This supports the validity of the conclusions generated using only OR values. Such error rates are comparable to those of visual inspection (18,25); however, they were based on our objective methodology and provided statistical substantiation of our conclusions. The error rates further indicate that TD values capture sufficient variation between individuals while minimizing variation within individuals to allow for powerful discrimination.

Not only were the error rates 0% but the confidence ratings derived from the LR verbal equivalency chart (40) were usually in the highest rating (i.e., "very strong"). Over 65% of pairs classified as an FTE were given a "very strong" rating and an additional 11% received a "strong" rating. Further, 91% of pairs excluded as originating from the same source received a "very strong" (<0.0001) rating. There were, however, some weak conclusions for both FTE and "Exclude" where 11.1% of FTE pairs and 7.1% of "Exclude" pairs received a "limited" rating. In our study, 11.1% of SS pairs and 7.1% of DS pairs would be classified as "Equivocal" conclusions and would require other means of corroborating the identification. The blind test demonstrated the usefulness of the technique in simulated case scenarios that were reflective of real casework.

The final goal of this study was to empirically test the theoretical predictions given by the OR for 45 pairs of radiographs from the Contemporary sample. This was achieved by finding the percentage of SS pairs correctly discriminated (OR > 1) and the percentage of DS pairs correctly discriminated (OR < 1) by two observers (MC and MM). Our results indicate that the OR is highly accurate, with 100% of SS pairs and 97% of DS pairs correctly discriminated, with only one false positive noted. This false positive did, however, only receive a "limited" confidence rating and would be stated as "Equivocal." These results support the validity of the technique and its potential as a powerful tool for objective comparison of FS morphology.

Our technique was designed to achieve the discriminatory power of EFA, while simplifying the procedure to allow wider application. The frequency distributions based on total Euclidean distance between EFA generated outlines were quite similar to those obtained in our study as both demonstrated significant differences ($p < 0.0001$). However, EFA did show an overlap in the actual scores, whereas, in our study the only overlap that occurred was in the extreme tail ends of the estimated normal curves. This may reflect the larger sample used in the EFA study (15).

In a subsequent study, LRs were calculated for SS replicates using EFA coefficients (29). By entering these coefficients into a LR formula, at 20 harmonics, the mean LR was $10^{21.22}$. These values are extremely high in comparison to those achieved in our study, however, the use of *z*-scores limited the calculation of the OR as the tables do not exceed ±4.0 SD. Therefore, OR values were stated as < 0.0001 to > 99,990 and would be more powerful with an extended z-table. Nonetheless, these values were considered strong enough to yield positive identifications.

Finally, EFA did not test the predictions of the LR and no error rates were stated. Therefore, the validity of these conclusions cannot be compared (15,29). In summary, the results obtained in our study are comparable to those using EFA, and the procedure is easier to apply.

Limitations

Our technique relies on adequate AM records being available for comparison. Clearly, those who have not had facial radiographs during their lifetime are not candidates for the application of the method. The individual must have been an adult (≥ 20 years) when the AM radiograph was taken in the Caldwell Projection and radiographs must be of sufficient quality to allow for a proper assessment. The Caldwell Projection, although standard at most hospitals for sinus radiographs, may not be available. Other projections, such as the Waters (occipito-mental) or even a lateral view, may be of use, but would require the construction of relevant reference distributions as well as testing of reliability and validity of discrimination.

Our technique also requires a complete and bilateral FS to be present above the defined baseline, limiting the extent to which this method can be applied. Without a complete sinus, bias occurs when comparing the TD to the reference distribution. Without all 59 measurements, the TD would be lower than if all measurements were present and would, therefore, be biased toward SS variation.

Radiographic identification has a general limitation in that it is difficult to obtain a conventional PM X-ray image with the exact alignment of the available AM radiograph. Our technique relies on maximal separation between the distributions; therefore, an important aspect is minimizing intra-individual variation by obtaining PM images as close to the AM record as possible. One method suggests that a more exact alignment may be achieved by taking consecutive computerized tomography (CT) slices of the PM skull and reconstructing them to obtain a 3D representation (41). Using specialized software, 2D conventional radiographs could be simulated from these CT data at any desired angle. The viewing angle of the projected image can be easily rotated until the desired orientation is achieved as displayed on a monitor.

One may question the redundancy of a technique such as this, particularly due to the fact that identity is rarely an issue challenged in court. Visual comparison methods have demonstrated comparably low error rates, are much quicker to apply, and require fewer resources. Perhaps, our technique need not be applied in all cases where the FS is used for identification, but should be reserved for cases in which the identity is contested, to demonstrate the validity of the identification process.

Summary and Conclusions

This technique has demonstrated reliability and validity when subjected to various tests. The TD measure was successful in capturing sufficient inter-individual variation while minimizing intra-individual variation to allow for reliable identification. The OR provided a continuous measure of the strength of the evidence (TD) against two hypotheses, allowing for an objective statement of confidence ratings. To standardize this methodology, making it useful in forensic cases, a few key areas require further research. First, refinement of the reference distributions is needed to improve representation of the relevant population. Second, a larger sample to test back the technique will yield error rates that can be extended to actual cases. Additionally, a more robust statistical analysis of intra- and inter-observer reliability would be beneficial for assessing not only congruency in patterns, but also consistency in absolute scores. Finally, computer automation of the method could reduce the time needed to apply the technique and eliminate the subjective element introduced by manual tracing.

Acknowledgments

We would like to thank the Manager and staff of Diagnostic Imaging at the Sudbury Regional Hospital, for arranging access to the Clinical sample of radiographs. Thank you to the employees who searched for and manually retrieved the radiographs necessary for our analysis. Our thanks also to the X-ray Department at the Sudbury Regional Hospital for taking the time to X-ray the skulls of the Contemporary sample and taking great care to achieve consistent alignment between simulated AM and PM radiographs. We would like to express our gratitude to Dr. M. Persinger and Prof. Linda St.-Pierre for reviewing our research and for offering advice and insight on various statistical approaches. Also, thanks to Chantal Hermann for her help with SPSS and general statistical guidance. We also thank Jane Cox who provided much needed encouragement and editorial assistance. Finally, we are indebted to the two anonymous reviewers of this manuscript for their helpful suggestions.

References

1. Camppbasso CP, Dell'Erba AS, Belviso M, Di Vella G. Craniofacial identification by comparison of antemortem and postmortem radiographs. Am J Forensic Med Pathol 2007;28(2):182–6.
2. Marlin DC, Clark MA, Standish SM. Identification of human remains by comparison of frontal sinus radiographs: a series of four cases. J Forensic Sci 1991;36(6):1765–72.
3. Nambiar P, Naidu MDK, Subramaniam K. Anatomical variability of the frontal sinuses and their application in forensic identification. Clin Anat 1999;12:16–9.
4. Quatrehomme G, Fronty P, Sapanet M, Grevin G, Bailet P, Ollier A. Identification by frontal sinus pattern in forensic anthropology. Forensic Sci Int 1996;83:147–53.
5. Angyal M, Derczym K. Personal identification on the basis of antemortem and postmortem radiographs. J Forensic Sci 1998;43(5):1089–93.
6. Culbert WL, Law FL. Identification by comparison of roentgenograms of nasal accessory sinuses and mastoid processes. JAMA 1927;88:1634–6.
7. Asherson N. Identification by frontal sinus prints. London: H. K., Lewis & Co. Ltd., 1965.
8. Duque CS, Casiano RR. Surgical anatomy and embryology of the frontal sinus. In: Stilianos E, Kountakis B, Wolfgang A, editors. The frontal sinus. Berlin, NY: Springer, 2005;21–31.
9. Figueroa RE, Sullivan J. Radiographic anatomy of the frontal sinus. In: Stilianos E, Kountakis B, Wolfgang A, editors. The frontal sinus. Berlin, NY: Springer, 2005;7–20.
10. Schuller A. A note on the identification of skulls by X-ray pictures of the frontal sinus. Med J Aust 1943;1:554–6.
11. Ubelaker DH. Positive identification from the radiographic comparison of frontal sinus patterns. In: Rathbun TA, Buikstra JE, editors. Human identification: case studies in forensic anthropology. Springfield, IL: Charles C Thomas Publisher, 1984;399–411.
12. Harris AMP, Wood RE, Nortje CJ, Thomas CJ. The frontal sinus: forensic fingerprint?—a pilot study. J Forensic Odontostomatol 1987;5: 9–15.
13. Reichs KJ, Dorion RBJ. The use of computerized axial tomography (CAT) scans in the comparison of frontal sinus configurations. Can Soc Forensic Sci J 1992;25:1–16.
14. Yoshino M, Miyasaka S, Sato H, Seta S. Classification system of frontal sinus patterns by radiography: its application to identification of unknown skeletal remains. Forensic Sci Int 1987;34:289–99.

15. Christensen AM. Assessing the variation in individual frontal sinus outlines. Am J Phys Anthropol 2004;127:291–5.
16. Seiden AM. Frontal headache. In: Kountakis SE, Senior BA, Draf W, editors. The frontal sinus. Berlin, NY: Springer, 2005;115–26.
17. Allen BC. Applied anatomy of paranasal sinuses. J Am Osteopath Assoc 1961;60:978.
18. Kirk NJ, Wood RE, Goldstein M. Skeletal identification using the frontal sinus region: a retrospective study of 39 cases. J Forensic Sci 2002;47(2):318–23.
19. Donald PL. Anatomy and histology. In: Donald PL, Gluckman JL, Rice DH, editors. The sinuses. New York, NY: Raven Press, 1995;24–48.
20. Meyer TK, Rhee JS, Smith TL. Frontal sinus fractures. In: Kountakis SE, Senior BA, Draf W, editors. The frontal sinus. Berlin, NY: Springer, 2005;132–42.
21. Fairgrieve SI. Forensic cremation: recovery and analysis. Boca Raton: CRC Press, 2008.
22. Moeller TB. Pocket atlas of radiographic anatomy, 2nd ed. Electronic resource. Germany: Thieme, 2000.
23. Rice DH. Embryology. In: Donald PL, Gluckman JL, Rice DH, editors. The sinuses. New York, NY: Raven Press, 1995;15–23.
24. Haglund WD, Fligner CL. Confirmation of human identification using computerized tomography. J Forensic Sci 1993;38(3):708–12.
25. Kullman L, Eklund B, Grundin R. Value of the frontal sinus in identification of unknown persons. J Forensic Odontostomatol 1990;8(1):3–10.
26. Daubert v. Merrell Dow Pharmaceuticals, Inc. 1993. 509 U.S. 579.
27. R v. Mohan. [1994] S.C.R. 9.
28. R v. J.-L.J. 2000 SCC 51, [2000] 2 S.C.R. 600.
29. Christensen AM. Testing the reliability of frontal sinuses in positive identification. J Forensic Sci 2005;50(1):1–5.
30. Reipert T, Ulmcke D, Schweden F, Nafe B. Identification of unknown dead bodies by X-ray image comparison of the skull using X-ray stimulation program FoXSIS. Forensic Sci Int 2001;117:89–98.
31. Ruiz CR, Wafae N. Anatomo-radiological and morpho-metrical study of the frontal sinus in humans. Braz J Mophol Sci 2004;21(1):53–6.
32. Schimmler JB, Helmer RP, Rieger J. Craniometric individuality of human skulls. In: Iscan MY, Helmer RP, editors. Forensic analysis of the skull. Craniofacial analysis, reconstruction, and identification. New York, NY: Wiley, 1993; 89–96.
33. Cameriere R, Ferrante L, Mirtella D, Rollo FU, Cingolani M. Frontal sinuses for identification: quality for classifications, possible error and potential corrections. J Forensic Sci 2005;50(4):1–4.
34. Reichs KJ. Quantified comparison of frontal sinus patterns by means of computed tomography. Forensic Sci Int 1993;61:141–68.
35. Tatlisumak E, Ovali GY, Aslan A, Asirdizer M, Zeyfeoglu Y, Tarhan S. Identification of unknown bodies using CT images of frontal sinus. Forensic Sci Int 2007;166:42–8.
36. Ribeiro FA. Standardized measurements of radiographic films of the frontal sinuses: an aid to identifying unknown persons. Ear Nose Throat J 2000;79:26–33.
37. Buckland-Wright JC. A radiographic examination of frontal sinuses in early British populations. Man 1970;5(3):512–7.
38. Koertvelyessy T. Relationship between the frontal sinus and climatic conditions: a skeletal approach to cold conditions. Am J Phys Anthropol 1972;37:161–72.
39. Adobe Systems Incorporated. Adobe Photoshop CS2® [computer program] 1990–2005. San Jose, CA: Adobe Systems Incorporated, 2005.
40. Rose P. Forensic speaker identification. New York, NY: Taylor and Francis, 2002.
41. Pfaelli M, Vock P, Dirnhofer R, Braun M, Bolliger SA, Thali MJ. Postmortem radiological CT identification based on classical ante-mortem x-ray examination. Forensic Sci Int 2007;171:111–7.

Additional information and reprint requests:
Scott I. Fairgrieve, Ph.D.
Chair, Department of Forensic Science
Laurentian University
935 Ramsey Lake Road
Sudbury, Ontario
Canada P3E 2C6
E-mail: sfairgrieve@laurentian.ca