JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, March 2010, Vol. 55, No. 2
doi: 10.1111/j.1556-4029.2009.01297.x
Available online at: interscience.wiley.com

TECHNICAL NOTE

ANTHROPOLOGY

Romina Ciaffi,[1,2] B.Sc.; Danilo De Angelis,[1] Ph.D., D.D.S.; Pier Federico Gherardini,[3] B.Sc.; Giovanni Arcudi,[2] M.D.; Renato Nessi,[4] M.D.; Gian Paolo Cornalba,[4] M.D.; Marco Grandi,[1] M.D.; and Cristina Cattaneo,[1] Ph.D., M.D.

Identification from Chest X-Rays: Reliability of Bone Density Patterns of the Humerus*

ABSTRACT: A critical review of Kahana and Hiss' study on identification from bone trabecular pattern and a test of their method conducted on the humerus are presented. Bone trabecular pattern was studied through the generation of a numerical file representing the gray scale. Using the correlation coefficient, several pairwise comparisons between numerical files were performed. The test gave nearly 30% of incorrect exclusions (the method did not recognize couples of radiographs belonging to the same subject) and 50% of misidentifications (the method recognized couples of radiographs belonging to different subjects, as belonging to the same subject); therefore, this research shows that at the present time, it is not possible to safely quantify identification through bone density patterns, of the proximal humerus taken from thoracic X-rays. Thus, an "easy"—but dangerous—use of trabecular density patterns on this specific type of radiogram as an identification method should be currently avoided.

KEYWORDS: forensic science, forensic anthropology, forensic radiology, human identification, trabecular bone pattern, quantification, thoracic radiograph, humerus

In cases of criminal investigations that involve unidentified decedents, one of the main issues is the attribution of identity. Whenever the remains are fragmentary, burned, water damaged, or badly decomposed and visual ID by next of kin, fingerprint identification, or DNA analysis are not applicable, positive identification is usually performed by forensic odontology. However, when this popular method cannot be applied, other methods of comparing antemortem and postmortem data are necessary. One of these is the comparison of bone morphology on antemortem and postmortem X-rays. In the case of bone morphology, however, it is impossible to give a quantitative response, as in the case of DNA; neither are there clear guidelines concerning, for example, how many points of similarity are to be considered sufficient to claim identity. For this reason, several groups in different disciplines, including forensic anthropology and odontology, are testing the methods brought to trials (1–9). However, several authors claim that there is no need for quantification, because expert knowledge, training, and experience are the only reliable features for a correct identification (10), and visual comparison proves sufficient to achieve identification, assuming that the expert has a high level of training and experience in the forensic practice (8,11–13). Regardless, the quest for quantitative or semiquantitative methods still goes on. The aim of this technical note is to analyze and test the applicability of one of the few techniques in the field, proposed by Kahana and Hiss (14) and by Kahana et al. (15), which attempts to quantify bone morphological similarity, on more difficult yet more common radiographic material. Kahana and Hiss' case reported in 1994 used hand radiographs and focused on metacarpal and phalanx bone architecture (14), while research conducted by Kahana et al., in 1998, used wrist radiographs and focused on the radius (15). Kahana et al. in their study state that the trabecular density pattern is unique to each individual and stable enough to be used as a forensic marker for the positive identification of human remains. They reached these conclusions through the study of postmenopausal women wrist radiographs, taken annually, for a period of 6 years, with the aim of studying the bone mineral status of these women. The lowest correlation coefficient that Kahana et al. obtained comparing same subject X-rays was 0.72 (r = 0.70–0.74, around 4% of same subject comparisons), while the highest correlation coefficient they obtained, comparing X-rays from different subjects, was 0.62 ($r > 0.5$, 5% of different subjects comparisons). Given these results, the authors claimed that the "cutting point" is 0.72 and the technique provides a quantifiable criteria to be presented in a court of law. Furthermore observation of such patterns seems reliable as it has been demonstrated that bone mineralization density does not change in a period of 10 years (14,16). The method is therefore described as successful; however, wrist and hand X-rays are not easy to come by in cases of unidentified decedents. Thoracic X-rays on the other hand are much more common. Clearly, they are performed for the purpose of visualizing soft tissues and are not optimal for the study of skeletal structures. However, in many cases of unidentified decedents, antemortem X-rays that are most commonly produced by family, physicians, or hospitals are thoracic

[1]Labanof, Istituto di Medicina Legale, University of Milan, Via Mangiagalli 37, 20133 Milano, Italy.
[2]Istituto di Medicina Legale, University of Rome "Tor Vergata", Via Montpellier 1, 00133 Rome, Italy.
[3]Centre for Molecular Bioinformatics, Department of Biology, University of Rome "Tor Vergata," Via della Ricerca Scientifica, 00133 Rome, Italy.
[4]Institute of Radiological Sciences, University of Milan, and S. Paolo Hospital, via Di Rudini 8, Milan, Italy.

*Presented at the XX[th] Congress of International Academy of Legal Medicine, August 23–26, 2006, in Budapest, Hungary.

Received 4 July 2008; and in revised form 28 Feb. 2009; accepted 15 Mar. 2009.

ones. The frequency of chest films in the US population is 43% (10) and in the Italian population is 21% (personal communication by Dr. C. Zocchetti, Direzione Generale Sanità, Regione Lombardia); furthermore, 53% of the identification cases reported by Murphy (17) were solved using these specific records. This is the reason why we chose to work on these types of radiographs. The scope was therefore to verify whether it is possible to apply Kahana et al.'s method on a different type of radiogram. As in Kahana et al.'s studies (14,15), we chose to work on what they describe as "trabecular bone pattern" but which in fact refers to bone density. The area chosen on the thoracic X-rays was the humerus, because among all the bones reproduced on a thorax X-ray, it is the one that offers the best visualization of the trabeculae. In this study, the trabecular pattern of the humerus visible in most thoracic X-rays was studied through the selection of a line ("reference line") and the generation of a numerical file representing trabecular density along the line of interest. Several comparisons were performed (comparisons between X-rays belonging to the same subject and comparisons between X-rays belonging to different subjects) using Pearson's correlation coefficient.

Materials and Methods

Two sets of thorax radiographs of living subjects were collected. The first set, taken between 1960 and 1990, included 191 X-rays from 88 subjects, only 25 of which were used. The second set, taken between 2005 and 2006, included 108 X-rays belonging to 59 subjects, among these X-rays, only 46 belonging to 25 subjects were used. X-rays from subjects in which skeletal development was not complete were excluded. For each subject, one, two, or more X-rays were available (taken within 10 years for the first set and between 1 min and 12 months for the second set); this allowed AM/PM simulated comparisons. X-rays from the first set (1960/1990) were placed on a lightbox, and then a black and white picture was taken using a camera (Nikon Coolpix 3200; resolution 2048 × 1536; exposure 1/60; ISO 50; focal length 17.4). Radiographs from the second set (2005/2006) were scanned using a "Scanner Epson Expression 1640 XL," and the digitization has been conducted using the software "Photoshop 7.0" (18), and the resolution used was 300 dpi. The same criteria for the image capture methods and the same softwares were used for the two sets and, as demonstrated by the results, the two methods are in fact comparable. Moreover, they are comparable to the image capture methodologies applied by Kahana et al. in their 1998 study. In fact, they digitized the X-rays using a view table and a CCD computer-controlled video camera, and the sole diversity between the two studies is inherent to technological advances in digital image acquisition. Accordingly, the resolution we used was higher than the one used by Kahana et al., and the gray scale level was the same (15).

A method to standardize the images was devised (the "standardized method") that involved resizing of the humeri and positioning of a "reference line." The latter issue was not analyzed in Kahana et al.'s research (15), but it was discussed in Israel's work (19). Initially, Pearson's correlation coefficient was used to individualize the best "reference line" between four lines positioned on the X-rays. In a preliminary part of the study, in fact, a test was performed to verify which area at the humeral neck was best for comparison. Four lines were traced, within the first set, beneath the humeral head in the following sequence: the first at the most proximal point of the diaphysis, the last one at the distal extremity of the humeral head, and the other two randomly placed in between (Fig. 1). Results of this preliminary test showed that the most user-friendly and efficient comparisons were obtained on the first line

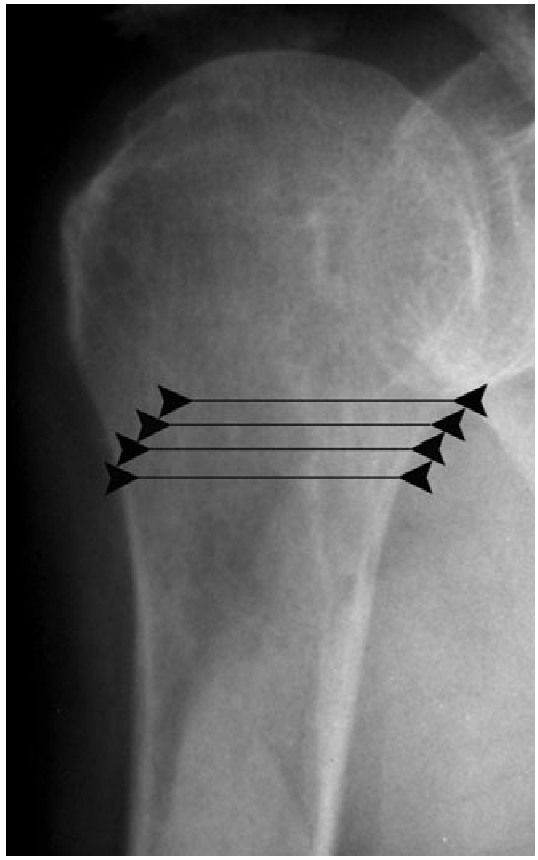

FIG. 1—*Detail from a chest film. The proximal epiphysis of a right humerus, with the four couples of arrows and the four lines used to select the best reference line (the lowest line in the image).*

(Table 1), i.e., the lowest one on the humerus, at the most proximal point of the diaphysis. Hence, this became the reference point for the rest of the study. The "standardized method" was devised using "Adobe Photoshop CS" software (20). Two humeri (one right humerus and one left humerus) were selected among the sample; an outline was shaped, and then two arrows were positioned at the reference point. These images (the "standard images") of the outlines and the arrows (after the removal of the original humerus) (Fig. 2) were then superimposed to all the humeri studied; each head of the humerus was brought to the standard dimension, without altering the proportions of the bone, and the arrows were then used to obtain a line at the reference point (the same line in each humerus analyzed). As in Kahana et al.'s research, the trabecular pattern was studied through the generation of numerical files representing the gray scale along the reference line; in the present research, these files were obtained using "ImageJ" (21), an open source scientific software, while Kahana et al. used a software specifically designed for their study (15); it should be noted that as the gray scale measurement is a straightforward calculation performed

TABLE 1—*Correlation coefficients obtained when testing the "standardized method" on the first set of X-rays (thorax X-rays,1960 and 1990). Each radiograph was compared to itself; all lines were considered (1n, 2n, 3n, and 4n).*

Correlation	Line 1n		Line 2n		Line 3n		Line 4n	
<0.60	0%	0	10%	2	10%	2	10%	2
>0.60	100%	22	90%	20	90%	20	90%	20

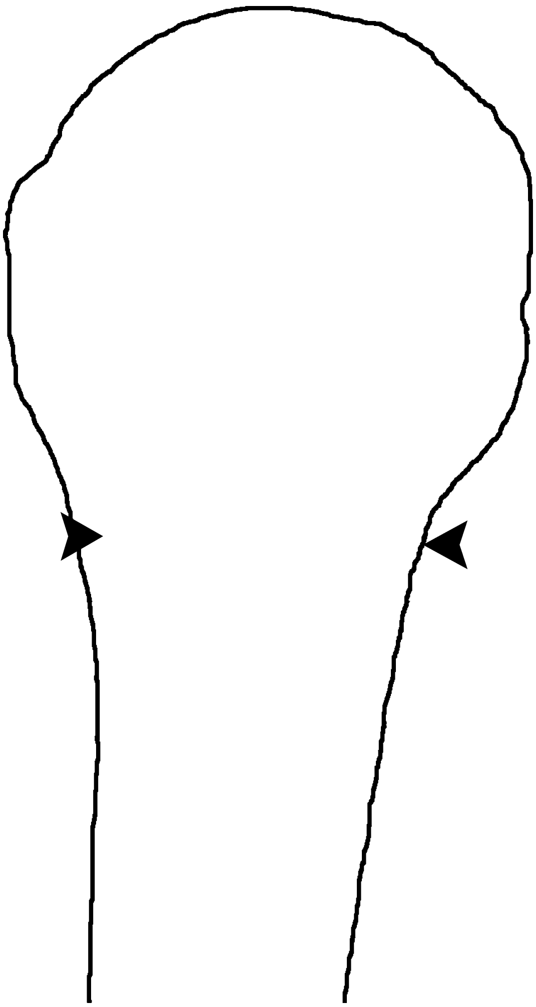

FIG. 2—*Outline and arrows used for the comparisons of the radiographs. This outline was superimposed on every humerus, in order both to resize the humeri and to place the reference line for the measurement of bone density patterns.*

on a digital file, the fact that different software packages were used does not influence the outcome of the analysis. The statistical analysis was conducted by comparing the numerical files using Pearson's correlation coefficient, as carried out in Kahana et al. study (15). All the possible pair-wise comparisons were made; numerical files obtained from different radiographs of the same subject were compared to each other, and comparisons between numerical files taken from radiographs of diverse subjects were also made. To test how many times the technique did not match two radiographs belonging to the same subject (false negatives, incorrect exclusions), different radiographs from the same subject were compared; to test how many times radiographs belonging to different subjects were matched as belonging to the same subject (false positives, misidentifications), comparisons were performed among X-rays from different subjects. As the aim of the comparisons was to verify whether the two lines had some common variation, Pearson's correlation coefficient was used. X-rays taken in different conditions might have a different luminosity and a different contrast; the former feature does not alter the correlation results because a variation in luminosity acts as a linear transformation on all the pixels of an image. Conversely, differences in contrast could theoretically alter the results although, both in this study and in Kahana et al.'s research (15), the outcome does not appear to be influenced by this feature. The statistical analysis was performed using "R" (22), a programming environment for data analysis and graphics.

Results and Discussion

As mentioned earlier, in the preliminary study performed to verify where to trace the reference line, tests on four lines were performed. Among the four lines selected, line 1n (the lowest line in Fig. 1), i.e., at the most proximal point of the diaphysis, gave the best outcome, resulting in 100% of the comparisons having $r > 0.60$ (in 82% of comparisons $r > 0.82$, Table 1). The results that follow in the rest of the study obviously include only line 1n, as this, from the beginning, was chosen as the reference line. The comparisons conducted between lines obtained from different radiographs, in the first set, belonging to the same subject gave an incorrect exclusion (false negative) in 26% of cases ($r < 0.60$, lowest $r = -0.24$) (Table 2). When comparing radiographs from different subjects, the quality of the results obtained did not improve; in fact, 49% of the comparisons were misidentifications (false positives) ($r > 0.60$, greatest $r = 0.94$) (Table 3). As the results obtained on the first set could have been attributed to the deterioration of the radiographic material used, a second set of more recent thoracic radiographs, taken between 2005 and 2006, were investigated. From the analysis of this set, when comparing X-rays from the same subjects, 38% of incorrect exclusions (false negatives, $r < 0.60$, lowest $r = -0.23$) were obtained (Table 2). Moreover, the comparisons between X-rays from different subjects gave 58% of misidentifications (false positives, $r > 0.60$, greatest $r = 0.97$) (Table 3). Although the number of comparisons conducted between different X-rays belonging to the same subject was not high (set I 15 comparisons and set II 26 comparisons, Table 2), the number

TABLE 2—*Correlations obtained comparing lines from different X-rays belonging to the same subject, thus simulating AM/PM comparisons, on the two sets of radiographs (considering only line 1n).*

	I Set (Thorax 1960/1990)		II Set (Thorax 2005/2006)	
Correlation	Percentage	No. of Comparisons	Percentage	No. of Comparisons
−0.30 to 0.00	6.66	1	3.85	1
0.01 to 0.20	0	0	11.54	3
0.21 to 0.40	0	0	3.85	1
0.41 to 0.60	20.00	3	19.23	5
0.61 to 0.80	20.00	3	23.1	6
0.81 to 1.00	53.33	8	38.46	10
Total		15		26

TABLE 3—*Correlations obtained comparing lines taken from radiographs belonging to different subjects, on the two sets of radiographs (considering only line 1n).*

	I Set (Thorax 1960/1990)		II Set (Thorax 2005/2006)	
Correlation	Percentage	No. of Comparisons	Percentage	No. of Comparisons
−0.80 to −0.41	0	0	3.41	20
−0.40 to 0.00	3.20	4	3.58	21
0.01 to 0.20	7.20	9	7.51	44
0.21 to 0.40	12.80	16	10.24	60
0.41 to 0.60	27.20	34	17.06	100
0.61 to 0.80	35.20	44	26.10	153
0.81 to 1.00	14.4	18	32.10	188
Total		125		586

of comparisons between different X-rays belonging to different subjects was consistent (set I 125 and set II 586, Table 3). This outcome, obtained using a methodology comparable to Kahana et al.'s methodology, shows how it is not possible to perform a reliable identification based on linear density patterns on the proximal humerus from thoracic X-rays. This study has focused exclusively on humeral bone structure as can be seen on chest X-rays; regardless, future research on X-rays of different regions of the skeleton might demonstrate the applicability of such a method for positive identification, as already demonstrated by Kahana et al. with wrist trabecular bone density pattern comparison, presented as an ideal method of identification (15), although the present work cautions against the "easy" adaptation of the method to other bones and radiographs. An issue which needs to be addressed, and which may be a possible source of error, when using thoracic radiographs, is the possible slight differences in the positioning of the arms. Further research should clarify the effect of slight posture variations on trabecular density patterns. In a real case of identification, however, one could exactly position the arm of the cadaver in the PM X-ray as it was positioned in the AM radiograph. As previously stated, if taken in a period of 10 years, two X-rays can be reliably compared for the trabecular bone density pattern analysis (16). In Kahana and Hiss' study (14), the time lag between the first X-rays and the second X-rays, belonging to the same person, was never more than 6 years, and in this study, it was never more than 10 years.

This research shows that at the present time, it is not possible to safely quantify identification through bone density patterns of the proximal humerus taken from thoracic X-rays. In other words, the authors only wish to warn against an "easy"—but dangerous—use of trabecular density patterns on this specific type of radiogram (commonly available) as an identification method. Thus, at the moment, as several authors state, the sole useful and reliable approach may be a morphological analysis by an expert, which can itself be evaluated through the measurement of observer accuracy (1,2).

Acknowledgments

The authors thank the Sant'Anna Hospital of Como and the Institute of Radiological Sciences of the University of Milan, for the collection of X-rays. Thanks are also due to Dr. C. Zocchetti, of the "Direzione Generale Sanità, Regione Lombardia."

References

1. Koot MG, Sauer NJ, Fenton TW. Radiographic human identification using bones of the hand: a validation study. J Forensic Sci 2005;50 (2):263–8.
2. Kuehn CM, Taylor KM, Mann FA, Wilson AJ, Harruff RC. Validation of chest X-ray comparisons for unknown decedent identification. J Forensic Sci 2002;47 (4):725–9.
3. Adams BJ. Establishing personal identification based on specific patterns of missing, filled, and unrestored teeth. J Forensic Sci 2003;48 (3):487–96.
4. Christensen AM. The impact of Daubert: implications for testimony and research in forensic anthropology (and the use of frontal sinuses in personal identification). J Forensic Sci 2004;49 (3):427–30.
5. Christensen AM. Testing the reliability of frontal sinuses in positive identification. J Forensic Sci 2005;50 (1):18–22.
6. Hogge JP, Messmer JM, Doan QN. Radiographic identification of unknown human remains and interpreter experience level. J Forensic Sci 1994;39 (2):373–7.
7. Komar D, Lathrop S. Frequencies of morphological characteristics in two contemporary forensic collections: implications for identifications. J Forensic Sci 2006;51 (5):974–8.
8. MacLean DF, Kogon SL, Stitt LW. Validation of dental radiographs for human identification. J Forensic Sci 1994;39 (5):1195–200.
9. Rogers TL, Allard TT. Expert testimony and positive identification of human remains through cranial suture patterns. J Forensic Sci 2004;49 (2):203–7.
10. Brogdon BG. Forensic radiology. New York, NY: CRC Press, 1998.
11. Mundorff AZ, Vidoli G, Melinek J. Anthropological and radiographic comparison of vertrebrae for identification of decomposed human remains. J Forensic Sci 2006;51 (5):1002–4.
12. Pretty IA, Pretty RJ, Rothwell BR, Sweet D. The reliability of digitized radiographs for dental identification: a web-based study. J Forensic Sci 2003;48:1325–30.
13. Soomer H, Lincoln MJ, Ranta H, Penttila A, Leibur E. Dentists' qualifications affect the accuracy of radiographic identification. J Forensic Sci 2003;48:1121–6.
14. Kahana T, Hiss J. Positive identification by means of trabecular bone pattern comparison. J Forensic Sci 1994;39 (5):1325–30.
15. Kahana T, Hiss J, Smith P. Quantitative assessment of trabecular bone pattern identification. J Forensic Sci 1998;43 (6):1144–7.
16. Roschger P, Gupta HS, Berzlanovich A, Ittner G, Dempster DW, Fratzel P, et al. Constant mineralisation density distribution in cancellous human bone. Bone 2003;32 (3):316–32.
17. Murphy WA, Spruill FG, Gantner GE. Radiologic identification of unknown human remains. J Forensic Sci 1980;25 (4):727–35.
18. Adobe Photoshop 7.0. Version 7.0. Adobe, 2002, http://www.adobe.com/support/downloads/product.jsp?platform=windows&product=39.
19. Israel H. Microdensitometric analysis for study of skeletal growth and aging in the living subject. Am J Phys Anthropol 1968;29 (2):287–94.
20. Adobe Photoshop CS. Version 8.0.1. Adobe, 2003, http://www.adobe.com/support/downloads/product.jsp?platform=Macintosh&product=26.
21. ImageJ. Version 1.34s. National Institutes of Health (USA), 2005, http://wareseeker.com/free-imagej-1.34s/.
22. R. Version 2.3.0. R Development Core Team, 2006-04-24, http://www.r-project.org.

Additional information and reprint requests:
Romina Ciaffi, B.Sc.
Istituto di Medicina Legale
Università di Milano
Via L. Mangiagalli 37
20133 Milano
Italy
E-mail: ciaffi@med.uniroma2.it