PAPER

PHYSICAL ANTHROPOLOGY

Joan A. Bytheway,[1] Ph.D. and Ann H. Ross,[2] Ph.D.

A Geometric Morphometric Approach to Sex Determination of the Human Adult Os Coxa

ABSTRACT: Sex determination of the human skeleton is best assessed from the os coxa. The present study explored the possibility of using three-dimensional landmark coordinate data collected from various landmarks located over the entire bone to determine whether there were significant sex differences local to the landmarks. Thirty-six landmarks were digitized on 200 African American and European American male and female adult human os coxae. MANCOVA results show that sex and size have a significant effect on shape for both European Americans (Sex, $F = 17.50$, d.f. = 36, 63, $p > F = 0.0001$; Size, $F = 2.56$, d.f. = 36, 63, $p > F = 0.0022$) and African Americans (Sex, $F = 21.18$, d.f. = 36, 63, $p > F = 0.0001$; Size, $F = 2.59$, d.f. = 36, 63, $p > F = 0.0005$). The discriminant analysis shows that sexing accuracy for European Americans is 98% for both males and females, 98% for African American females, and 100% for African American males.

KEYWORDS: forensic science, forensic anthropology, sex determination, os coxa, geometric morphometrics, discriminant function analysis

The os coxa is the most reliable bone of the human skeleton to determine sex and when used alone it can achieve approximately 90% accuracy (1–4). Sex determination using the os coxa can be accomplished through visual observation of anthroposcopic characteristics or metric analysis of linear measurements (5) and both methods have shown high accuracy percentages. For example, Steyn and İşcan's (6) metric analysis on a modern Greek population showed that most measurements were "highly repeatable" and when all variables are present, the average accuracy is 95.4%. Phillip Walker (7) found that visual observation of the greater sciatic notch produces "about the same proportion of correct sex assignments" as using metric methods.

There have been a number of observed pelvic features used over the years that have been successful in determining sex. Traditionally, the pelvic inlet (8,9) was used. The female pelvic inlet increases in diameter during puberty. Its shape changes from a pre-adolescent elliptical form to a more open, rounded aperture. Males retain the more elliptical, constricted aperture. This change in the female pelvic inlet initially was believed to be the result of the increase of hormonal activity (9). More recent studies (10–17) attribute pelvic inlet or pelvic canal change to "functional adaptations to certain modes of locomotion" and maternal–fetal size relationships (16). In addition to the pelvic inlet, features such as the subpubic angle, pubic symphysis height, and the shape of the obturator foramen (18,19) were used. With growth, the female subpubic angle becomes more obtuse than in the male. The pubic symphysis becomes shorter and the obturator foramen becomes more triangular (20). One problem using the subpubic angle is that although sex assessment can be made without complete articulation of the pelvis,

it is best assessed when it is. Therefore, it could be argued that using only one side does not result in the most accurate evaluation (20). Because of the frequent fragmentary nature of pelvic remains articulation is not always possible. The second problem exists when observing the pubic symphysis height. With the frequency of damage at the symphysis, symphysis height could only be estimated (20), or the pelvis would be excluded from the analysis. Thirdly, the observer would have to estimate the amount of cartilage at the pubic symphysis to determine height (20). This may lead to interobserver error. In general, these features are now considered to be less reliable than other features.

More reliable features observed over the years include the angle of the greater sciatic notch, the width of the pubis, and the preauricular sulcus. Because of the increase in diameter of the pelvic inlet, the inferior portion of the ilium (21) and the greater sciatic notch become wider in females as does the pubic bone. Some researchers have found that these features tend to "show varying degrees of male and femaleness" (22) and there is no set standard as to which feature should be more heavily weighted when features have intermediate values. In 1969, Phenice developed a new observational method for sexing the pelvis that has a 95% accuracy and can be used by less experienced as well as experienced observers. It is also less time consuming than measurement techniques. Phenice's work (22) showed that the ventral arc, the subpubic concavity, and the medial aspect of the ischiopubic ramus are each "essentially discrete in character" and that if one feature were missing, there will almost always be one of the features present and will be "obviously indicative of male or female" (22). Phenice did claim, however, that there are racial differences in the accuracy of the technique. Lovell (23) found that age difference also affects the accuracy of the technique. She found that older individuals (than those in Phenice's study) showed degenerative changes in the pubic bone. Lovell reported that these changes may disguise the manifestation of one or more of the features used by Phenice. Lovell's average accuracy was $83 \pm 7\%$. She attributed this to the

[1]College of Criminal Justice, Forensic Science, Sam Houston State University, Huntsville, TX 77341-2525.
[2]Department of Sociology and Anthropology, North Carolina State University, Raleigh, NC 27695-8107.

uncertainty of the observers when observing bones of older individuals whose bony surfaces were more irregular than younger individuals.

In 1990, McLaughlin et al. (24) found that the accuracy of Phenice's method on three European populations decreased dramatically, with the ventral arc and the ischiopubic ramus being the least reliable indicators of sex. Because Phenice's descriptions of each feature are somewhat ambiguous and the addition of Sutherland and Suchey's (25) addendum to the ventral arc description, McLaughlin claimed that the variability in interpretation of the defined features leaves the method vulnerable to higher levels of interobserver error, which ultimately results in differing levels of accuracy. Rogers and Saunders (5) tested the reliability and accuracy of 17 pelvic traits. Of the 17 examined, four of the most effective features were the true pelvis shape, length of the pubis, subpubic concavity, and the ventral arc. Rogers and Saunders found a higher than acceptable level of intraobserver error (11.3%) for all pelvic traits. The four most problematic areas of error included the acetabulum shape and size, auricular surface height, preauricular sulcus, and ischiopubic ramus shape. Although the acetabulum can be extremely accurate, it is hard to assess (5). Rogers and Saunders (5) found that "with precision and accuracy," the ventral arc of the os coxa is the best method for determining sex. In addition to the ventral arc, other reliable features for sex determination include the subpubic concavity and the medial aspect of the ischiopubic ramus (22). More recent contributions to sex determination include the McBride et al. (26) quantitative approach using ID3 algorithm and other visual approaches using the preauricular sulcus, inferior pelvis, ischiopubic region, (27) and the greater sciatic notch (27,28).

Some of the problems with using anthroposcopic traits of the pelvis are (i) the negative correlation that exists between age and accuracy of identification because of degeneration of bony features (23), (ii) some features, such as the ventral arc, do not become distinct until the third decade (22,23,25,29,30), and (iii) some of the features are commonly missing because of taphonomic or postmortem disturbances, such as scavenging, erosion, or breakage.

Traditional metric analyses of the human skeleton have been claimed by some to be more repeatable and more case-inclusive methods of determining sex (31,32), which can be performed by less experienced practitioners (22) and can sometimes expose significant areas of variation that may not be readily recognizable via visual observation. Furthermore, quantitative methods can also identify differences when "within-group variability is low" (33). One example of a popular traditional metric analysis is the ischiopubic index developed by Washburn in 1948. Washburn found that the index alone has a reliability of 90% when used on pelves of one major population. Other more recent metric approaches include Steyn and Işcan's (6) discriminant function for sex determination using a modern Greek population, Kurki's work (10) on obstetric canal protection in small-bodied populations, and Pretorius et al.'s (34) study using geometric morphometric analysis of three bony elements: the eye orbit, the ascending ramus, and the sciatic notch. Traditional metric techniques that rely strictly on linear measurements are not always able to capture the underlying shape variation because that variation may not lie along the span of the calipers (35,36). There have been several three-dimensional studies of the os coxa; however, they have been limited to clinical applications for orthopedic treatment of patients with acetabular fractures (37) and assessments of acetabular fractures (38).

The purpose of this study is to apply modern morphometric techniques to quantify the shape of the os coxae and possibly identify new areas on the os coxa that could be used for sex estimation. Because methods from the geometric morphometrics may make it possible to locate regions of shape variation using a reduced set of landmarks, locating new areas of the pelvis to use for sex estimation could have applicability for forensic fragmentary skeletal remains and applicability in bioarchaeology.

Materials and Methods

Sample

A sample of 200 intact os coxae of documented sex, age, and population was selected from the Terry Collection, Smithsonian Institute, Washington, D.C. The sample consists of the left os coxae of 50 European American adult males and females and 50 African American adult males and females. The mean age and standard deviation of the entire sample is 34 ± 7 years.

Landmarks

Thirty-six landmarks were identified on male and female os coxae, as shown in Figs. 1 and 2. These are also listed in Table 1. The landmarks chosen are based on various landmarks found throughout metric and geometric landmark literature (13,39–42). They were chosen to potentially capture the entire shape variation of the os coxa. Landmarks 11, 13–16, 27, 29–32, 34–36 are newly derived landmarks developed by the first author (43).

The 36 landmarks are categorized using three general categories based on Lele and Richtsmeier (39) and Bookstein's (40) landmark classifications. An attempt was made to categorize all landmarks under Bookstein's three landmark types; however, because of the variety of landmark locations, this was not possible. Definitions of some landmarks (e.g., constructed) did not fit precisely into one of the three landmark types and some, by definition, fit much easier into Lele and Richtsmeier's categories. These landmarks were chosen for various reasons: (i) they are considered useful and significant in the literature, (ii) they are repeatable, (iii) they represent the object being studied, (iv) they represent regions that are considered reliable for sexing the pelvis, and (v) they represent regions that have not been metrically considered for sex determination but are addressed in this research.

Traditional or Type 2 landmarks are defined as "precisely delineated points corresponding to the location of features of some biological significance" (39), or "points of application of real
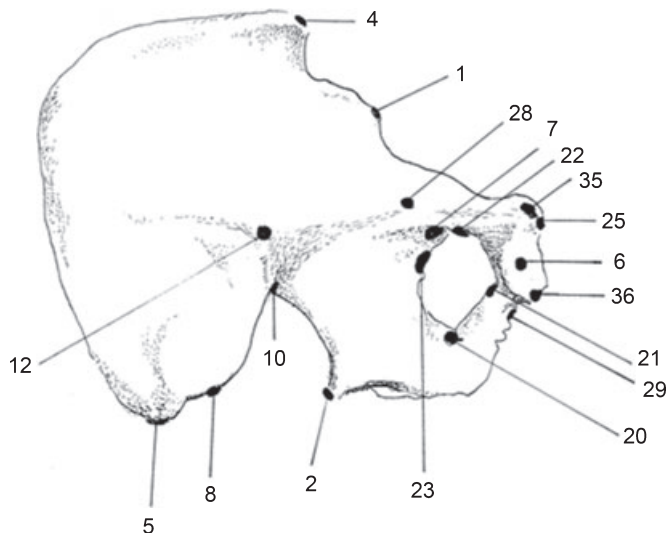

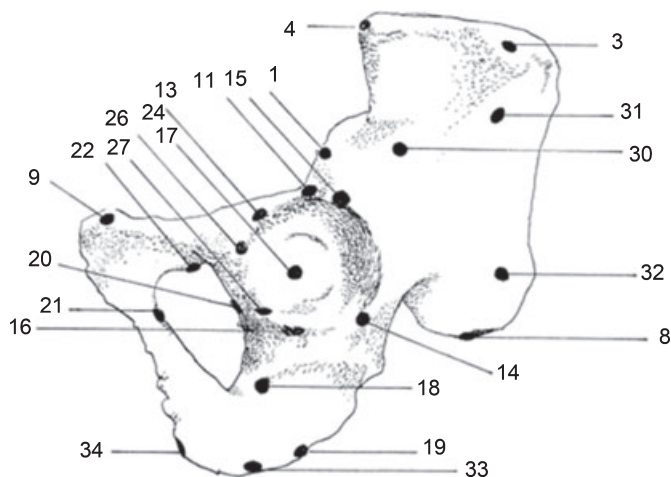FIG. 1—*Medial view of the human adult os coxa showing landmarks.*

FIG. 2—*Lateral view of the human adult os coxa showing landmarks.*

TABLE 1—*List of landmarks.*

Landmarks and Descriptions
Traditional landmarks
1. Anterior inferior iliac spine
2. Ischial spine
3. Iliac tubercle
4. Anterior superior iliac spine
5. Posterior superior iliac spine
6. Pubic symphysis
7. Obturator groove
8. Posterior inferior iliac spine
9. Pubic tubercle
10. Apex inside greater sciatic notch/maximum curvature point
11. Directly inferior to the anterior inferior iliac spine to a point on the acetabular rim
12. Auricular surface apex (Seidler 1980)
Constructed landmarks
13. and 14. The two points on the acetabular rim that produce the maximum horizontal diameter of the acetabulum
15. and 16. The two points on the acetabular rim that produce the maximum vertical diameter of the acetabulum
17. The acetabular point ("Found by drawing two lines at right angles to each other across the acetabulum") (Lele and Richtsmeier 2001)
18. and 19. The two points on the ischium that produce maximum width of the ischial tuberosity (Steudel 1981)
20. and 21. The two points on the obturator foramen rim that produce the maximum breadth of the obturator foramen (perpendicular to the line of maximum length) (Coleman 1969) without using the obturator groove.
22. and 23. The two points on the obturator foramen rim that produce maximum length of the obturator foramen (Coleman 1969)
24. and 25. The two points that produce pubic length (using the acetabular point to "the most medial point on the body of the pubis") (Steudel 1981)
Extremal landmarks
26. The tip of the inferior acetabular lip (Coleman 1969) (the most inferior point at the beginning of the internal edge of the anterior rim of the acetabulum)
27. The most inferior point at the beginning of the internal edge of the posterior rim of the acetabulum.
Traditional (Fuzzy) landmarks
28. Arcuate (iliopubic) eminence
29. Ischiopubic ramus—at the narrowest point inferior to the pubic symphysis
30. Inferior gluteal line—directly above #11 on the gluteal line
31. Anterior gluteal line—at or near the large foramen on the gluteal line
32. Posterior gluteal line—directly superior to the superior posterior iliac spine on the gluteal line
Extremal (Fuzzy) landmarks
33. Most posterior point on ischial tuberosity
34. Most inferior point on the ischial tuberosity
35. Most superior point on the pubic bone at symphysis
36. Most inferior point on the pubic bone at symphysis

biomechanical forces, pushes or pulls" (40). Landmarks in this category are single points on an object.

Constructed landmarks are "points corresponding to locations that are defined using a combination of traditional landmarks and geometric information" (e.g., an arc formed by connecting two traditional landmarks and a point designated along that arc) (39). These landmarks are also single points.

Extremal landmarks or Type 3 are those located at an extreme point, such as "the most inferior point…" (39) or "points taken as 'farthest' from other points" (40). Although not as meaningful as other landmarks (40), they are included in this study.

Traditional and *Extremal* landmarks also have a subcategory labeled *Fuzzy* landmarks. *Fuzzy* landmarks are landmarks of precise delineation and acknowledged as being an area of biological significance, having no single point but rather occupying an area greater than a single point (e.g., arcuate eminence) (39). Although the stylus of the digitizer touches the landmark in a single point, the area of the landmark is not a single point.

Each os coxa was initially examined for the quality of preservation. Os coxae that were damaged or were missing any of the 36 landmarks were excluded. Those excluded were replaced by another os coxa that contained all 36 landmarks. Each os coxa was then clamped into a vertically oriented vise and the 36 landmarks were digitized. Prior to digitizing, landmarks 13 through 25 (constructed landmarks) had to be measured with a sliding caliper to obtain maximum measurements, which were marked with a pencil. A Microscribe® 3D digitizer (Revware, Raleigh, NC) was used to register the x, y, and z coordinates in a Microsoft Excel spreadsheet.

Geometric Morphometric Analysis

Raw coordinate data are not directly comparable as shape variables to compare specimens because each set was collected in its own coordinate system. A generalized Procrustes analysis or GPA was used to bring all specimens into a common coordinate system to minimize the sum of squared distances between landmarks of each os coxa and those of an iteratively computed mean scaling all specimens to unit centroid size (CS). The GPA superimposition was carried out with the software Morpheus et al., written by Dennis Slice (44). The resulting shape variables and CS were used in the subsequent multivariate analyses (for a more in-depth discussion of Geometric Morphometrics see Slice [45]). A principal component analysis (PCA) of the covariance matrix was conducted on

the GPA transformed variables to reduce the dimensionality of the data. A multivariate analysis of covariance (MANCOVA) was performed using the PCA scores to test whether size and sex have significant effects on the average shape of males and females of each population. In addition, independent group *t*-tests were used to compare the mean CS of sexes for each population. Separate discriminant function analyses using the cross-validation or *leave-one-out* method were performed for each population using the PCA scores. The multivariate analyses were performed using the SAS System for Windows version 9.1.3 (SAS Institute) (46).

Results

The MANCOVA results are presented in Table 2. The MANCOVA procedure did not detect a significant size-by-sex interaction for either population. However, both size and sex were

found to have a significant effect on both European and African Americans.

The independent group *t*-test shows that the male CS mean is significantly different from the female mean for both populations (European Americans $t = -3.98$, $p < 0.0001$, d.f. 98; African Americans $t = -4.99$, $p < 0.0001$, d.f. 98). In addition, the sexes themselves differ significantly across groups (females $t = -4.38$, $p < 0.0001$, d.f. 98; males $t = -3.08$, $p < 0.0021$, d.f. 98).

The cross-validated discriminant function analysis using the first 23 principal component scores accounting for 86% of the total variation produced sexing accuracies for European Americans of 98% for females and 98% for males (Table 3). The cross-validated discriminant function analysis using the first 22 principal component scores accounting for 85% of the total variation produced sexing accuracies for African Americans of 98% for females and 100% for males (Table 4).

The sex-specific pelvic shape variation is illustrated by plots showing the difference vectors, which convey the relative direction and magnitude of difference between one configuration and another. Figures 3 and 4 are a medial and lateral view of the African American female mean (large gray circles) compared to the African American male mean (small black circles) after GPA superimposition. The dissimilarity between the two groups is reflected by the length of the vectors, which have been magnified by a factor of 3. As would be expected, the landmarks on the pubic symphysis (landmarks 6, 25, 35, and 36) of the female mean are much more medially placed than in the male mean. And those landmarks associated with the ischium and ischiopubic ramus (landmarks 2, 18, 19, 29, 33, and 34) are more posteriorly placed

in females. The landmarks along the iliac crest (1, 3, and 4) are more inferiorly oriented, and those along the posterior iliac spine (5 and 8) are more posteriorly oriented in the female configuration. The auricular point (14) has a more lateral orientation, and the arcuate eminence (30) is positioned more posteriorly in the female mean. Interestingly, the shape of the obturator foramen (landmarks 7, 21, 22, and 23) and acetabulum (landmarks 11, 13, 14, 15, 16, 24, and 26) do not appear to be significantly different between males and females. The sex variation pattern between females (large gray circles) and males (small black circles) is similar in the European American sample (Fig. 5). However, the difference vectors are somewhat shorter in the European sample suggesting that they may be less sexually dimorphic than the African American sample.

Discussion

The os coxa was chosen for the study because of its characteristics that differentiate males from females better than any other single bone. The pelvis, however, is subjected to taphonomic damage. Skeletal remains from archeological or forensic circumstances repeatedly demonstrate the effects of postmortem trauma to bone (30,47–49). In cases where the os coxa is fragmented and reliable,

TABLE 2—*Multivariate analysis of covariance results.*

	Wilks' Lambda	F	d.f.	p > F
African Americans				
Size*sex	0.642	0.95	36, 61	0.5629
Sex	0.091	17.50	36, 63	0.0001
Size	0.436	2.56	36, 63	0.0022
European Americans				
Size*sex	0.627	1.03	36, 61	0.4550
Sex	0.076	21.18	36, 63	0.0001
Size	0.403	2.59	36, 63	0.0005

TABLE 3—*Cross-validated European American classification summary of shape variables using the first 23 PC scores accounting for 86% of the total variation.*

Sex	Frequency of Female Classification (n)	Frequency of Male Classification (n)
Female	98% (49/50)	0
Male	0	98% (4950)
Total	49% (49/100)	49% (49/100)

TABLE 4—*Cross-validated African American classification summary of shape variables using the first 22 PC scores accounting for 85% of the total variation.*

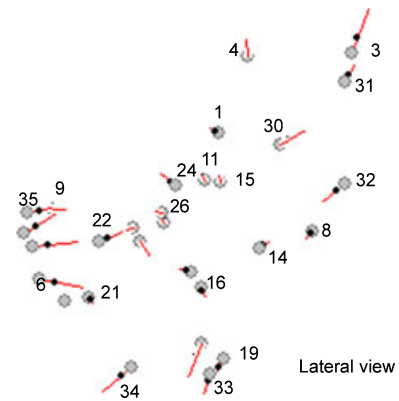| Sex	Frequency of Female Classification (n)	Frequency of Male Classification (n)
Female	98% (49/50)	2% (1/50)
Male	0	100% (50/50)
Total	49% (49/100)	51% (51/100)

FIG. 3—*Medial view showing the difference vectors of the African American female mean (large gray circles) to male mean (small black circles) (X Factor of 3). The difference vectors illustrate the direction and magnitude of difference between the corresponding landmarks of the mean configurations.*
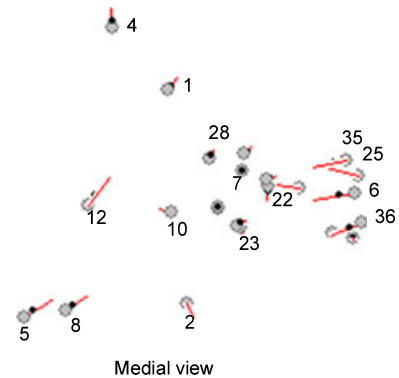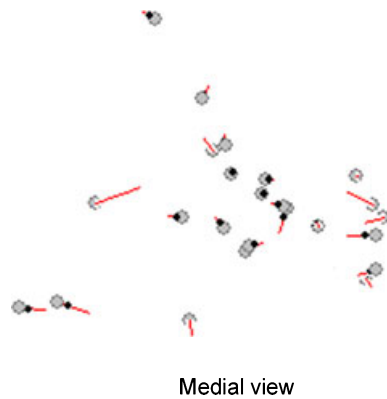

FIG. 4—*Lateral view showing the difference vectors of the African American female mean (large gray circles) to male mean (small black circles) (X Factor of 3). The difference vectors illustrate the direction and magnitude of difference between the corresponding landmarks of the mean configurations.*

Medial view

FIG. 5—*Medial view showing the difference vectors of the European American female mean (large gray circles) to male mean (small black circles) (X Factor of 3). The difference vectors illustrate the direction and magnitude of difference between the corresponding landmarks of the mean configurations.*

observational, bone features are absent, the os coxa is excluded from the analysis, resulting in a decrease in accuracy of determining sex.

Previous studies have shown that there is a high degree of accuracy when using the os coxae for sex determination (3,4,7,50). Historical literature on the problems with observational methods (22,23,25,30,48,49) and metric analysis was a factor in the design of the present study. The minimal amount of current research utilizing digitized 3D landmark data on the os coxa and the concept of it being used in future studies were also components of the study's design. Other factors that were considered for the use of a morphometric technique were (i) landmarks are repeatable and (ii) inexperienced users could use the method without affecting accuracy of the technique. Although digitizing is straightforward, the accuracy of identifying these landmarks by an experienced or inexperienced user (other than the authors) is not known. An inter/intraobserver study is underway and will be addressed in a separate publication.

The goal of the present study was to identify three-dimensional landmarks and/or regions on the human adult os coxae that could be used to determine sex. These high correct classification rates obtained in the present study further demonstrate that the os coxa is a reliably sexually diagnostic region of the skeleton. Furthermore, the vector plots illustrate that the pubis, ilium, and ischium are the most sexually dimorphic regions of the os coxa. The obstetric modifications observed in the pubis that allow for an increased pelvic inlet in females would have to be compensated for in other areas such as the ischium and ilium to retain bipedal locomotion. The shape of the male pelvis, in contrast, is related to evolutionary adaptations to bipedality. Interestingly, differences were not evident in either the acetabulum or the obturator foramen. Although African Americans and European Americans have a similar pattern of sexual dimorphism, they appear to show population differences in the level of differentiation, which could potentially affect the accuracy of sex determination across different populations. The *t*-test results show that European American females and African American females differ in size as do the males, also suggesting variation in levels of sexual dimorphism across populations. These results identified the exact areas and direction of shape changes in the os coxa between males and females that is not possible with more traditional analyses. Knowledge of the areas of variability will allow us to develop sexing criteria using a reduced suite of landmarks, which could have applicability to both the forensic and bioarchaeology setting, particularly for fragmentary remains.

References

1. Byers SN. Introduction to forensic anthropology. Boston, MA: Pearson/Allyn & Bacon, 2008.
2. Krogman WM. The human skeleton in forensic medicine. Springfield, IL: Charles C. Thomas, 1962.
3. Krogman WM, İşcan MY. The human skeleton in forensic medicine, 2nd edn. Springfield, IL: Charles C. Thomas, 1986.
4. Stewart TD. Sex determination of the skeleton by guess and by measurement. Am J Phys Anthropol 1954;12(3):385–92.
5. Rogers T, Saunders S. Accuracy of sex determination using morphological traits of the human pelvis. J Forensic Sci 1994;39(4):1047–56.
6. Steyn M, İşcan MY. Metric sex determination from the pelvis in Modern Greeks. Forensic Sci Int 2008;179:86.e1–e6.
7. Walker PL. Greater sciatic notch morphology: sex, age, and population differences. Am J Phys Anthropol 2005;127:385–91.
8. Greulich WW, Thoms H. The dimensions of the pelvic inlet of 789 white females. Anat Rec 1938;72(1):45–51.
9. Greulich WW, Thoms H. The growth and development of the pelvis of individual girls before, during, and after puberty. Yale J Biol Med 1944;17:91–7.
10. Kurki HK. Protection of obstetric dimensions in a small-bodied human sample. Am J Phys Anthropol 2007;133:1152–65.
11. Tague RG, Lovejoy CO. The obstetric pelvis of A.L. 288-1 (Lucy). J Hum Evol 1986;15:237–55.
12. Leutenegger W, Larson S. Sexual dimorphism in the postcranial skeleton of new world primates. Folia Primatol 1985;44:82–95.
13. Steudel K. Sexual dimorphism and allometry in primate ossa coxae. Am J Phys Anthropol 1981;55:209–15.
14. Leutenegger W. Functional aspects of pelvic morphology in simian primates. J Hum Evol 1974;3:207–22.
15. Leutenegger W. Sexual dimorphism in the pelves of African lorises. Am J Phys Anthropol 1973;38:251–4.
16. Leutenegger W. Newborn size and pelvic dimensions of *Australopithecus*. Nature 1972;240(Dec):568–9.
17. Gingerich PD. The development of sexual dimorphism in the bony pelvis of the squirrel monkey. Anat Rec 1971;172:589–96.
18. Hooton EA. Up from the apes, revised edn. New York, NY: Macmillan Company, 1946.
19. Hrdlicka A. Practical anthropometry, 2nd edn. Philadelphia, PA: Wistar Institute, 1939.
20. Washburn SL. Sex differences in the pubic bone. Am J Phys Anthropol 1948;6:199–207.
21. Straus WL Jr. The human ilium: sex and stock. Am J Phys Anthropol 1927;11:1–28.
22. Phenice TW. A newly developed visual method of sexing the os pubis. Am J Phys Anthropol 1969;30:297–302.
23. Lovell NC. Test of Phenice's technique for determining sex from the os pubis. Am J Phys Anthropol 1989;79(1):117–20.
24. McLaughlin SM, Bruce MF. The accuracy of sex identification in European skeletal remains using the Phenice characters. J Forensic Sci 1990;35(6):1384–92.
25. Sutherland LD, Suchey JM. Use of the ventral arc in sex determination of the os pubis. Proceedings of the 39th Annual Meeting of the American Academy of Forensic Sciences; 1987 Feb 16–21; San Diego, CA. Colorado Springs, CO: American Academy of Forensic Sciences, 1987.
26. McBride DG, Dietz MJ, Vennemeyer MT, Meadors SA, Benfer RA Jr, Furbee NL. Bootstrap methods for sex determination from the os coxae using the ID3 algorithm. J Forensic Sci 1992;46(3):427–31.
27. Bruzek J. A method for visual determination of sex, using the human hip bone. Am J Phys Anthropol 2002;117:157–68.
28. Listi GA, Bassett HE. Test of an alternative method for determining sex from the os coxae: application for modern Americans. J Forensic Sci 2006;51(2):248–52.
29. Kelley MA. Sex determination with fragmented skeletal remains. J Forensic Sci 1979;24(1):154–8.

30. Tague RG. Variation in pelvic size between males and females. Am J Phys Anthropol 1989;80:59–71.
31. Giles E, Elliot O. Sex determination by discriminant function analysis of crania. Am J Phys Anthropol 1963;21(1):53–68.
32. Henke W. On the method of discriminant function analysis for sex determination of the skull. J Forensic Sci 1977;6:95–100.
33. Richtsmeier JT, Cheverud JM, Lele S. Advances in anthropological morphometrics. Annu Rev Anthropol 1992;21:283–305.
34. Pretorius E, Steyn M, Scholtz Y. Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. Am J Phys Anthropol 2006;129:64–70.
35. Ross AH, McKeown AH, Konigsberg LW. Allocation of crania to groups via the ''new morphometry.'' J Forensic Sci 1999;44(3):584–7.
36. Kimmerle EH, Ross A, Slice D. Sexual dimorphism in America: geometric morphometric analysis of the craniofacial region. J Forensic Sci 2008;53(1):54–7.
37. Adam P, Labbe JL, Alberge Y, Austry P, Delcroix P, Ficat RP. Role of computed tomography in the assessment and treatment of acetabular fractures. Clin Radiol 1985;36(1):13–8.
38. Guy RL, Butler-Manuel PA, Holder P, Brueton RN. The role of 3D CT in the assessment of acetabular fractures. Br J Radiol 1992;65(773):384–9.
39. Lele S, Richtsmeier J. An invariant approach to statistical analysis of shapes. Boca Raton, FL: Chapman and Hall/CRC, 2001.
40. Bookstein FL. Morphometric tools for landmark data. Geometry and biology. Cambridge: Cambridge University Press, 1991.
41. Seidler H. Sex-diagnosis of isolated os coxae by discriminant functions. J Hum Evol 1980;9:597–600.
42. Coleman WH. Sex differences in the growth of the human bony pelvis. Am J Phys Anthropol 1969;31:125–52.
43. Bytheway JA. Sex determination of the adult human fragmented pelvis utilizing Euclidean distance matrix analysis. Ph.D. thesis, Pittsburgh, PA: University of Pittsburgh, 2003.
44. Slice DE. Morpheus et al. [computer program]. Recent beta version. Stony Brook, New York: Department of Ecology and Evolution, State University of New York, 1998.
45. Slice DE. Geometric morphometrics. Annu Rev Anthropol 2007;36:261–81.
46. SAS. SAS system [computer program]. Windows version 9.1.3. Cary, NC: SAS, 2004.
47. Waldron T. The relative survival of the human skeleton: implications for paleopathology. In: Boddington A, Garland AN, Janaway RC, editors. Approaches to archaeology and forensic science. Manchester, NH: Manchester University Press, 1987;55–64.
48. Mays S. Taphonomy (sic) factors in a human skeletal assemblage. Circaea 1992;9:54–8.
49. Brain CK. A re-interpretation of the Swartkrans site and its remains. S Afr J Sci 1976;72(5):141–6.
50. Rhine S. In bone voyage. Albuquerque, NM: Albuquerque University of New Mexico Press, 1998.

Additional information and reprint requests:
Joan A. Bytheway, Ph.D.
Sam Houston State University
1003 Bowers Boulevard
Chemistry and Forensic Science Building
Office 221D
Huntsville, TX 77341-2525
E-mail: bytheway@shsu.edu

PAPER

PHYSICAL ANTHROPOLOGY

Guillaume Moskovitch,[1,2] *M.S.; Fabrice Dedouit,*[1,2,3] *M.D.; José Braga,*[2] *Ph.D.;*
Daniel Rougé,[2,3] *M.D., Ph.D.; Hervé Rousseau,*[1] *M.D., Ph.D.; and Norbert Telmon,*[2,3] *M.D., Ph.D.*

Multislice Computed Tomography of the First Rib: A Useful Technique for Bone Age Assessment

ABSTRACT: Macroscopic study of the first rib has been described by Kunos et al. as an efficient method of age estimation. We retrospectively reviewed a test sample of 160 first right ribs obtained from multislice computed tomography (MSCT) clinical investigations of living individuals aged 15–30 years old. Based on the descriptions of Kunos et al., we analyzed the morphological appearance of the sternal end of the first rib on two- and three-dimensional MSCT reconstructions and defined changes in appearance in stages (from 2 to 5). We also studied the calcification and ossification of the costal cartilage as an independent feature. By statistical analysis (ANOVA), we determined for each gender the correlation between the estimated stage of the costal face, the appearance of the costal cartilage, and the documented age of the subjects. We demonstrated that MSCT of the first rib appears to be an efficient noninvasive modality for bone age estimation.

KEYWORDS: forensic science, forensic anthropology, age estimation, first rib, multislice computed tomography, volume rendering technique

The estimation of bone age is an important part of physical and forensic anthropology. It can be carried out in two ways, either by macroscopic or by radiological study. Macroscopic study of bones is commonly used to estimate age at death. Several indicators, such as the pubic symphysis, the fourth right rib, and the auricular face of the ilium, have yielded good results (1–6). Radiological studies are useful for living persons, in particular when age estimation is requested by legal authorities for foreign nationals whose identity is uncertain (7,8). It also avoids the destruction of materials during bone preparation. In view of the increasing technical development of computed tomography (CT), some authors have tested this imaging modality for bone age and sex determination (9,10). CT scan, and especially multislice acquisition, significantly improves spatial definition. Progress in computer technology offers two- and three-dimensional reconstructions of high quality, with no spatial deformation of the feature studied. The combination of these advances allows ''macroscopic'' evaluation thanks to the volume rendering (VR) technique that gives easy soft tissue suppression and good spatial resolution of bone shape, and multiplanar (MP) reconstruction that provides easy and reproducible measurements. A further advantage of digitally stored data is that it may be recalled at will and provide fresh, intact reconstruction. Several studies have confirmed the value of CT scan for the assessment of changes in the medial epiphysis of the clavicle in relation to the civil age of the

subjects examined (11–13). Schmeling et al. had used this criterion earlier in conventional radiography (14). Dedouit et al. examined the reproducibility of measurements obtained on scanned and dry fourth ribs using Iscan's method (15). In our study, we focused on the sternal end of the right first rib and in particular the morphology of its sternal face on CT scan reconstructions in living persons. Kunos et al. previously described the variation of the macroscopic morphology of the first rib as a good indicator of bone age (16). They analyzed the age-related morphological changes of the first rib and of its anatomical features. They made a precise description of the morphological changes of the costal face. These changes concern the geometric shape of the rib face, the surface topography, and the appearance of the face margins, and they are correlated with age. The face of the first rib showed evidence of morphological changes from birth to the fourth decade, with particularly marked changes during the second and third decades. Their descriptive study of the morphological changes of the first rib did not establish statistical correlations between the age intervals and the different stages nor was the gender of the subjects taken into consideration. DiGangi et al. recently reported a new method for bone age-at-death assessment based on the descriptions of Kunos (17). They collected data on three aspects of the first rib (costal face, rib head, and tubercle facet) for 470 known-age male bodies. After statistical analysis of several variables, they reported that the two variables most strongly correlated with age were the geometric shape of the costal face and the surface texture of the tubercle facet.

The aim of our study was age estimation in living subjects based on CT scan of the right first rib, using the earlier descriptions of Kunos et al. On examination of the CT scans, we observed other features (osseous beaks extending from the articular face, calcifications within the costal cartilage) that were significantly related to age. McCormick et al. had previously studied costal cartilage

[1]Service de Radiologie Générale, Hôpital de Rangueil, 1 avenue du Professeur Jean Poulhès, TSA 50032, 31059 Toulouse Cedex 9, France.
[2]Laboratoire d'Anthropobiologie, CNRS, FRE 2960, 37 allées Jules Guesde, Faculté de Médecine, 31073 Toulouse Cedex 3, France.
[3]Service de Médecine Légale, Hôpital de Rangueil, 1 avenue du Professeur Jean Poulhès, TSA 50032, 31059 Toulouse Cedex 9, France.

mineralization on conventional radiography of the plastron and had reported a good correlation between age and costal cartilage modifications on radiography (18). We evaluated the overall accuracy and reproducibility of our method.

Materials and Methods

Sample

We retrospectively reviewed clinical thoracic CT scans of living subjects of documented age, performed in the radiology departments of Toulouse University Hospitals (Toulouse, France), using a picture archiving and communication system (McKesson Medical Imaging Group, Richmond, BC, Canada). The patients had undergone CT scan for various reasons such as thoracic trauma or cardiopulmonary disease. As we wanted to focus on the threshold age of criminal liability and because Kunos et al. reported a better reliability between age and the morphology of the costal face of the first rib during the first four decades, we selected 160 patients from 15 to 30 years old according to the following inclusion criteria:

- no pathological modification of the sternal end of the first rib,
- suitable CT scan parameters: sharp reconstruction filter, slice thickness <2 mm, overlapping reconstruction interval <1.5 mm,
- lastly, for each year of age between 15 and 30, we randomly selected five men and five women.

Multislice Computed Tomography (MSCT)

MSCT was performed on Sensation-16 (Siemens, Erlangen, Germany) and Brilliance-16 systems (Philips Brilliance, Philips Medical Systems, Amsterdam, The Netherlands). The CT reconstruction parameters are summarized in Table 1. The image matrix was 512×512 pixels for Sensation 16 (Siemens) and 1024×1024 pixels for Brilliance 16 (Philips Brilliance).

Postprocessing was performed on a Leonardo workstation (Siemens). Two-dimensional (2D) (MP) and three-dimensional (3D) reconstructions (VR) were obtained for all cases. MP reconstructions were performed along the long axis of the sternal end of the right first rib in two planes (slice thickness 2 mm, reconstruction

interval 1 mm), one anteroposterior and one craniocaudal. VR reconstructions enabled 3D analysis of the sternal end of the rib.

Study Method

Three-Dimensional Analysis—We analyzed the sternal end of the right first ribs. For costal face analysis, the descriptions and drawings of the original article by Kunos et al. were applied to the VR reconstructions (16). This enabled us to stage the sternal end of the first rib, as Iscan did for the right fourth rib (19,20). Three main features based on the description of Kunos et al. were analyzed for staging:

- geometric shape,
- surface topography,
- margin form.

Based on these features, five stages were defined (Table 2). The appearance of stages 2–5 on VR reconstructions is presented in Fig. 1a–d.

Two-Dimensional Analysis—An additional feature was analyzed separately. We observed that the bony changes of the anterior and posterior margins of the costal face extended medially like beaks and were associated with calcifications within the costal cartilage. This feature was termed osseous and calcified projections (OCP).

On cephalocaudal MP reconstructions, we analyzed the beaks or calcifications projecting medially in the costal cartilage (Fig. 2a–d).

These were quantified and scored as absent = 0, one = 1, two = 2.

When projections were present, we observed whether their appearance was linear:

- If it was linear, we measured their length from the medial end of the beak or of the calcification (if present) to the costal face. We compared this length with the distance between the costal face and the sternum. If the length was <50% it was scored 1, if it was >50% it was scored 2 (Fig. 2b,c).
- A nonlinear appearance with an extension at the cranial or caudal faces of the costal cartilage was scored 3 (Fig. 2d).

Analysis and staging of OCP was performed by two observers (one student in anthropology and one forensic pathologist and anthropologist) who were not aware of the sex or age of the subjects.

Statistical Analysis

Statistical analysis was performed with R 2.6.2. (21). We determined for each gender the correlation between the estimated stage

TABLE 1—*Multislice computed tomography (MSCT) reconstruction parameters.*

Slice thickness/interval (mm)	1.0–0.5	1.0–1.0	2.0–1.0	2.0–2.0
Number of MSCT investigations	31	22	61	46

TABLE 2—*Characteristics of first rib stages for three-dimensional analysis based on the descriptions of Kunos et al.*

| Stage 3D	Geometric Shape	Surface Topography	Margin Forms
1	Anteroposterior ovoid	Immature, smooth homogeneous bone	Rounded, undefined
2	Anteroposterior slanted face	Slight undulating ridges and depressions	Superoinferior bone deposition about margin creates rimmed appearance
3	Superoinferior groove separating the face into an anterior two-thirds and a posterior one-third	Transverse ridges bounded by ovoid cavities	Raised margins, defined
4	Anterior thickening of margin projects medially Anteroposterior bone deposition creates superior projection medially	Surface becomes flat Development of pitting	Distinct, rounded, thin profile except for anterior thickening
5	Cartilage encapsulated by shell of cortical bone creating a center concave cavity	Increasing excavation of surface topography	Margins begin projecting. Thin shell transforms into local thickening of margin with anterior margin extended greater than posterior. Scalloped edges
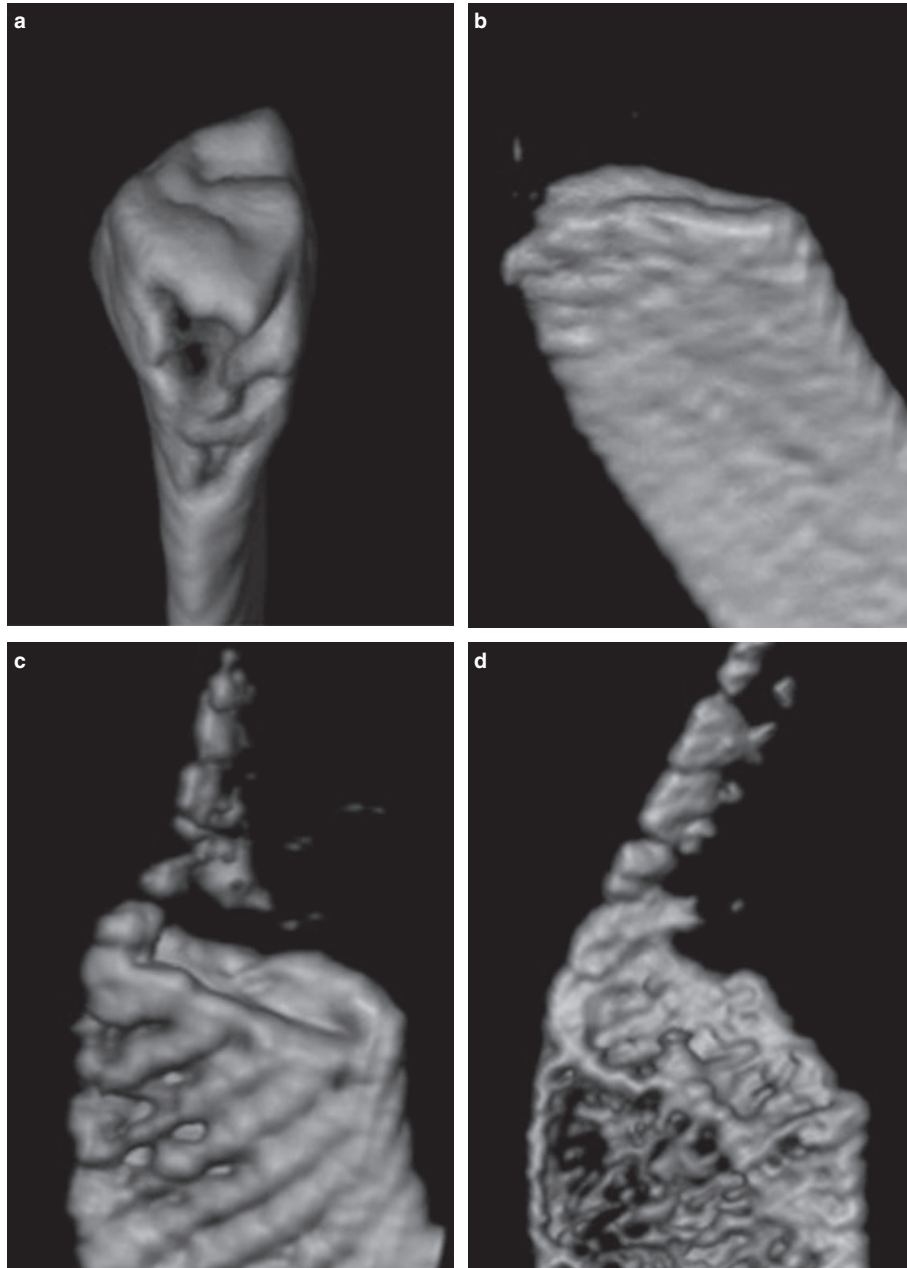
FIG. 1—*(a–d) Stages of the sternal face of the first rib on volume rendering reconstructions (3D analysis). a, Stage 2; b, stage 3; c, stage 4; d, stage 5.*

of the costal face, the appearance of the OCP, and the documented age of the subjects. Mean values, standard deviation, minimum and maximum values, and statistical difference between the distribution of women and men (two-way analysis of variance) were calculated. Cohen's kappa coefficient was used to evaluate intra- and interobserver variability (22).

Results

Stages (Three-Dimensional Analysis)

Table 3 shows the minimum, maximum, mean, and standard deviation for stages 2–5 for each gender. The results are expressed as histograms for each stage, for men (Fig. 3*a*) and women (Fig. 3*b*) separately. As the age range of our population was 15–30 years, no subject had stage 1 characteristics. Male and female

data showed statistically significant differences for all stages ($p = 0.04$), as the age ranges for each stage differed significantly between men and women, and for a given stage, men were also younger than women. Intra- and interobserver reproducibility was good with a kappa coefficient of 0.65 and 0.61, respectively (Table 4).

Osseous or Calcified Projections (Two-Dimensional Analysis)

Results of the number and appearance of OCP are shown as box-plots (Figs. 4*a*,*b* and 5*a*,*b*). There was no statistical difference between men and women for number ($p = 0.06$) or appearance ($p = 0.08$) of OCP. Both features, number and appearance, of the osseous or calcified medial projections showed good reproducibility with kappa coefficients always >0.80 (Table 4).
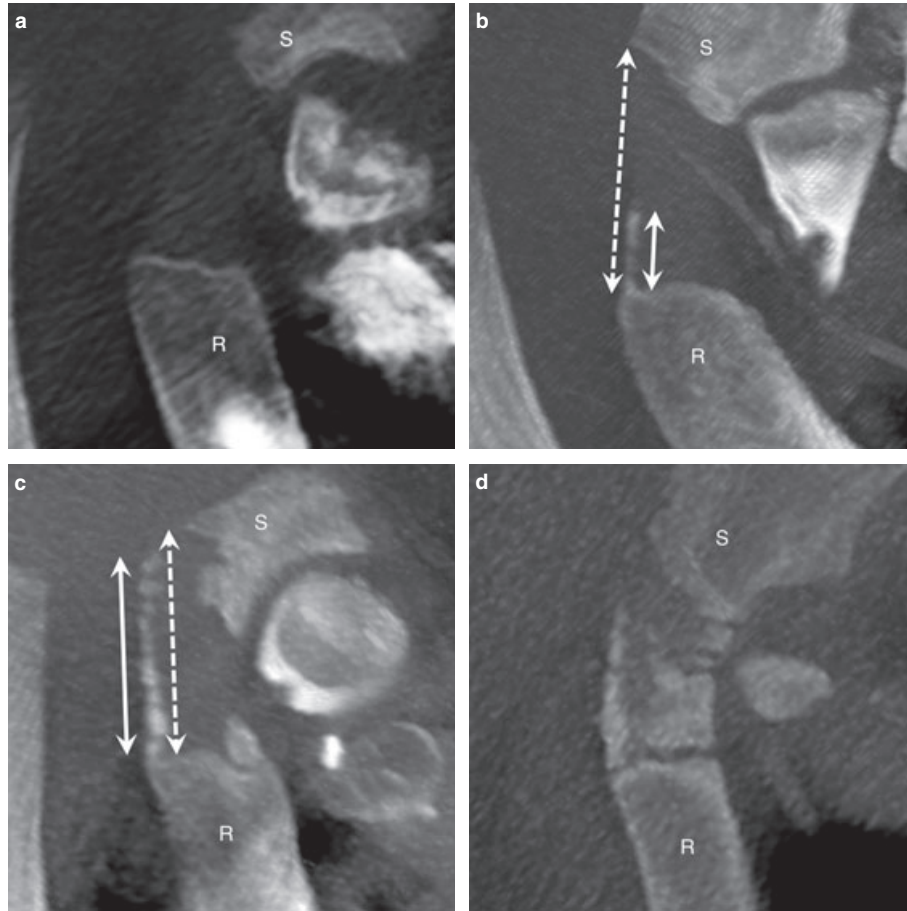
FIG. 2—(a–d) Osseous or calcic projections (OCP): appearance on craniocaudal multiplanar reconstructions (2D analysis). R: first rib, S: sternum. a, Stage 0; b, stage 1 (double arrow: length of the linear OCP, dotted double arrow: distance between the rib surface and the sternum); c, stage 2 (double arrow: length of the linear OCP, dotted double arrow: distance between the rib surface and the sternum); d, stage 3.

TABLE 3—Age distribution (in years): mean values, standard deviation, minimum and maximum values for each stage.

	Women	Men
	Mean (years) ± SD [min, max]	Mean (years) ± SD [min, max]
Stage 3D		
2	16.0 ± 1.4 [15, 17]	15.3 ± 0.6 [15, 16]
3	19.4 ± 3.4 [15, 26]	18.5 ± 2.7 [15, 24]
4	25.5 ± 3.4 [18, 30]	24.2 ± 3.4 [18, 30]
5	28.0 ± 1.4 [27, 29]	27.6 ± 1.9 [24, 30]
ANOVA	$F = 24.7$, Df = 3 ($p \leq 0.01$)	$F = 43.7$, Df = 3 ($p \leq 0.01$)

Discussion

Numerous studies have been conducted on the time frame of ossification of the medial clavicular epiphyseal cartilage using CT scan (11–13). This radiological modality is now recommended in the age group involved in criminal liability procedures (7,8). The authors demonstrated good sensitivity of CT scan for analysis of the ossification and fusion of the medial epiphysis of the clavicle. In addition, recent publications have shown that CT scan is an efficient modality for macroscopic bone study. Dedouit et al. demonstrated the reproducibility of the results obtained applying Iscan's method to MSCT reconstructions and dry bone specimens (15). The MSCT investigation

parameters gave excellent spatial resolution with a low level of irradiation, which is compatible with current clinical practice for living patients.

To the best of our knowledge, age assessment based on VR reconstructions has never been previously evaluated in living subjects.

We chose the sternal end of the first rib for two reasons:

- It develops major changes around 20 years old (criminal liability threshold).
- It is easily examined on thoracic CT scans with no movement artefact.

On VR reconstructions, we found good correlation between documented and the morphological changes of the sternal face of the first rib described by Kunos et al. For identical stages, women were older than men. It was clearly possible to distinguish stage 5 and stage 2 for both genders. For forensic practice, it is interesting to note that

- Stage 2 is reached for both genders at 18 years old.
- The earliest occurrence of stage 4 for both genders is 18 years old.
- The kappa coefficients confirmed the feasibility and reproducibility of this evaluation.

The OCP are well visualized on MP CT scan, in particular isolated calcifications within the costal cartilage which are not fused
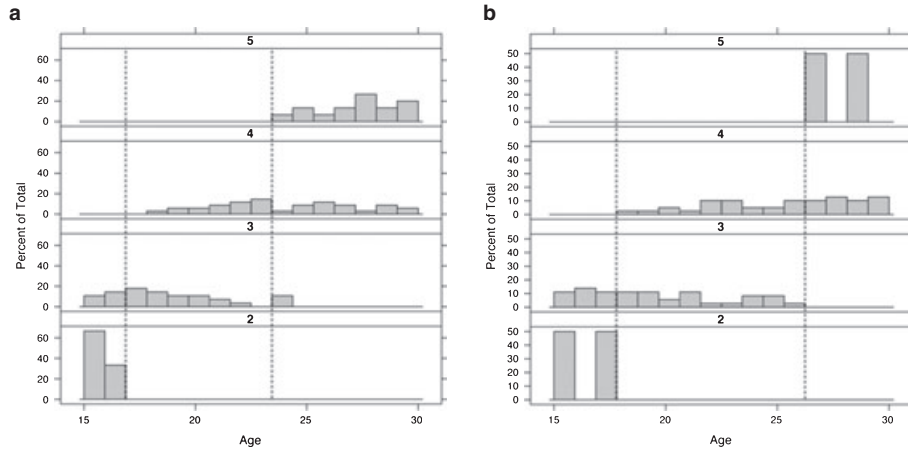
FIG. 3—(a and b) Distribution by age for each stage (dotted lines represent the boundaries between stages 2 and 5). a, Males; b, females.

TABLE 4—Intra- and interobserver variability (Cohen's kappa coefficient).

Criteria Studied		Intraobserver Variability	Interobserver Variability
Stages (3D)		0.65	0.61
Osseous or calcified	Appearance	0.95	0.95
projection(s) (2D)	Number	0.94	0.89

with the sternal end of the rib. CT scan is thus the only way to examine these nonfused calcifications, because they are no longer present on dry bone specimens after soft tissue removal. This analysis gave some significant results:

- In men, absence of OCPs indicates that the subject is <25 years old.
- In women, presence of two OCPs indicates that the subject is more than 20 years old.
- Extensive cartilaginous calcification was only found in individuals more than 20 years old for men and 25 years old for women.

Study of the OCP can be considered as a highly reproducible semi-quantitative criterion.

Conclusion

Our CT adaptation of the descriptions of Kunos et al. for staging the sternal end of the right first rib is well correlated with age and is reproducible. MSCT enabled us to examine calcifications that are not fused to the rib and are no longer present on dry bone specimens, and to confirm their utility in bone age assessment. In forensic practice, MSCT evaluation of the sternal end of the first rib enables estimation of the age of young (<30 years old) living or deceased human beings, as has already been carried out by MSCT investigation of the medial epiphysis of the clavicle. Technical advances in MSCT, computer systems, and data storage have opened up a new field of research in forensic sciences and anthropology that deserves further exploration (23). MSCT evaluation of skeletal sites historically used for age estimation on dry bone (such as the fourth rib or the auricular face of the ilium) may make it possible to determine age precisely with a single scan.
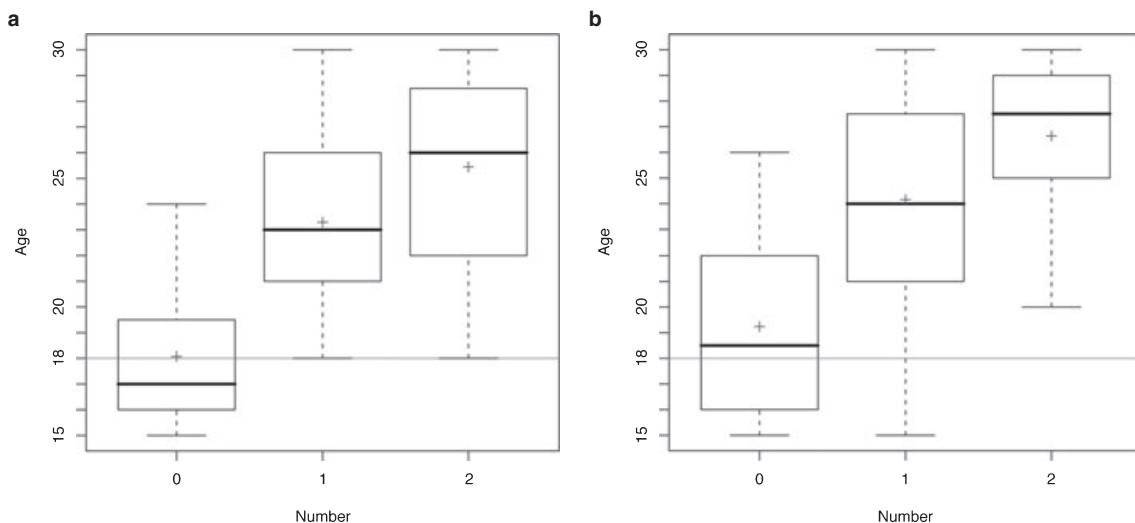

FIG. 4—(a and b) Number of osseous and calcified projections: box represents interquartile range, thick black horizontal line represents median, t-bars represent extremes and (+) represents mean values (gray line: age 18). a, Males; b, females.
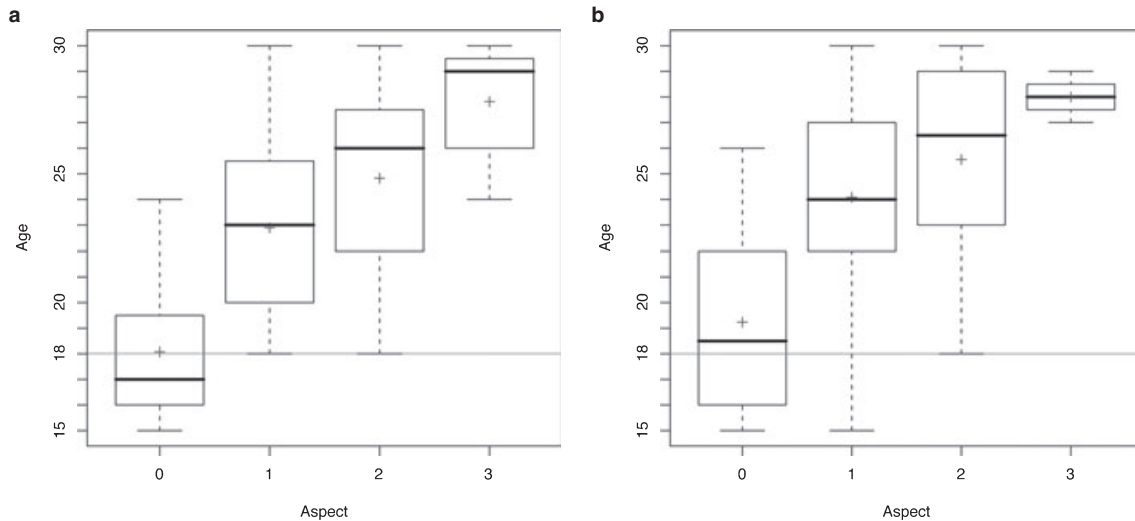
FIG. 5—(a and b) Appearance of osseous and calcified projections: box represents interquartile range, thick black horizontal line represents median, t-bars represent extremes and (+) represents mean values (gray line: age 18). a, Males; b, females.

References

1. Bass WM, Missouri Archaeological Society. Human osteology: a laboratory and field manual, 5th edn. Springfield, MO: Missouri Archaeological Society, 2005.
2. Bedford ME, Russell KF, Lovejoy CO, Meindl RS, Simpson SW, Stuart-Macadam PL. Test of the multifactorial aging method using skeletons with known ages-at-death from the Grant Collection. Am J Phys Anthropol 1993;91:287–97.
3. Martrille L, Ubelaker DH, Cattaneo C, Seguret F, Tremblay M, Baccino E. Comparison of four skeletal methods for the estimation of age at death on white and black adults. J Forensic Sci 2007;52:302–7.
4. Katz D, Suchey JM. Age determination of the male os pubis. Am J Phys Anthropol 1986;69:427–35.
5. Iscan MY, Loth SR, Wright RK. Metamorphosis at the sternal rib end: a new method to estimate age at death in white males. Am J Phys Anthropol 1984;65:147–56.
6. Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. Am J Phys Anthropol 1985;68:15–28.
7. Schmeling A, Olze A, Reisinger W, Geserick G. Age estimation of living people undergoing criminal proceedings. Lancet 2001;358:89–90.
8. Schmeling A, Olze A, Reisinger W, Rosing FW, Geserick G. Forensic age diagnostics of living individuals in criminal proceedings. Homo 2003;54:162–9.
9. Telmon N, Gaston A, Chemla P, Blanc A, Joffre F, Rouge D. Application of the Suchey-Brooks method to three-dimensional imaging of the pubic symphysis. J Forensic Sci 2005;50:507–12.
10. Pasquier E, De Saint Martin Pernot L, Burdin V, Mounayer C, Le Rest C, Colin D, et al. Determination of age at death: assessment of an algorithm of age prediction using numerical three-dimensional CT data from pubic bones. Am J Phys Anthropol 1999;108:261–8.
11. Kreitner KF, Schweden FJ, Riepert T, Nafe B, Thelen M. Bone age determination based on the study of the medial extremity of the clavicle. Eur Radiol 1998;8:1116–22.
12. Schulz R, Muhler M, Mutze S, Schmidt S, Reisinger W, Schmeling A. Studies on the time frame for ossification of the medial epiphysis of the clavicle as revealed by CT scans. Int J Legal Med 2005;119:142–5.
13. Schulze D, Rother U, Fuhrmann A, Richel S, Faulmann G, Heiland M. Correlation of age and ossification of the medial clavicular epiphysis using computed tomography. Forensic Sci Int 2006;158:184–9.
14. Schmeling A, Schulz R, Reisinger W, Muhler M, Wernecke KD, Geserick G. Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. Int J Legal Med 2004;118:5–8.
15. Dedouit F, Bindel S, Gainza D, Blanc A, Joffre F, Rouge D, et al. Application of the Iscan method to two- and three-dimensional imaging of the sternal end of the right fourth rib. J Forensic Sci 2008;53:288–95.
16. Kunos CA, Simpson SW, Russell KF, Hershkovitz I. First rib metamorphosis: its possible utility for human age-at-death estimation. Am J Phys Anthropol 1999;110:303–23.
17. DiGangi EA, Bethard JD, Kimmerle EH, Konigsberg LW. A new method for estimating age-at-death from the first rib. Am J Phys Anthropol 2009;138:164–76.
18. McCormick WF, Stewart JH. Age related changes in the human plastron: a roentgenographic and morphologic study. J Forensic Sci 1988;33:100–20.
19. Iscan MY, Loth SR, Wright RK. Age estimation from the rib by phase analysis: white males. J Forensic Sci 1984;29:1094–104.
20. Iscan MY, Loth SR, Wright RK. Age estimation from the rib by phase analysis: white females. J Forensic Sci 1985;30:853–63.
21. Team RDC. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2008, http://www.r-project.org.
22. Siegel S, Castellan NJ. Nonparametric statistics for the behavioral sciences, 2nd edn. New York, NY: McGraw-Hill, 1988.
23. Dedouit F, Telmon N, Guilbeau-Frugier C, Gainza D, Otal P, Joffre F, et al. Virtual autopsy and forensic identification—practical application: a report of one case. J Forensic Sci 2007;52:960–4.

Additional information and reprint requests:
Guillaume Moskovitch, M.S.
Service de Radiologie Générale
Centre Hospitalier Universitaire Rangueil
1 avenue du Professeur Jean Poulhès
TSA 50032
31059 Toulouse Cedex 9
France
E-mail: moskovitchguillaume@free.fr

PAPER

PHYSICAL ANTHROPOLOGY

Sabrina B. Sholts,[1] *M.A.; Sebastian K.T.S. Wärmländer,*[1,2,3] *Ph.D.; Louise M. Flores,*[1] *B.A.;*
Kevin W.P. Miller,[4] *Ph.D.; and Phillip L. Walker,*[1,†] *Ph.D.*

Variation in the Measurement of Cranial Volume and Surface Area Using 3D Laser Scanning Technology

ABSTRACT: Three-dimensional (3D) laser scanner models of human crania can be used for forensic facial reconstruction, and for obtaining craniometric data useful for estimating age, sex, and population affinity of unidentified human remains. However, the use of computer-generated measurements in a casework setting requires the measurement precision to be known. Here, we assess the repeatability and precision of cranial volume and surface area measurements using 3D laser scanner models created by different operators using different protocols for collecting and processing data. We report intraobserver measurement errors of 0.2% and interobserver errors of 2% of the total area and volume values, suggesting that observer-related errors do not pose major obstacles for sharing, combining, or comparing such measurements. Nevertheless, as no standardized procedure exists for area or volume measurements from 3D models, it is imperative to report the scanning and postscanning protocols employed when such measurements are conducted in a forensic setting.

KEYWORDS: forensic science, forensic anthropology, 3D laser scanning, digital morphometrics, cranial morphology, digital imaging

The problem of accurately measuring the size and shape of the human cranium is a subject of long-standing interest among both forensic and physical anthropologists. Volume and surface area measurements of different craniofacial components have been used to elucidate ontogenetic and phylogenetic relationships in human evolution (1–4), as well as to determine the age, sex, or ethnicity of modern skeletal remains for forensic and medical purposes (5–9).

The volume of the braincase is traditionally measured by filling crania with water, or packing them with granular material such as seed or lead shot (10–16). Similar methods have also been used to measure the capacity of the bony orbits (17–19) and the maxillary sinuses (20). Alternative means of estimating cranial capacity include immersing the calvarium in water (21,22) and creating latex or plaster casts from the endocranial surface (23–26). In addition, mathematical formulae based on linear distance measurements have been used to predict neurocranial, orbital, and maxillary sinus volumes (6,27–33). More recently, advances in computed tomography (CT) scanning, magnetic resonance imaging (MRI), and laser scanning have led to the development of less invasive and more accurate methods of measurement (7,26,34–38).

[1]Department of Anthropology, University of California, Santa Barbara, CA 93106-3210.
[2]Division of Biophysics, Arrhenius Laboratories, Stockholm University, 106 91 Stockholm, Sweden.
[3]UCLA/Getty Conservation Program, A410 Fowler Bldg., University of California, Los Angeles, CA 90095-1510.
[4]Departments of Chemistry and Criminology, California State University, Fresno, 2576 E San Ramon Avenue, M/S ST 104, Fresno, CA 93740.
[†]Posthumous.

Three-dimensional (3D) laser scanners are increasingly being used by forensic scientists, physical anthropologists, and conservators to document, reconstruct, and analyze objects and human remains, including craniofacial features (39–41). This work has served to illustrate the many advantages of laser scanners over other types of 3D imaging technology in terms of cost, speed, and portability. In addition, the spatial resolution of 3D models produced by laser scanners is equal to or greater than many 3D models produced under standard clinical conditions by CT and MRI machines, and the low-energy light used by laser scanners is nonhazardous to both the scanner operator and the biomolecules (e.g., DNA) preserved within the skeletal remains that are scanned (42).

Perhaps one of the most promising aspects of 3D laser scanning technology for forensic applications is the potential to quantify differences in skeletal features that are traditionally limited to visual assessment. Sexual dimorphism in the skull is often evident in the size of the mastoid processes, for example, but standard procedures for sex determination measure this variation on an ordinal scale of 1–5 (43). By analyzing 3D models produced by laser scanners, sexually dimorphic traits can be digitally isolated and measured separately from the rest of the skull, resulting in continuous variables such as volume and surface area (unpublished research). Recent work shows that metric data generally yield more correct classifications of sex estimations than those based on qualitative traits (44), suggesting that a volumetric approach to sex determination may contribute to more accurate biological profiles of unknown human remains.

Although laser scanners appear to be ideally suited for forensic work both in the laboratory and in the field, there is presently a lack of information about the precision and reproducibility of volume and area measurements from the 3D models they produce.

The potential for inter- and intraobserver error is particularly high with 3D laser scanning, as subjective decisions made by the scanner operator can affect the resulting 3D model and its geometric properties, possibly obscuring or distorting meaningful aspects of morphological variation. These decisions include the positioning of the object to be scanned, the number of scans to be recorded, and especially how scans taken from different angles are merged. Such factors are potentially problematic for forensic anthropologists who may present findings based on 3D model measurements in expert witness testimony of their casework (45,46). As Williams and Rogers (47:729) assert, "If the techniques utilized by the forensic anthropologist are not based on reliable principles and methodology, and the accuracy and error rates are not known, the credibility of the expert and professional reputation could be damaged." The same concerns, of course, also apply to scientific research in general. Hence, the use and repeatability of 3D laser scanners must be validated before their digital models can be used to measure and interpret human skeletal morphology.

To investigate this issue, we conducted a study where 3D laser scanner models of five human crania were created independently by two scanner operators who used different protocols for data collection and processing. Volume and surface areas were calculated for all resulting 3D models, allowing us to quantify both inter- and intraobserver errors for the measurements of cranial area and volume, as well as to investigate the levels of geometric variation associated with different instrument settings. The results allowed us to estimate the likelihood that case workers operating under different conditions (e.g., time constraints) will produce 3D models that yield incompatible measurements or are otherwise affected by observer-related differences.

Materials and Methods

Specimens

3D models were created from five human crania excavated from an archaeological site in Malibu, California (CA-LAN-264: accession numbers 505-5, 573-46, 573-47, 573-67B, and 573-73A). These crania are currently housed at the University of California in Santa Barbara and were selected because of their relative completeness and high degree of preservation. For the purposes of this paper, their identification numbers were shortened to 5, 46, 47, 67, and 73.

Data Capture

The NextEngine Desktop 3D scanner (NextEngine Inc., Santa Monica, CA) used to create the 3D models operates with both laser and normal white light. The instrument's laser beams sweep across the surface of the object being digitized, allowing detectors to measure the distance from the light source to the object's surface. The resulting data set, or "point cloud," contains hundreds of thousands of 3D coordinates, or vertices (Fig. 1). Triangles are drawn between these vertices to create a continuous "mesh" surface that approximates the surface contours of the original object. In addition, a built-in digital color camera captures high-resolution color images of the object's surface, and these images are mapped onto the mesh surface, resulting in a 3D full-color rendering of the scanned object. In this study, the scanner mode was set to a geometric point resolution of 75 dots per inch (DPI) while color information was recorded at 150 DPI, which are the standard capture settings for objects of larger size.

In the NextEngine system, objects are positioned on an auto-incremented rotating platform and held in place with an adjustable
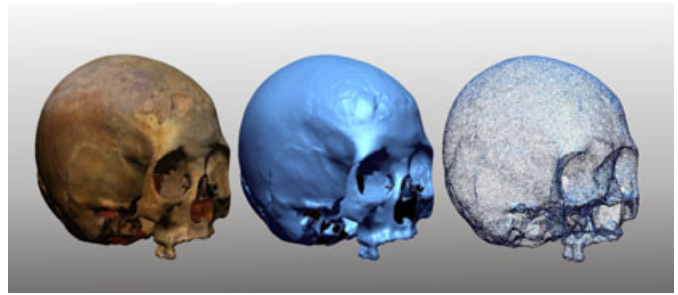

FIG. 1—*Images generated from 3D laser scans of a human cranium. Left: a model in which color and texture information has been mapped onto the mesh surface. Center: the same model in which the color information has been subtracted to create a continuous tone mesh surface. Right: a 3D model generated using the "point cloud" employed to produce the other images.*

part gripper. The scanner operator specifies whether a single scan will be taken from one angle only, or whether consecutive scans will be taken from multiple angles. When a "bracket" scan (three consecutive angles) or "360" scan (scanning during one complete revolution of the platform) is specified, the operator must select the number of scans to be taken and the number of degrees that the platform will rotate between each scan.

As the scanner itself remains stationary during the scanning process, the object must be manually repositioned between data captures to allow all aspects of its surface to be recorded. Thus, the number of scans and the number of re-orientations of the object required to create a complete 3D model depend on the object's size and shape. Ultimately, the operator decides how many scans and angles to record, weighing the time required for each scan against the relative importance of specific regions of the object for the question at hand. To fully capture the complex morphology of the human cranium in 3D can be time-consuming, especially as certain aspects of the cranial surface—notably the posterior aspect of the zygomatic bone and the interior walls of the orbits—have limited visual exposure, so that separate scans need to be made to properly capture these features.

Differences Between the Two Scanning Protocols

Each cranium was scanned three separate times by two independent observers (SBS and LMF). The cranial scanning protocol developed by SBS involved taking 16 scans of two different types. One "360" scan with eight (45°) divisions was performed while the cranium is positioned "face up," with the occiput resting on the platform and the rubber tip of the part gripper placed at *nasion* for stability. The purpose of this scan was to capture the majority of the ectocranial surface cranial vault. To capture the remaining surfaces, eight single scans were taken at various angles. The cranium was then re-positioned in a way that maximizes the visibility to the scanners of each of the following elements: the depth of the orbits, the superior walls of the orbits, the inferior walls of the orbits, the medial wall of the right orbit, the medial wall of the left orbit, the posterior aspect of the right zygomatic, the posterior aspect of the left zygomatic, and the posterior aspect of the alveolar bone. The total procedure took approximately 45 min per cranium to complete.

The cranial scanning protocol developed by LMF involved taking 29 scans of two different types. One "360" scan with 10 (36°) divisions was performed while the cranium was positioned "face up," with the occiput resting on the platform and the rubber tip of

the part gripper placed at *nasion* for stability. As before, the purpose of this scan was to capture the majority of the ectocranial surface of the cranial vault. Additionally, four bracket scans with a division of 12 were taken to capture data from the posterior vault, the cranial base, the inferior walls of the orbits, and the superior walls of the orbits. An extra bracket scan, with a smaller angle of rotation at 14 divisions, was taken to capture the postorbital region, particularly the spheno-frontale juncture. Four single scans were also taken to capture the superior and medial surfaces of the zygomatic arches. For each of these target surfaces, the cranium was re-positioned in a way that maximizes the visibility of the bone surface. The LMF cranial scanning procedure took about 90 min per cranium to complete.

Data Processing

Both observers used the same scanner settings for speed and exposure when the data for the 3D models were captured. Once the data were captured, both observers used the ScanStudio PRO 1.6.3 software (NextEngine Inc. 2006–2008) to align the different scans by manually identifying the common points between them. Aligned scans were then subjected to a process that eliminated the overlap between multiple scans. The built-in function to "intelligently" fill holes in the 3D mesh was not used.

The ScanStudio software allows the scanned data to be reduced, meaning that only a specified fraction of the raw data points is used to render the 3D model. Available reduction levels range from 1 (no reduction) to 10 (maximum reduction). A reduction setting of 1 (no reduction) yields a data file of roughly 230 MB, which was too large to be handled by the 512 MB graphics card used in this study. Therefore, the scan data were reduced to a maximum level of 3, where only one of nine of the raw data points collected from a cranium was used to create a 3D model of mesh triangles with an average size of 0.11 cm (Fig 1).

To investigate how measurements of cranial area and volume were affected by this data point reduction, 3D models were also created using reduction level 6 (1 of 36 raw data points used, mesh triangle size 0.23 cm) and reduction level 9 (1 of 81 raw data points used, mesh triangle size 0.34 cm).

After data processing was complete, the 3D models were imported into the RapidWorks 2.3.2 software (Rapidform, Sunnyvale, CA, 2006), which offers an extensive palette of tools for data analysis, including the functions used to obtain the total volume and surface area values for each 3D model.

Results

Each cranium was scanned three separate times by each observer. The mean surface area (cm^2) and volume (cm^3) values for the five cania are reported in Tables 1 and 2, where they are listed by mesh triangle size (0.11, 0.23, and 0.34 cm) and observer (SBS and LMF).

Surface Area

For the area measurements of 3D models created by SBS with mesh triangle sizes of 0.11, 0.23, and 0.34 cm, the mean standard deviation (SD) values were ±3.1, ±12.4, and ±13.2 cm^2, respectively (Table 1). For 3D models created by LMF, the corresponding SD values were consistently somewhat smaller, i.e., ±2.1, ±10.0, and ±12.9 cm^2, showing that LMF's scanning protocol produced 3D models that yielded slightly more precise area measurements. The differences were, however, too small to be statistically

significant. For both observers, the measurements became more precise when the mesh triangle size was decreased.

When surface area measurements of 3D models of different mesh triangle sizes were compared, the differences between 3D models of 0.11-cm and 0.34-cm mesh sizes were small, i.e., −2.4 cm^2 for SBS and +10.4 cm^2 for LMF. Somewhat unexpectedly, the area differences were larger when the 3D models of 0.11-cm and 0.23-cm mesh sizes were compared, i.e., −15.0 cm^2 for SBS and −14.0 cm^2 for LMF, where the latter value corresponds to roughly 1.5% of the total area. We currently have no explanation for this nonlinear relationship between mesh triangle size and measured surface area.

The differences between surface areas obtained by SBS and by LMF ranged between −19 cm^2 (Specimen 46) and −43 cm^2 (Specimen 73) at the 0.11-cm mesh level, with an average of −30 cm^2. At the 0.23-cm mesh level, the average difference was −31 cm^2, and at the 0.34-cm mesh level, the average difference was −17 cm^2. Hence, for all specimens, the 3D models created by LMF produced larger mean surface area values than those made by SBS. Combining all mesh levels, the total average was −25 cm^2, corresponding to roughly 2.5% of the total area, a number that is small but not insignificant.

Volume

The mean SD values of volume measurements for 3D cranial models created by SBS with mesh triangle sizes of 0.11, 0.23, and 0.34 cm were ±3.8, ±12.9, and ±20.3 cm^3, respectively (Table 2). For 3D models created by LMF, the corresponding mean SD values were ±1.5, ±23.3, and ±29.4 cm^3. Although neither observer produced 3D models that yielded consistently more precise results than the other, it was again clear that a finer mesh size yielded measurements of a higher precision.

When mean volumes of 3D models obtained at different mesh sizes were compared, the volume differences between 3D models of 0.11-cm and 0.34-cm mesh sizes were small (−20.6 cm^3 for SBS and −8.6 cm^3 for LMF). When 3D models of 0.11-cm and 0.23-cm mesh sizes were compared, the differences were even smaller (−3.4 cm^3 for SBS and −4.4 cm^3 for LMF). As 4 cm^3 corresponds to roughly 0.2% of the total average volume, these results indicated that the measured volumes did not vary significantly with mesh size. In other words, the mesh size appeared to affect the precision but not the accuracy of the volume measurements.

At the 0.11-cm mesh level, the differences between the volumetric values obtained by SBS and by LMF ranged between +9 cm^3 (Specimen 47) and +35 cm^3 (Specimen 73), with an average of +26 cm^3. At the 0.23-cm mesh level, the average difference was +25 cm^3, and at the 0.34 cm level, the average difference was +38 cm^3. Combining all mesh levels, the total average was +29.5 cm^3, which corresponds to roughly 1.5% of the total average volume (Table 2). Thus, for all specimens, the 3D models created following SBS's protocol produced larger volumes than those created by LMF's protocol, although the difference is rather small.

Discussion

Observer Error

The 3D model acquisition protocols used by the scanner operators in our study were very different in terms of the type and number of scan captures performed. For each cranium, LMF spent about 90 min collecting 29 scans from different angles, amounting to a total of about 5 million data points, while SBS only spent

TABLE 1—*Mean values and standard deviations for surface area measurements (cm^2).*

Observer	Specimen	Area from 0.11-cm Mesh Size Models	Area from 0.23-cm Mesh Size Models	Area from 0.34-cm Mesh Size Models	Area Difference 0.11–0.23 cm Mesh Size	Area Difference 0.11–0.34 cm Mesh Size
SBS	5	980 ± 2.2	994 ± 7.9	969 ± 7.1	−14	+11
	46	934 ± 2.1	954 ± 19.1	942 ± 31.8	−20	−8
	47	909 ± 1.6	928 ± 17.6	929 ± 18.2	−19	−20
	67	979 ± 3.4	987 ± 8.3	980 ± 2.8	−8	−1
	73	912 ± 5.9	926 ± 9.2	906 ± 6.2	−14	+6
Mean deviation		±3.1	±12.4	±13.2		
Mean difference					−15.0	−2.4
LMF	5	1013 ± 1.4	1020 ± 4.9	1000 ± 9.6	−7	+13
	46	953 ± 4.1	965 ± 10.2	939 ± 3.7	−12	+14
	47	937 ± 1.6	953 ± 9.8	928 ± 4.9	−16	+9
	67	1007 ± 1.1	1029 ± 9.9	1004 ± 7.4	−22	+3
	73	955 ± 2.3	968 ± 15.1	942 ± 39.1	−13	+13
Mean deviation		±2.1	±10.0	±12.9		
Mean difference					−14.0	+10.4
SBS–LMF mean area difference		−30 ± 7.8	−31 ± 11.7	−17 ± 16.3		

*Values are listed by observer, specimen, and mesh size. To find out whether the mesh size affects the accuracy of the area measurements, the differences between areas obtained from models with different mesh sizes were calculated.

TABLE 2—*Mean values and standard deviations for volume measurements (cm^3).*

Observer	Specimen	Volume from 0.11-cm Mesh Size Models	Volume from 0.23-cm Mesh Size Models	Volume from 0.34-cm Mesh Size Models	Volume Difference 0.11–0.23 cm Mesh Size	Volume Difference 0.11–0.34 cm Mesh Size
SBS	5	1988 ± 2.6	1988 ± 4.7	1990 ± 7.8	0	−2
	46	1843 ± 1.2	1823 ± 7.3	1882 ± 59.4	+20	−39
	47	1684 ± 3.5	1700 ± 18.7	1713 ± 18.5	−16	−29
	67	1920 ± 2.4	1930 ± 17.5	1922 ± 6.2	−10	−2
	73	1747 ± 9.2	1758 ± 16.4	1778 ± 9.5	−11	−31
Mean deviation		±3.8	±12.9	±20.3		
Mean difference					−3.4	−20.6
LMF	5	1965 ± 1.3	1963 ± 24.9	1978 ± 32.9	+2	−13
	46	1813 ± 1.1	1831 ± 7.4	1835 ± 10.1	−18	−22
	47	1675 ± 1.7	1669 ± 15.9	1659 ± 26.6	+6	+16
	67	1888 ± 1.6	1878 ± 24.7	1886 ± 37.5	+10	+2
	73	1712 ± 1.6	1734 ± 43.7	1738 ± 39.9	−22	−26
Mean deviation		±1.5	±23.3	±29.4		
Mean difference					−4.4	−8.6
SBS–LMF mean volume difference		26 ± 9.2	25 ± 19.3	38 ± 14.3		

*Values are listed by observer, specimen, and mesh size. To find out whether the mesh size affects the accuracy of the volume measurements, the differences between volumes obtained from models with different mesh sizes were calculated.

45 min collecting 16 scans, with a total of about 2.5 million data points. After the different scans were merged into complete 3D models, and the data size reduced by one-ninth, LMF's and SBS's respective models comprised about 160,000 and 140,000 points each. Thus, the original 100% difference in raw data points was reduced to only a 10% difference in processed points. Visual inspection of the resulting 3D models revealed that the majority of 3D mesh surface gained by LMF's scanning protocol was located in the postorbital region, on the posterior aspect of the zygomatic bones and on the zygomatic arches, where there were small areas of missing data in SBS's models. Several craniometric landmarks such as *sphenion*, *jugale*, and *zygion*, are located on the zygomatic bones and arches of the cranium (48). As these cranial features are often damaged or broken by postmortem processes, especially in ancient remains and forensic remains that have been subjected to perimortem and postmortem trauma, they are generally excluded from many craniometric analyses.

Despite the large difference in collected data points between SBS's and LMF's 3D models, the differences in cranial area and volume measurements were relatively small. The mean difference in volume between 3D models made by SBS and LMF is 29.4 ± 16 cm^3, a quantity that represents about 1.6% of the average model volume, and the mean area difference is −26.3 ± 14 cm^2, representing about 2.6% of the average model surface. Thus, LMF's 3D models had a tendency to be somewhat smaller in volume, but with a larger surface area. A possible explanation of these results is that the additional surface area in LMF's models corresponds to concave features in the cranium that subtract from its volume. In the absence of a standardized method or protocol for measuring external cranial surface area or volume, it is not possible to tell whose measurements—LMF's or SBS's—are the more accurate. The aforementioned presence of small gaps in the surfaces of SBS's models might indicate that the more extensive scanning protocol of LMF yields more reliable data. However, in both cases, the mean differences were just slightly smaller than twice the corresponding standard deviations, meaning that neither the volume nor the area difference between the two observers was statistically significant at a two-sigma level.

When the same cranium was scanned by the same observer multiple times at a mesh size of 0.11 cm, both observers obtained 3D

models that varied by <3 cm^2 (0.3%) for surface area and <4 cm^3 (0.2%) for volume. Also for models of larger mesh size, the SD values for the volume and surface area measurements were rather similar between the two observers. These results suggested that the 3D data captured with a systematic scanning protocol were highly repeatable and was unlikely to be influenced by errors from the scanning process itself. The results also demonstrated that the 3D model approach yields significantly more precise cranial volume measurements than traditional techniques, which can produce standard errors of up to ±50 cm^3 (16,21).

As the NextEngine system requires more operator decisions for creating the 3D models than most other laser scanners, 3D models created with the NextEngine scanner presumably display the largest operator-related errors. If so, the errors reported in this study would constitute upper limits to the precision—but not the accuracy—of area and volume measurements obtained from 3D models created with currently available commercial laser scanners.

Mesh Triangle Size

The results show that mesh size directly impacts the precision of the area and volume measurements, which is not surprising. The standard deviations for repeated volume measurements were about 10 times larger for 3D models with a mesh triangle size of 0.34 cm than for 3D models with a mesh triangle size of 0.11 cm (Table 2). For repeated surface area measurements, 3D models with a mesh triangle size of 0.34 cm mesh yielded standard deviations about five times larger than those for 3D models with a 0.11-cm mesh triangle size (Table 1). Even with a 0.34-cm mesh triangle size, however, the SD values represented an acceptable level of error for most anthropometric applications, (i.e., $<1.5\%$ of the total area and volume). In addition, the differences in actual area and volume values associated with different mesh triangle sizes were $<2\%$ of the total volume and surface areas. As limitations of computer capacity prevented the use of cranial models at the finest mesh size setting available (0.04 cm), the current limiting factor to area and volume measurement precision was not the resolution of the laser scanner, as might have been expected, but rather the resolution of the 3D mesh models. With the continuing advancement of computer and laser scanner technology, the precision of these measurements are expected to improve in the near future. In the meantime, when laser scanner models are used for forensic measurements, it remains important for operators to report not only the protocol used to scan the objects, but also the choices made for postscan data processing and the mesh size used to render the 3D model surfaces.

Conclusions

The results of the present study show that portable, low-cost 3D laser scanners are valuable tools for measuring human remains recovered from forensic and archaeological contexts. Measurements of surface area and volume from human crania were used to illustrate the validity of this technology. Following different protocols for scan data collection, both scanner operators produced 3D models that yielded intraobserver measurement errors amounting to $<0.3\%$ of the total area and volume. When the results obtained from the two protocols were compared, the interobserver errors are likewise small, i.e., on the order of 2%. As measurement precision is correlated with the number of data points used to render the 3D models, which is limited by computer capacity, we expect that improved computer technology together with further automation of the software involved in assembling the 3D models will reduce

inter- and intraobserver measurement errors in the near future. Nonetheless, our present results show that geometric data obtained from 3D models created with the same type of scanner but by different operators and procedures can be shared and combined without uncertainties regarding the comparability of metrical data. This information is particularly important for forensic anthropologists, whose study samples may be more easily analyzed and exchanged in digital form, because of fragility or inaccessibility/legal issues. Moreover, by analyzing human remains via 3D models, forensic anthropologists can construct biological profiles using precise and accurate metrical data to determine key aspects of identity. However, as there is currently no standardized method for measuring cranial area or volume from 3D models, it is imperative that the protocol used is described whenever such measurements are carried out.

References

1. Freedman L, Blumer WFC, Lofgren M. Endocranial capacity of Western Australian Aboriginal crania: comparisons and association with stature and latitude. Am J Phys Anthropol 1991;84(4):399–405.
2. Lorenzo C, Carretero JM, Arsuaga JL, Gracia A, Martínez I. Intrapopulational body size variation and cranial capacity variation in Middle Pleistocene humans: the Sima de los Huesos sample (Sierra de Atapuerca, Spain). Am J Phys Anthropol 1998;106(1):19–33.
3. Steven RL. Cranial capacity evolution in *Homo erectus* and early *Homo sapiens*. Am J Phys Anthropol 1992;87(1):1–13.
4. Brown P, Maeda T. Post-Pleistocene diachronic change in East Asian facial skeletons: the size, shape and volume of the orbits. Anthropol Sci 2004;112:29–40.
5. Ariji Y, Kuroki S, Moriguchi E, Kanda S. Age changes in the volume of the human maxillary sinus: a study using computed tomography. Dentomaxillofac Radiol 1994;23(3):163–8.
6. Fernandes CL. Forensic ethnic identification of crania: the role of the maxillary sinus—a new approach. Am J Forensic Med Pathol 2004;25(4):302–13.
7. Teke HY, Duran S, Canturk N, Canturk G. Determination of gender by measuring the size of the maxillary sinuses in computerized tomography scans. Surg Radiol Anat 2007;29(1):9–13.
8. Duray SM, Martel SS. A quantitative method for estimation of volume changes in arachnoid foveae with age. J Forensic Sci 2006;51(2):238–43.
9. Jones MW. Facial reconstruction using volumetric data. In: Ertl T, Girod B, Greiner G, Niehann H, Seidel HP, editors. Proceedings of Vision, Modeling, and Visualization; 2001 Nov 21–23; Stuttgart, Germany. Amsterdam: IOS Press, 2001;135–42.
10. Bushkovitch VJ. An automatic apparatus for the measurement of cranial capacity. Am J Phys Anthropol 1927;10(3):355–63.
11. Hrdlička A. A modification in measuring cranial capacity. Science 1903;17(443):1011–4.
12. Hrdlička A. Anthropometry. Philadelphia, PA: Wistar Institute, 1920.
13. Matthews W. Use of rubber bags in gauging cranial capacity. Am Anthropol 1898;11(6):171–6.
14. Soemmering ST. Über die körperlichen Verschiedenheiten des Negers vom Europäer. Frankfurt and Mainz: Varrentrapp, 1785.
15. Stewart TD. Cranial capacity studies. Am J Phys Anthropol 1934;18(3):337–47.
16. Todd TW, Kuenzel W. The estimation of cranial capacity: a comparison of the direct water and seed methods. Am J Phys Anthropol 1925;8(3):251–9.
17. P'an TH. Measurement of the Chinese orbit. J Anat 1933;67(4):596–8.
18. Schultz AH. The size of the orbit and of the eye in primates. Am J Phys Anthropol 1940;26(1):389–408.

19. Adachi B. Die Orbita und die Hauptmasse des Schädels der Japaner und die Methode der Orbitalmessung. Z Morphol Anthropol 1904;7:379–480.
20. Shea BT. Eskimo craniofacial morphology, cold stress and the maxillary sinus. Am J Phys Anthropol 1974;47:289–300.
21. Jørgensen JB, Paridon E, Quaade F. The correlation between external cranial volume and brain volume. Am J Phys Anthropol 1961;19(4): 317–20.
22. Jørgensen JB, Quaade F. External cranial volume as an estimate of cranial capacity. Am J Phys Anthropol 1956;14(4):661–4.
23. Diamond MK. Endocasts and meningeal vascular patterns. Am J Phys Anthropol 1994;95(3):355–9.
24. Falk D, Redmond JC Jr, Guyer J, Conroy GC, Recheis W, Weber GW, et al. Early hominid brain evolution: a new look at old endocasts. J Hum Evol 2000;38(5):695–717.
25. Holloway RL. Australopithecine endocast (Taung specimen, 1924): a new volume determination. Science 1970;168(934):966–8.
26. Schoenemann PT, Avants JGB, Holloway RL, Monge J, Lewis J. Validation of plaster endocast morphology through 3D CT image analysis. Am J Phys Anthropol 2007;132(2):183–92.
27. Dekaban A, Lieberman JE. Calculation of cranial capacity from linear dimensions. Anat Rec 1964;150:215–9.
28. Finarelli JA. Estimation of endocranial volume through the use of external skull measures in the Carnivora (mammalia). J Mammal 2006;87(5): 1027–36.
29. Olivier G, Tissier H. Determination of cranial capacity in fossil men. Am J Phys Anthropol 1975;43(3):353–62.
30. Rao CR, Shaw DC. On a formula for the prediction of cranial capacity. Biometrics 1948;4(4):247–53.
31. Todd TW, Russell M. Cranial capacity and linear dimensions, in White and Negro. Am J Phys Anthropol 1923;6(2):97–194.
32. Shea BT. Eskimo craniofacial morphology, cold stress and the maxillary sinus. Am J Phys Anthropol 1977;47:289–300.
33. Gayat J. Essais de mensuration de l'orbite. Ann Ocul 1873;70:1–20.
34. Acer N, Sahin B, Bas O, Ertekin T, Usanmaz M. Comparison of three methods for the estimation of total intracranial volume: stereologic, planimetric, and anthropometric approaches. Ann Plast Surg 2007;58(1):48–53.
35. Mazonakis M, Karampekios S, Damilakis J, Voloudaki A, Gourtsoyiannis N. Stereological estimation of total intracranial volume on CT images. Eur Radiol 2004;14(7):1285–90.
36. Sahin B, Ergur H. Assessment of the optimum section thickness for the estimation of liver volume using magnetic resonance images: a stereological gold standard study. Eur J Radiol 2006;57:96–101.
37. Marsh JL. Measurement of orbital volume by a 3-dimensional software program: an experimental study. J Oral Maxillofac Surg 2000;58:648.
38. Gault D, Brunelle F, Renier D, Marchac D. The calculation of intracranial volume using CT scans. Childs Nerv Syst 1988;4(5):271–3.
39. Karas BV, Beaubien HF. Three-dimensional laser scanning of cultural heritage: the deer stones of Mongolia. Scanning 2006;28(3):187–8.
40. Park H-K, Chung J-W, Kho H-S. Use of hand-held laser scanning in the assessment of craniometry. Forensic Sci Int 2006;160:200–6.
41. Friess M, Marcus LF, Reddy DP, Delson E. The use of 3D laser scanning techniques for the morphometric analysis of human facial shape variation. BAR Int Series 2002;1049:31–5.
42. Grieshaber BM, Osborne DL, Doubleday AF, Kaestle FA. A pilot study into the effects of X-ray and computed tomography exposure on the amplification of DNA from bone. J Archaeol Sci 2008;35(3):681–7.
43. Buikstra JE, Ubelaker DH editors. Standards for data collection from human skeletal remains. Fayetteville, AK: Arkansas Archaeological Survey, 1994.
44. Kimmerle EH, Ross A, Slice D. Sexual dimorphism in America: geometric morphometric analysis of the craniofacial region. J Forensic Sci 2008;53(1):54–7.
45. 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):1–4.
46. Rogers T, Allard TT. Expert testimony and positive identification of human remains through cranial suture patterns. J Forensic Sci 2004;49(2): 203–7.
47. Williams BA, Rogers TL. Evaluating the accuracy and precision of cranial morphological traits for sex determination. J Forensic Sci 2006;51(4):: 729–35.
48. Howells WW. Skull shapes and the map: craniometric analyses in the dispersion of modern Homo. Cambridge, MA: Peabody Museum, Harvard University, 1989.

Additional information and reprint requests:
Sabrina B. Sholts, M.A.
Department of Anthropology
University of California
Santa Barbara
CA 93106-3210
E-mail: sbsholts@umail.ucsb.edu

PAPER

PHYSICAL ANTHROPOLOGY

J. Christopher Dudar,[1] *Ph.D.*

Qualitative and Quantitative Diagnosis of Lethal Cranial Neural Tube Defects from the Fetal and Neonatal Human Skeleton, with a Case Study Involving Taphonomically Altered Remains

ABSTRACT: Cranial neural tube defect, or anencephaly, is the absence of normal brain development because of severe developmental defect in the fetus. While the current incidence of human anencephaly ranges between 1 to 5 per 1000 births, and was higher prior to folic acid supplementation, there is no discussion of anencephaly diagnosis in the forensic literature and only one published example from the archeological record. This article presents both qualitative observations of abnormal cranial elements and an osteometric method to quantitatively determine anencephaly from forensic recovery contexts where taphonomic variables may otherwise mask diagnostic characteristics. Evidence is presented for only the second case of anencephaly diagnosed from a burial context, and the first not involving soft tissue mummification. The initial recognition and accurate prediction of anencephaly is a significant contribution to investigators recovering found human fetal remains.

KEYWORDS: forensic science, anencephaly, neural tube defects, fetal skeleton, neonatal skeleton, fetal abandonment, forensic anthropology

Anencephaly is classified as a lethal neural tube defect (or NTD), which are among the leading causes of perinatal death in the developed world (1). The incidence of NTDs range between 1 to 5 per 1000 births and show marked geographic, ethnic, and temporal variation (2). Anencephaly is defined as the absence of brain development because of a failure of closure in the cranial neuropore sometime between the 23rd and 26th embryonic day. Figure 1*a,b* illustrates the clinical appearance of anencephaly. Some rudimentary lower brain development is usually present with a small proportion of neonates surviving a few days to weeks postpartum; however, the mean gestational period is eight lunar months (3). Appearing frequently with anencephaly is rachischisis, the partial or complete exposure of the vertebral canal (Fig. 2). This combination of defects represents a more severe expression of cranial NTD and is technically referred to as craniorachischisis. This results from an earlier and more complete failure of neural tube closure between the 17th and 23rd embryonic day (4). In this article, anencephaly will be used to generally refer to all cranial NTDs except where noted.

In 2004, the Centers for Disease Control reported that 1640 anencephaly-affected pregnancies occurred annually in the United States during the period immediately prior to mandatory folic acid supplementation of grain products in 1996 (5). The year after supplementation saw a 16% decline in anencephaly-affected

pregnancies to 1380 cases; however, no significant change in the prevalence has been observed in subsequent years (6). These figures include: live births, stillbirths, fetal deaths, and elective abortions, and are acknowledged by the CDC as probable underestimates. The precise etiology of anencephaly, craniorachischisis, and other NTDs such as spina bifida are unknown; however, socioeconomic status, environmental conditions, and the genetics of both population and familial ancestry are indicated (7,8).

While anencephaly is a relatively common birth defect even today, and was likely more frequent historically and prehistorically, no cases of anencephaly appear in the forensic literature. Only one case has been diagnosed from archeologically recovered human remains, an Egyptian mummy from Hermopolis described by the 19th century French zoologists Etienne Geoffroy Saint-Hillare and his son Isidore (9). This specimen has been lost to history but the original illustration is reproduced in the paleopathology literature with various degrees of fidelity (9–12).

With the exception of the Hermopolis mummy, the apparent absence of forensic and archeologically recovered anencephaly cases in the published literature can be attributed to several reasons. Investigator bias is possible as there is a complete deficiency of diagnostic criteria to identify anencephaly from skeletal remains recovered from nonclinical contexts, such as forensic recovery and archeological excavations. At a recent Paleopathology Association annual meeting, Dupras and coworkers presented a case involving a naturally mummified 15-week gestation fetus from the Roman village of Kellis, in Egypt (13). The presence of an abnormally large sagittal opening at the apex of the cranium initially suggested a diagnosis of anencephaly; however, subsequent CT scans of the

[1]Director Repatriation Osteology Lab, Department of Anthropology, National Museum of Natural History, Smithsonian Institution, PO Box 37012, MRC 138 Washington, DC 20013-7012.

FIG. 1—(a) Clinical example of anencephaly, anterior view. Image courtesy of Dr. B. Ragsdale. (b) Clinical example of anencephaly, posterior view. Image courtesy of Dr. B. Ragsdale.

FIG. 2—Range in osteological expression of anencephaly. The case on the right displays a vertebral column with complete rachischisis and vestigial formation of the squamous occipital. The case on the left displays a narrow hiatus in the squamous occipital. Note the differential development of the frontal and parietal bones. Specimens from the National Museum of Health and Medicine.

skull revealed that the thin fetal vault bones had simply curled inwards during the mummification process. In fact Guy and coworkers (14) described how the burial environment and other taphonomic conditions may bias against the preservation and recovery of small, poorly mineralized, fetal and infant bones. Any number of taphonomic variables such as soil environment to maggot and scavenger activity could severely influence the recovery of certain elements as well as affect the visual appearance of both soft and hard tissues from fetal or neonatal remains in forensic contexts. And finally, there may be cultural bias involving stillbirth funeral or disposal custom, as anencephalic neonates could be considered monstrous or even nonhuman. The Hermopolis fetus was mummified in a manner consistent with ancient Egyptian methods for a baboon and thus may not have been perceived as human at the time.

Therefore, the initial recognition of abnormal skeletal morphology and the subsequent accurate and reliable prediction of anencephaly from found fetal and neonatal remains is a significant contribution for workers involved in the forensic recovery of concealed, buried, or otherwise abandoned human remains that may be damaged or incomplete. Toward that goal Miller and Simon

provided rudimentary skeletal illustrations to raise awareness (12), while Dudar presented more refined qualitative morphological indicators for recognition (15), and Dudar and Ousley a quantitative osteometric approach for anencephaly diagnosis (16). Mathews discovered further abnormal postcranial traits on axial skeletal elements (17).

The purpose of this article is to provide both qualitative cranial characteristics for the initial *in situ* recognition of anencephaly and a quantitative method for diagnosis from found human remains, such as in cases of neonatal abandonment where the determination of fetal viability is paramount. The overall appearance of affected cranial elements as well as the physical dimensions and prediction equations are provided. This approach was developed through observations on clinically documented and anatomically prepared anencephalic skeletons from the physical collections at the National Museum of Natural History, Smithsonian Institution. The model was then tested on a taphonomically altered fetal skeleton from a historic cemetery excavation where burial conditions resulted in relatively poor preservation of all human bone recovered. While this test case does not involve forensic significance, it provides an excellent example of the recovery and analysis of taphonomically damaged anencephalic fetal remains, and thus provides a good evaluation of the diagnostic method as presented.

Developmental Osteology of Cranial NTDs

Nomenclature of NTDs is dependent on location and extent of the developmental defect. Total spinal dysraphism is because of a very early and complete failure of fetal neural tube closure resulting in all neural arches being compromised. Failure of closure at the cranial end will result in anencephaly and possibly craniorachischisis if cervical and thoracic neural arches are involved. Failure of only the caudal end will result in spina bifida of the lumbar and/or sacral region.

As the skull is a mosaic of individual but integrated components, there is an essential need for coordination of their growth for normal development. The neurocranium, or those parts of the skull involved in surrounding and protecting the brain, is developmentally divided into the calvarium (or cranial vault) formed from membranous desmocranium, while the basicranium is formed from

TABLE 1—*Clinically documented cases of anencephaly in the Physical Collections at the Smithsonian Institution, NMNH. Comments and spelling from the original catalog cards.*

Catalog Number	Original Clinical Accession Record Comments
218141	Anencephalus & Spina bifida. Male fetus at about 6.5 months; showing anencephalus and Rachischisis. From a white woman, unmarried. Length, vertex-breech: 15 cm
218142	Human fetus, Pseudencephalus. Spina bifida. Female human fetus at about 7 months, showing absence of cranial vault and an open spine its entire length. Length, vertex-breech: 14.6 cm.
218143	Anencephalous. Human fetus (male) at about term.
218146	Skeleton of an anencephalus fetus, Prob. White
219199	Anencephalic fetus, form, with sm spina bifida (superiorly)
	Notes: Facial parts all well developed. No neck—but chin-line + thorax-neck line present. A sack in which remnants of brain projecting backward from a little behind the hair beginning on the frontal a ring of hair all around the cephalic defect. Head in complete retroflexion so that face projects almost directly above. Limbs + H. Buttocks-top 13.5 cm. Fetus of about 7 months (estimate).
219200	Wh. Ananceph. M. fetus. L. abt 28 cm from head to vertex. The case is a cerebral hernia, such lack of parts of frontal & parietals. Fair body + neck + Anencephalous male fetus. white. Fetus apparently at term.

the cartilaginous chondrocranium. Normal formation of the cranial flat bones (frontal, parietal, squamous temporal, and squamous occipital) depends on the presence of an osteogenic membrane derived from embryonic mesoderm and neural crest ectomesenchyme that surrounds the brain. The abnormal or complete absence of higher brain development results from neural tube failure, and thus the osteogenic membrane fails to form resulting in the partial or complete absence of the vault (18). However, it is apparent that even in extreme cases of anencephaly all vault elements are present but are severely reduced in size and morphologically deformed beyond recognition to the uninitiated. The development of the basilar elements of the cranium (sphenoid, petrous temporal, lateral occipital, and basal occipital) proceeds through endocondral ossification, but nonetheless shows marked deformities in cases of anencephaly (19), and in particular the midline structures of these elements.

The osteological manifestation of craniorachischisis, a more severe form of anencephaly that also involves exposure of the vertebral neural canal, occurs because of the failure of the vertebral lamina to fuse at midline resulting in the lateral displacement of the posterior margins and bulging of the underlying neurological tissue. Like the flat bones of the skull, normal lamina initially form at each hemi-arch via intramembranous ossification on the inner surface of the mesenchymal template surrounding the neural tube (20); however, the inductive signal is absent or insufficient to close the mesenchymal template at the dorsal aspect following a NTD. Thus, normal vertebral lamina are not formed, and the spinal cord and canal are left exposed.

Methods and Materials

Seven clinically documented and anatomically prepared anencephalic skeletons from the physical collections at the Smithsonian Institution, National Museum of Natural History, are listed in Table 1. They were observed for morphological abnormalities and their cranial elements measured using the osteometric definitions and illustrative figures provided by Fazekas and Kósa (21) and summarized by Scheuer and Black (20). Two other fetal skeletons, catalog numbers 224884 and 290932, display possible qualitative evidence of anencephaly (17) but are not clinically documented according to the accession records, and so are not included in the reference sample of this investigation. The fetal osteometric data reported by Fazekas and Kósa (21) provide an independent comparative normal sample to generate logistic regression equations to predict anencephaly from recovered remains. The resulting equations were then tested on Burial 492

(B.492), an archeologically recovered fetus that displayed no apparent cranial vault bones on excavation and disinterment (Fig. 3).

B.492 was recovered from the 19th century Elmbank Pioneer Cemetery, formerly on the grounds of Toronto's Pearson International Airport in Canada. The cemetery was excavated under contract by Archaeological Services Inc. (ASI) of Toronto from March through August of 2001, with the author serving as principal physical anthropologist. The cemetery was in use from the 1830s through the 1930s and included 622 burials, of which 14 contained fetal remains. B.492 is an 8.5- to 9.5-lunar-month fetus based on diaphyseal length of leg and arm bones compared to growth and development standards (20,21). Excavation and recovery of B.492 revealed no identifiable cranial elements, and subsequent investigations to sterile subsoil beneath the floor of the small partially preserved coffin indicated that no cranial fragments had migrated because of ground water passing through the coffin. In addition, there was no evidence of rodent burrowing in the excavated feature and preservation of other small bones, such as ribs and metacarpals from B.492, argued against decomposition as an explanation for the absence of recognizable cranial base and vault bones. All of the other fetal burials at Elmbank had preservation of cranial vault flat bones although they were damaged and fragmentary because of the collapse of the coffin and grave shaft overburden.

FIG. 3—*Exposed view of B.492 from Elmbank Pioneer Cemetery, Toronto, Ontario. Note complete absence of observable in situ cranial vault bones.*

Qualitative Osteological Observations

The heterogeneity of NTDs results in a wide range of morphological expression of anencephaly on the vault, or flat bones, of the cranium (Fig. 2). Therefore, any attempt to comprehensively describe the overall appearance and characteristics of these bones for diagnostic purposes will be less than successful. That said, the cranial vault bones generally present a greatly reduced and misshapen appearance with no resemblance to their normal counterpart, except for the orbital regions of the frontal bone (Fig. 2). The squamous portion of the frontal bone typically presents a strongly posterior slope, and the orbital region is significantly more shallow than normal, which likely causes the 'pop-eyed' appearance of clinical or embalmed anencephalic specimens (Fig. 1a). The margins at the defect are usually rounded because of the absence of articulation, however, some sharp margins are also observed along certain sutural segments in some specimens, but there is no apparent pattern.

More diagnostic and quantifiable osteological indicators of anencephaly occur along the midline elements of the cranium, none more so than the sphenoid bone which appears morphologically 'fish-like' and is not easily recognized or oriented even in the clinical anencephaly specimens (Fig. 4a,b). Other anencephalic elements are recognizable but present abnormal shape and size, such as the mandible (Fig. 5a,b), and the lateral and basilar portion of the occipital bones that surround the foramen magnum. The basilar occipital (BO) is observed to be morphologically abnormal as well as smaller in physical dimensions when compared to normal fetal bones of comparable postcranial development (Fig. 6).

Most of the recovered and subsequently identified cranial bones from B.492 displayed abnormal shape despite damage resulting from the burial environment; these included the temporal, lateral and BO, and fragments of the frontal and mandible. Additional abnormal bones from B.492, such as the 'fan-shaped' structures seen to the right in Fig. 7, have no identifiable counterparts in normal fetal anatomy. Comparison of these 'fan-shaped' structures to the skeletons with documented craniorachischisis from the Smithsonian Institution illustrates that these bones are created from fused vertebral lamina and dorsal rib heads as a result of their lateral displacement causing the exposed vertebral canal. Other bones and fragments recovered from the cranial end of B.492 were too damaged to identify.

These qualitative observations prompted the investigation of whether a quantitative model could be created to diagnose anencephaly when burial context and other taphonomic variables have altered the cranial bones. The midline structures of the cranial base are more likely to be protected from taphonomic damage, and thus became the focus of the statistical analysis although the zygomatic bone was also included because it is easily recognized (if difficult to orient) when well preserved in anencephalics.

Quantitative Model to Predict Anencephaly

The parameters of a binary logistic regression model are derived using the maximum-likelihood approach, which allows the direct calculation of the probability of an event, in this case the diagnosis of anencephaly. The dependent variable should have a dichotomous outcome, either normal or anencephalic, and the predicting equation can contain a number of independent variables such as standard fetal and neonatal osteometric measurements (20,21). The statistical package SPSS™ 10.0 (22) was used to analyze the normal and anencephalic craniometric data and to create the prediction model as follows.
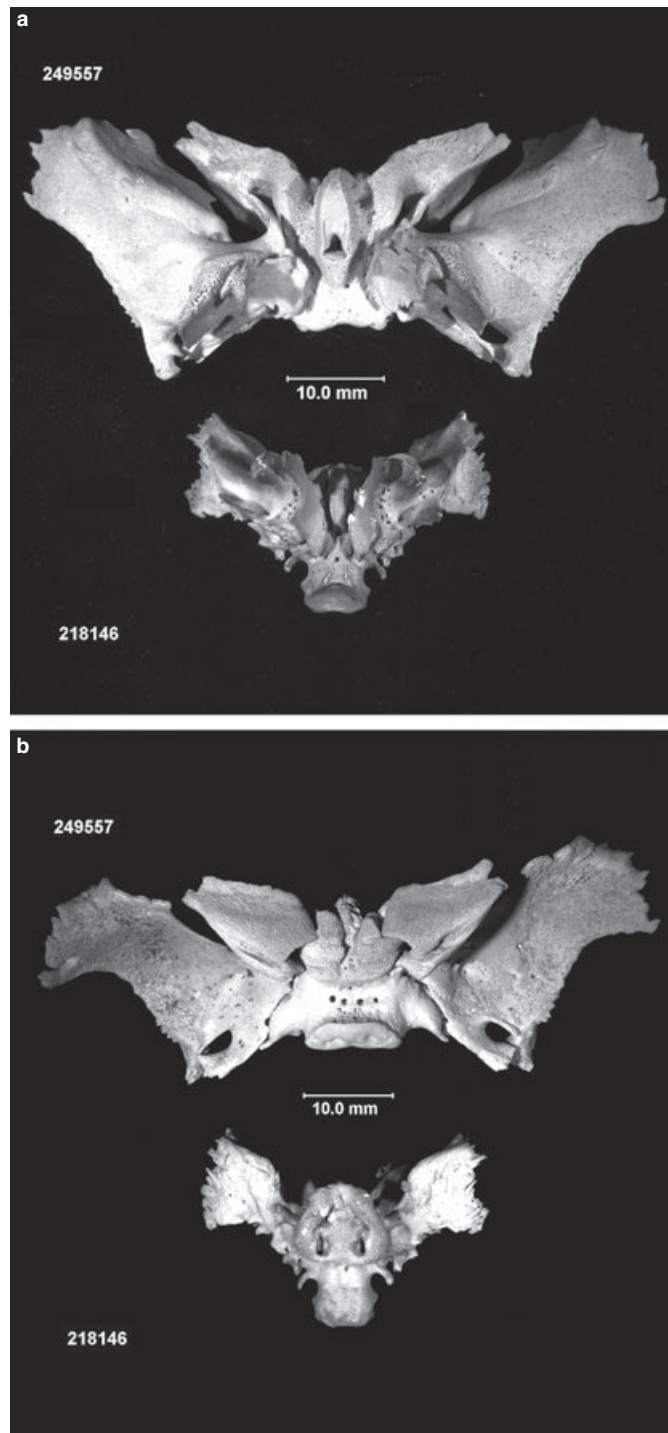

FIG. 4—*(a) Inferior view of normal fetal sphenoid (249557 above) and clinically documented anencephalic (218146 below) of approximately the same postcranial size and development. Specimens from the Smithsonian Institution, National Museum of Natural History. (b) Superior view of normal fetal sphenoid (249557 above) and clinically documented anencephalic (218146 below) of approximately the same postcranial size and development. Specimens from the Smithsonian Institution, National Museum of Natural History.*

Probability of Event (anencephaly) = $1/1 + e^{-z}$
where e = the base of the natural logarithm (2.71828182845904)
$Z = B + B1 (x1) + B2 (x2) + \ldots + Bi (xi)$
x1 to xi = independent variables (osteometric measurements),
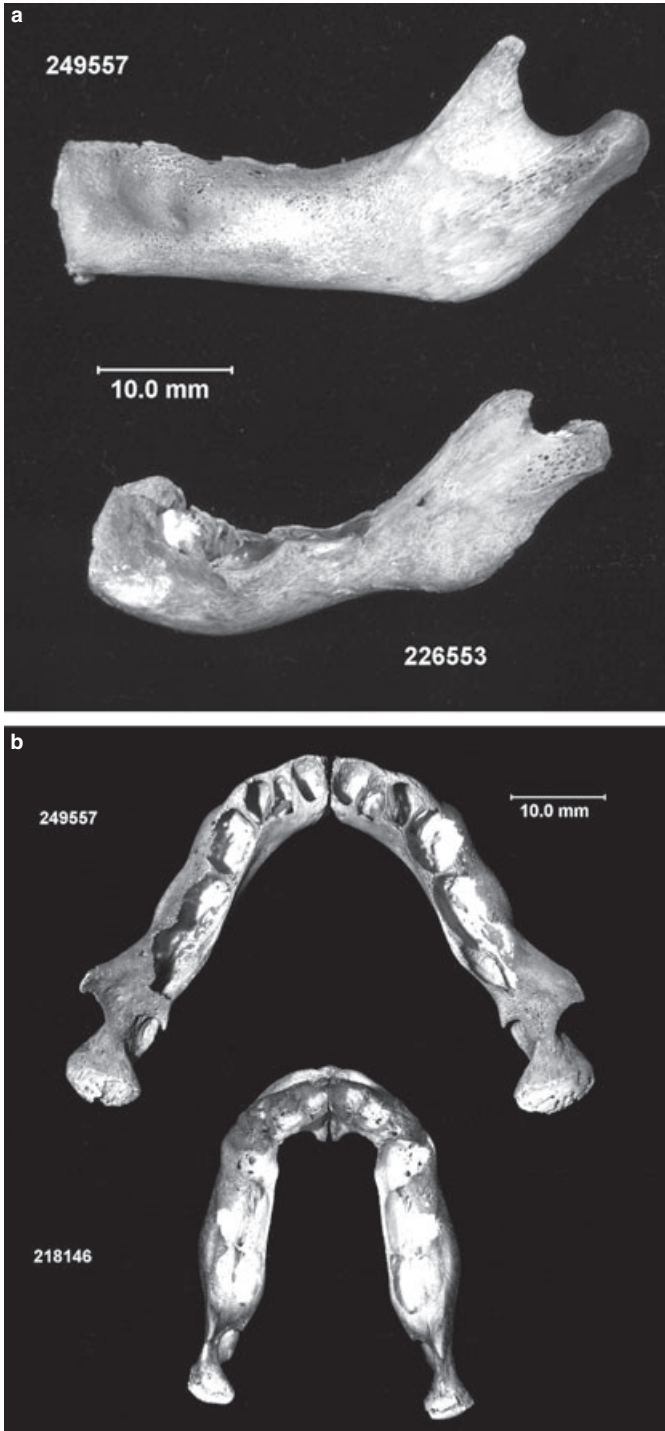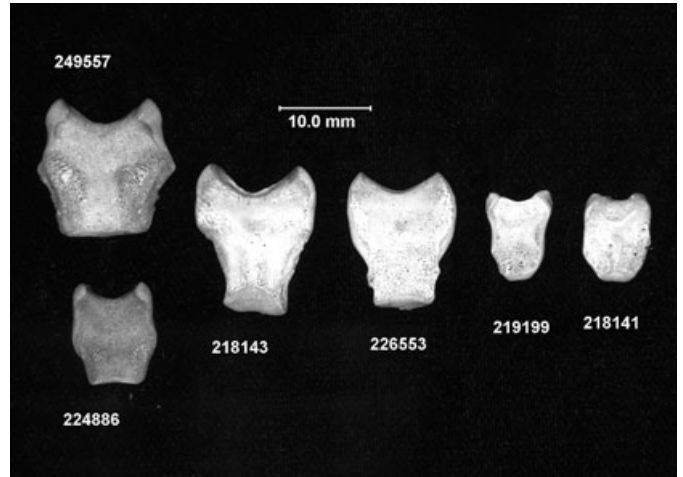B1 to Bi = calculated constants

FIG. 6—*Inferior view of normal fetal basilar occipital bones on the left (catalog number 249557 gestational age 10 lunar months, and 224886 age 7 lunar months). Compare clinically documented anencephalics (218143, 226553, 219199, 218141, gestational ages 10–6.5 lunar months). Note the range of abnormal size and morphological shape. Specimens from the Smithsonian Institution, National Museum of Natural History.*

FIG. 7—*Fusion of the vertebral lamina and ribs in a clinically documented case of craniorachischisis (218142), and similarly appearing osseous structures recovered from B.492 on the right (original scale enlarged and enhanced for contrast).*

FIG. 5—*(a) Lateral view of normal fetal mandible (249557 above) and clinically documented anencephalic (226553 below) of approximately the same postcranial size and development. Specimens from the Smithsonian Institution, National Museum of Natural History. (b) Superior view of normal fetal mandible (249557 above) and clinically documented anencephalic (218146 below) of approximately the same postcranial size and development. Specimens from the Smithsonian Institution, National Museum of Natural History.*

In this way, several equations were created with varying levels of performance; their predictive power depends on the degree of size and shape deformity expressed by the bone as a result of the severity of the NTD itself. Some of the less reliable models are included here to provide the widest possible range of diagnostic tools when damaged and incomplete fetal remains are suspected to be anencephalic. For example, Fig. 8 plots the clinical anencephalic remains and normal fetal data (21) used in the creation of the mandibular model that predicts anencephaly with 100% accuracy. Applying these logistic regression models to B.492 resulted in a maximum-likelihood prediction of anencephaly of $p = 0.98$ to $p = 0.99$ using the BO bone (Fig. 9, and Appendix for a sample calculation).

Mandible (m): (100% correct prediction)

$Z = -52.12 + 9.958$ (m body length) $- 20.747$ (m ramus width)

Basilar Occipital (BO): (62.5% and 97% correct prediction)

$Z = -25.313 - 3.305$ (BO width) $+ 5.227$ (BO length)

$Z = 21.488 - 71.94$ (BO [width/length]) $+ 0.831$ (humerus length)

Lesser Wing Sphenoid (LWS): (100% correct prediction)

$Z = 52.566 - 6.983$ (LWS length)

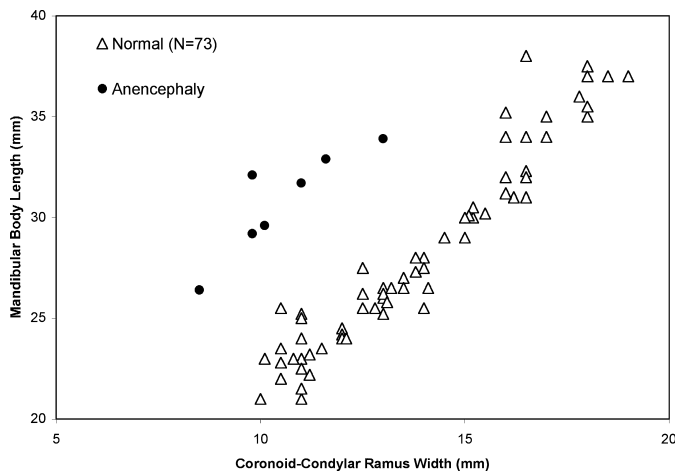$Z = 14.959 - 6.719$ (LWS length) $+ 4.288$ (LWS width)

FIG. 8—*Plot of mandibular variables used in the logistic regression model achieving 100% correct prediction. Normal fetal data as reported by Fazekas and Kosa (21). The mandible of B.492 was not adequately preserved to permit metric assessment.*
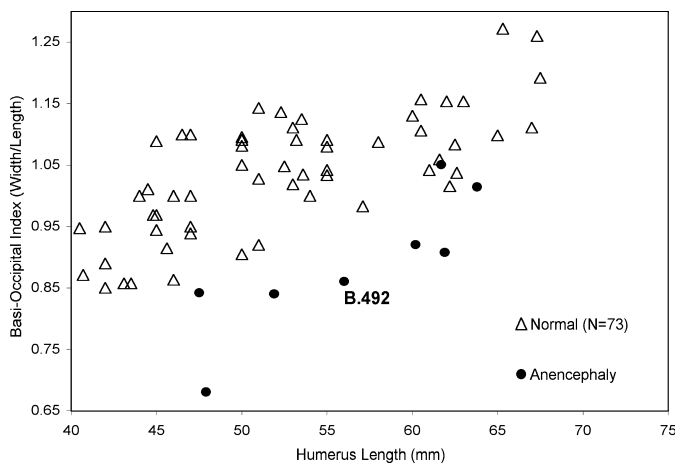

FIG. 9—*Plot of basilar occipital (BO) variables used in the logistic regression model achieving 97.5% correct prediction. The BO of B.492 is predicted to be anencephalic with* p = 0.98. *Normal fetal data as reported by Fazekas and Kosa (21).*

Greater Wing Sphenoid (GWS): (75% and 100% correct prediction)

$Z = 26.613 - 1.455$ (GWS length)

$Z = 86.004 - 11.219$ (GWS length) $+ 9.263$ (GWS width)

Body Sphenoid (BS): (55%, 77%, and 100% correct prediction)

$Z = -11.655 + 0.926$ (BS length)

$Z = 11.284 - 1.069$ (BS width)

$Z = -23.023 + 103.134$ (BS length) $- 72.287$ (BS width)

Maxilla (M): (100% correct prediction)

$Z = 2.899 + 1.603$ (M length) $- 2.0$ (M width)

$Z = -188.688 + 225.116$ (M height) $- 242.228$ (M width)

Zygomatic (Z): (100% correct prediction)

$Z = -10.591 - 14.684$ (Z length) $+ 16.37$ (Z width)

A few of these equations do not include two measurements of the bone involved and therefore the model does not take into account the relationship between the measurements, or proportional shape, of the bone itself. While this limitation may result in initial confusion between the abnormally small anencephalic cranial bones with chronologically younger (and thus smaller) normal fetal skeletal remains, it should be noted that the size of anencephalic postcranial remains was observed to be unaffected by the NTD at the cranial end. The postcranial long bones of the clinically documented anencephalics used in the creation of the diagnostic models were indistinguishable from the long bones of normal fetuses and neonates in the same estimated chronological age range. Possible differences in anencephalic postcranial size were not investigated because the chronological age of the documented anencephalic individuals was itself historically estimated by a pathologist from crown-rump dimensions found in the museum accession records. Regardless, if the skeleton of an individual fetus is carefully observed for gross size and shape inconsistencies between the postcranial and cranial bones, then young and overall small normal fetal skeletal remains cannot be confused with abnormally small cranial bones (but normal postcranial long bones) from anencephalic remains. Even when the postcranial skeleton is damaged and fragmentary it is possible to get an approximate estimate of the developmental size and age from the mid shaft diameter of most long bones. Therefore, it is highly recommended that both quantitative cranial observations and qualitative measurements be used to support the conclusion of anencephaly.

Discussion and Conclusions

Goodman and Gorlin state that the clinical features of anencephaly are so characteristic that it cannot be confused with other birth defects (4). Outside of the clinical setting forensic recovery contexts may involve taphonomic variables such as scavenging animals, decomposition, erosion, or other alteration which may create situations resulting in the absence of the cranial vault, and perhaps even mimic the abnormal morphology of certain osseous elements as seen in anencephaly (14). In addition, field recovery workers may not initially recognize taphonomically altered anencephalic remains because of ignorance of the condition and/or the lack of published diagnostic osteological criteria. This article has discussed and presented qualitative osteological characteristics of anencephaly to facilitate initial recognition, as well as a quantitative osteometric method to predict and diagnose anencephaly using the maximum-likelihood approach. It is highly recommended that both quantitative observations and qualitative measurements be used to support the conclusion of anencephaly.

Based on the implementation of the logistic regression analysis and morphological appearance of the test case, B.492 from Elmbank Cemetery, it is concluded that there is little doubt that these fetal remains are only the second case of anencephaly published in the forensic anthropological or bioarchaeological literature, and the first from a burial context not involving mummification of soft tissue. Further application of the quantitative approach to the suspected cases of anencephaly from the physical collection of the National Museum of Natural History (17), catalog numbers 290932 and 224884 (gestational ages of 3.5–4 lunar months and 4–5 lunar months, respectively), resulted in probabilities of $p < 0.1$ or not diagnostic of anencephaly. This is likely because of damage of these developmentally small fetal bones rendering accurate measurement impossible, as well as the absence of the more strongly diagnostic midline structures from the inventory of these skeletons. This finding strongly cautions us against the sole use of the osteometric approach, and promotes the implementation of both

qualitative and quantitative methods to diagnose anencephaly from incomplete and taphonomically damaged fetal and perinatal human remains.

Acknowledgments

References

1. Kalter H. Folic acid and human malformations: a summary and evaluation. Reprod Toxicol 2000;14:463–76.
2. Birnbacher R, Messerschmidt AM, Pollack AP. Diagnosis and prevention of neural tube defects. Curr Opin Urol 2002;12(6):461–4.
3. Elwood JM. Anencephalus, spina bifida, and potato blight in Canada. Can J Public Health 1976;67:122–6.
4. Goodman RM, Gorlin RJ. The malformed infant and child: an illustrated guide. Oxford: Oxford University Press, 1983.
5. Centers for Disease Control. Spina bifida and anencephaly before and after folic acid mandate—United States, 1995–1996 & 1999–2000. MMWR 2004;53(17):362–5, http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5701a7.htm.
6. Mathews TJ, Honein MA, Erickson JD. Spina bifida and anencephaly prevalence—United States, 1991–2001. MMWR 2002;51(RR13):9–11, http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5113a3.htm.
7. Dickel DN, Doran GH. Severe neural tube defect syndrome from the early archaic of Florida. Am J Phys Anthropol 1989;80:325–34.
8. Sever LE. Looking for causes of neural tube defects: where does the environment fit in? Environ Health Perspect 1995;103(Suppl. 6):165–71.
9. Collins-Cook D. Neglected ancestors: Etienne and Isidore Geoffroy Saint-Hillare. Paleopathol Newsl 2001;116:17–21.
10. Barnes E. Developmental defects of the axial skeleton in paleopathology. Niwot: University Press of Colorado, 1994.
11. Brothwell DR, Powers R. Congenital malformations of the skeleton in earlier man. In: Brothwell DR, editor. The skeletal biology of earlier human populations. Oxford: Pergamon Press, 1968;173–204.
12. Miller E, Simon SK. Anencephaly—something missing from the archaeological record? Paleopathol Newsl 2001;115:9–11.
13. Dupras TL, Breach T, Tocheri M, Molto JE, Sheldrick P. A naturally mummified fetus from Roman period Egypt. Proceedings of the Paleopathology Association's Annual Meeting; 2003 April 22–23; Tempe (AZ). Paleopathology Association, 2003;4.
14. Guy H, Masset C, Baud C-A. Infant taphonomy. Int J Osteoarchaeology 1997;7:221–9.
15. Dudar JC. A probable case of anencephaly with rachischisis in pioneer Upper Canada: qualitative and quantitative observations. Proceedings of the Paleopathology Association's Annual Meeting; 2003 April 22–23; Tempe (AZ). Paleopathology Association, 2003;3–4.
16. Dudar JC, Ousley SD. Diagnosis of anencephaly, a common lethal neural tube defect, from taphonomically altered fetal or neonatal skeletal remains. Proceedings of the 57th Annual Meeting of the American Academy of Forensic Sciences; 2005 Feb 21–26; New Orleans (LA). Colorado Springs, CO: American Academy of Forensic Sciences, 2005;24.
17. Mathews SL. Diagnosing anencephaly in archaeology: a comparative analysis of nine clinical specimens from the Smithsonian Institution National Museum of Natural History, and one from the Archaeological Site of Kellis 2 Cemetery in Dakhleh Oasis, Egypt [Masters Thesis]. Orlando (FL): University of Central Florida, 2008.
18. Sperber GH. Craniofacial embryogenesis: normal developmental mechanisms. In: Mooney MP, Siegel MI, editors. Understanding craniofacial anomalies: the etiopathogenesis of craniosynostoses and facial clefting. New York, NY: Wiley-Liss, 2002;31–60.
19. Ortner DJ. Identification of pathological conditions in human skeletal remains. Smithsonian Contributions to Anthropology Number 28. Washington, DC: Smithsonian Institution Press, 2003.
20. Scheuer L, Black S. Developmental juvenile osteology. San Diego, CA: Academic Press, 2000.
21. Fazekas IG, Kósa F. Forensic fetal osteology. Budapest: Akademiai Kiadó, 1978.
22. SPSS 10.0 [computer program] MS-Windows XP version. Chicago (IL): SPSS Inc., 2002.

Additional information and reprint requests:
J. Christopher Dudar, Ph.D.
Director Repatriation Osteology Lab
Department of Anthropology
National Museum of Natural History
Smithsonian Institution
PO Box 37012
MRC 138 Washington, DC 20013-7012
E-mail: dudarc@si.edu

Appendix

Sample calculation using the basilar occipital (BO) predictive logistic regression equation and the osteometric measurements from B.492, with measurement standards as defined by Fazekas and Kósa (21).

Step 1: The predictive equation, variables, and constants

$$Z = 21.488 - 71.94 \ (\text{BO [width/length]}) + 0.831 \ (\text{humerus length})$$
$$Z = 21.488 - 71.94 \ (10.5/12.2) + 0.831 \ (56)$$
$$Z = 3.7539$$

Step 2: Calculating probability with the logistic regression model

Probability of anencephaly $= 1/1 + e^{-z}$

where e = the base of the natural logarithm (2.71828182845904)

In Microsoft Excel^{TM}, the equation is entered into the cell as

$$=1/[1 + \text{EXP}(-Z)]$$

so, $p = 1/(1 + e^{(-3.7539)})$

or, $p = 1 \ (1 + 2.71828182845904^{(-3.7539)})$

$p = 0.977$ which is strongly diagnostic of anencephaly

PAPER

PHYSICAL ANTHROPOLOGY

Irene Atef Fawzy,[1] *M.D. and Nashwa Nabil Kamal,*[2] *M.D.*

Stature and Body Weight Estimation from Various Footprint Measurements Among Egyptian Population

ABSTRACT: Analysis of footprints can reveal very important clues which can be used as a forensic evidence and help in the estimation of stature and body weight of an individual. In this work, bilateral footprints were obtained from 50 male Egyptian medical students ranging in age between 18 and 25. Nine measurements were taken on each footprint. The result revealed significant bilateral asymmetry ($p < 0.001$) except foot breadth at ball. The significant and positive highest correlation coefficients with stature were shown by toe-5 length on right side ($R = 0.58$) and with body weight by foot breadth at ball on left side ($R = -0.52$). Regression equations presented smaller standard errors of estimate (3.52–4.69) in determination of stature than those in estimation of body weight (4.05–5.28). In conclusion, this study has provided equations that help to estimate stature and body weight from footprint measurements among Egyptians.

KEYWORDS: forensic science, forensic anthropology, stature estimation, body weight estimation, footprint measurements, Egyptians

The human foot has been studied for a variety of reasons, i.e., for forensic as well as for nonforensic purposes by anatomists, forensic scientists, anthropologists, physicians, podiatrists, and numerous other groups (1).

Footprints are of immense value in establishing personal identity of individuals in forensic examinations. They are found as a kind of evidence at the crime site and a potential link between the crime and the perpetrator (2).

Although footprints can be collected from almost every kind of crime scene, the possibility of their recovery at the scenes of sexual offenses and homicides is relatively greater. Examination of barefoot impressions is important in developing countries where the majority of the rural population like to walk barefooted because of socioeconomic and climatic reasons. The partial or complete footprints can be found on rain-covered surfaces, newly waxed floors, freshly cemented surfaces, moistened surfaces, in dust, mud, sand, oil, paint, and blood (3).

Analysis of footprints helps in estimation of an individual's stature because of the existence of a strong positive correlation between one's stature and foot size (4). Similarly, a relationship has been suggested between one's body weight and footprints (5).

So, the aim of this study was to verify the possibility of estimation of stature and body weight from various dimensions of footprints among Egyptians and to derive regression formula for stature and body weight estimation from these dimensions.

Materials and Methods

This study was carried out in the Forensic Medicine and Clinical Toxicology Department, Faculty of Medicine, El-Minia University in collaboration with the Public Health Department.

The study included 50 Egyptian men. They were medical students of El-Minia University. Their ages ranged from 18–25. All subjects were free from any apparent foot deformity. An informed consent was taken from all participants, and the procedures followed were in accordance with the ethical standards of El-Minia University committee on human experimentation. Footprints were obtained from the right and left feet of all subjects according to the steps of obtaining the footprints by Robbins (6).

Each subject was asked to clean his soles, and the ink used for obtaining footprints in Egypt was applied to the cleaned soles of the subjects; then, each subject was asked to put his right sole followed by his left sole on a white paper. Before lifting his sole from the paper, anatomical land marks of the feet were marked on the papers which are calcaneal tubercle lateral, calcaneal concavity medial, mid-rear heel point, lateral metatarsal point, and medial metatarsal point. The longitudinal axis and the base line were drawn on the footprint. The longitudinal axis is from the mid-rear heel point to the lateral side of toe 1. The base line is drawn at the rear edge of the foot perpendicular to the longitudinal axis.

Nine measurements (Fig. 1) were all taken in centimeters on the footprints (right and left):

Five measurements were taken from the mid-rear heel point to the most anterior point of each toe (T_1, T_2, T_3, T_4, and T_5).

Foot breadth at ball was measured from metatarsal lateral, the most lateral point on the metatarso-phalangeal joint of toe number 5, and metatarsal medial, the most medial point on the metatarso-phalangeal joint of toe number 1.

Big toe pad breadth measurement from lateral metatarsal point to medial metatarsal point.

[1]Department of Forensic Medicine and Toxicology, Faculty of Medicine, El-Minia University, Minia, Egypt.

[2]Department of Public Health, Faculty of Medicine, El-Minia University, Minia, Egypt.

Big toe pad length from anterior terminal landmark to posterior terminal landmark.

Foot breadth at heel measured from calceneal concavity medial to calcaneal tubercle lateral.

Measuring stature and weight were carried out according to Lohman et al. (7). The stature was measured in centimeters by asking the subject to stand barefoot on the base board of a standard metric height measuring stand, both feet in close contact with each other, trunk straight along the vertical board with eyes looking forward. The measurement was taken by bringing the projecting horizontal sliding bar to the vertex. The weight was measured in kilograms by using a standard balanced scale with the subject standing on the center of the scale barefoot.

Statistical Analysis

The data were collected and analyzed using Statistical Program for Social Sciences (SPSS) version 8 (IBM Company, Chicago, IL). The difference between left and right footprint measurements was determined to create a new variable. Next, the mean and standard deviation of this new difference variable were calculated and the mean compared to "0" by way of a one-sample *t*-test. Linear regression analysis was performed to calculate regression equations following Jasuja et al. (8) and Krishan (9). Regression equations were also computed for estimating stature and body weight. The standard error of estimate (SEE) was calculated to predict the deviation of the estimated stature and body weight from the actual stature and body weight. Karl Pearson's correlation coefficient was used to determine the association between various footprint measurements and stature or body weight.

Results

The descriptive statistics of right and left footprint measurements are presented in Table 1. A range of variation can be observed in the measurements of variables.

Table 2 lists the bilateral differences (left−right) in footprint measurements by one-sample *t*-test in all cases. All variables except foot breadth at ball show statistically significant difference (asymmetry), and values are greater on the left side. The differences between left and right feet are small and of limited forensic significance.

Tables 3 and 4 show the linear regression equations for estimation of stature and body weight from footprint measurements on right and left sides. The tables also present the *R*-square, *F* value, *p*
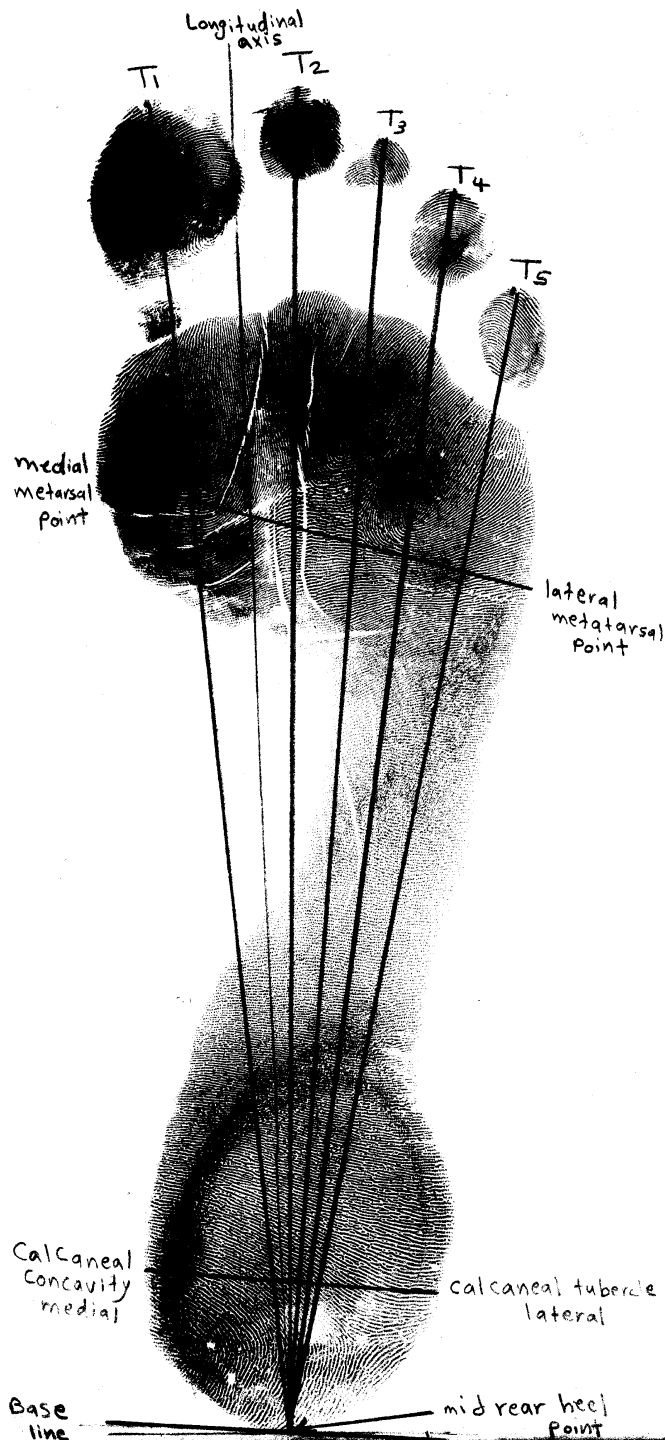

FIG. 1—*Showing footprint measurements.*

TABLE 1—*Descriptive statistics of footprint measurements in centimeters (n = 50).*

Variables	Min. No. 50	Max. No. 50	Mean No. 50	SD No. 50
T_1 Right	22.10	27.30	24.82	1.26
T_1 Left	22.70	27.60	25.31	1.21
T_2 Right	21.30	27.40	24.69	1.29
T_2 Left	22.50	27.80	25.16	1.24
T_3 Right	20.60	26.20	23.80	1.26
T_3 Left	21.50	26.80	24.27	1.19
T_4 Right	19.70	25.30	22.67	1.23
T_4 Left	20.50	25.70	23.14	1.17
T_5 Right	18.10	23.10	20.94	1.14
T_5 Left	18.90	23.80	21.49	1.07
BTB Right	1.90	3.20	2.67	0.30
BTB Left	2.00	3.30	2.78	0.28
BTL Right	2.70	4.90	3.62	0.56
BTL Left	2.60	4.90	3.72	0.49
FBB Right	7.00	11.20	9.00	0.77
FBB Left	7.00	11.00	9.09	0.71
FBH Right	3.40	6.70	4.80	0.67
FBH Left	3.70	6.40	5.02	0.65

Min., minimum; Max., maximum; SD, standard deviation; T_1, length from anterior part of 1st toe to mid-rear heel point; T_2, length from anterior part of 2nd toe to mid-rear heel point; T_3, length from anterior part of 3rd toe to mid-rear heel point; T_4, length from anterior part of 4th toe to mid-rear heel point; T_5, length from anterior part of 5th toe to mid-rear heel point; BTB, big toe breadth; BTL, big toe length; FBB, foot breadth at ball; FBH, foot breadth at heel.

value, and the SEE that predicts the deviation of estimated stature and body weight from the actual stature and body weight.

Equation for stature or body weight estimation $= a + bx$, where a is the regression coefficient of the dependent variable (stature or body weight), and b is the regression coefficient of the independent variable (measurement of footprint), and x is the measurement of the footprint. Regression equations have been computed for stature and body weight estimation for each side and for each variable.

The regression equation is a better indicator of stature estimation from toe 5 length on right side (SEE = 3.52) and body weight determination from foot breadth at ball on left side (SEE = 4.05). This result showed also that the regression equation is a better indicator for stature estimation (SEE = 3.52–4.69) than for body weight determination (SEE = 4.05–5.28).

TABLE 2—*One-sample* t-*test of bilateral differences (left−right) in footprint measurements (n = 50).*

Variables	Mean Difference (Left−Right) No. 50	Standard Difference No. 50	t Value No. 50	p Value No. 50
T_1	0.48	0.46	7.262	0.000
T_2	0.47	0.50	6.658	0.000
T_3	0.47	0.41	8.185	0.000
T_4	0.46	0.41	7.991	0.000
T_5	0.55	0.43	9.019	0.000
BTB	0.11	0.21	3.709	0.001
BTL	0.10	0.20	3.634	0.001
FBB	0.009	0.43	1.465	0.149
FBH	0.22	0.32	4.838	0.000

T_1, length from anterior part of 1st toe to mid-rear heel point; T_2, length from anterior part of 2nd toe to mid-rear heel point; T_3, length from anterior part of 3rd toe to mid-rear heel point; T_4, length from anterior part of 4th toe to mid-rear heel point; T_5, length from anterior part of 5th toe to mid-rear heel point; BTB, big toe breadth; BTL, big toe length; FBB, foot breadth at ball; FBH, foot breadth at heel.

$p < 0.05$ is significant.

TABLE 3—*Linear regression equations for stature estimation (in centimeters) from different footprint measurements on right and left sides.*

Variables	Regression Equations (No. 50)	SEE (No. 50)	R-Square	F	p
T_1 Right	$91.88 + 3.1 \times T_1R$	3.55	0.293	19.489	0.000
T_1 Left	$88.34 + 3.25 \times T_1L$	3.63	0.290	19.178	0.000
T_2 Right	$113.38 + 2.32 \times T_2R$	3.92	0.162	9.104	0.004
T_2 Left	$106.98 + 2.53 \times T_2L$	4.03	0.186	10.706	0.002
T_3 Right	$112.02 + 2.46 \times T_3R$	3.86	0.173	9.863	0.003
T_3 Left	$110.65 + 2.47 \times T_3L$	4.00	0.159	8.897	0.005
T_4 Right	$110.01 + 2.67 \times T_4R$	3.86	0.197	11.540	0.001
T_4 Left	$111.86 + 2.54 \times T_4L$	3.96	0.168	9.499	0.003
T_5 Right	$92.57 + 3.72 \times T_5R$	3.52	0.332	23.362	0.000
T_5 Left	$93.89 + 3.57 \times T_5L$	3.67	0.271	17.439	0.000
BTB Right	$172.71 + 0.76 \times BTBR$	4.69	0.002	0.084	0.773
BTB Left	$154.84 + 5.68 \times BTBL$	4.27	0.045	2.212	0.144
BTL Right	$162.93 + 2.13 \times BTLR$	4.48	0.018	0.882	0.352
BTL Left	$158.59 + 3.23 \times BTLL$	4.17	0.047	2.331	0.134
FBB Right	$144.63 + 2.89 \times FBBR$	4.14	0.094	487.5	0.032
FBB Left	$146.54 + 2.65 \times FBBL$	4.16	0.064	3.223	0.079
FBH Right	$162.83 + 1.62 \times FBHR$	4.48	0.018	0.876	0.354
FBH Left	$61.19 + 1.80 \times FBHL$	4.39	0.025	1.185	0.282

SEE, standard error of estimate; T_1, length from anterior part of 1st toe to mid-rear heel point; T_2, length from anterior part of 2nd toe to mid-rear heel point; T_3, length from anterior part of 3rd toe to mid-rear heel point; T_4, length from anterior part of 4th toe to mid-rear heel point; T_5, length from anterior part of 5th toe to mid-rear heel point; BTB, big toe breadth; BTL, big toe length; FBB, foot breadth at ball; FBH, foot breadth at heel.

$p < 0.05$ is significant.

The coefficient of determination (R_2) or the predictive value was best for stature from toe 5 length on right side ($R_2 = 0.332$) and for body weight from foot breadth at ball on left side ($R_2 = 0.266$). Tables 5 and 6 illustrate the correlation coefficients between stature or body weight and footprint measurements on both sides.

All length measurements exhibit statistically positive significant correlation coefficients with stature and body weight except the big toe length on right and left sides. As regards breadth measurements,

TABLE 4—*Linear regression equations for body weight estimation (in kilograms) from different footprint measurements on right and left sides.*

Variables	Regression Equations (No. 50)	SEE (No. 50)	R-Square	F	p
T_1 Right	$31.88 + 1.28 \times T_1R$	4.32	0.172	9.746	0.003
T_1 Left	$33.53 + 4.20 \times T_1L$	5.04	0.159	8.907	0.004
T_2 Right	$30.32 + 3.99 \times T_2R$	4.09	0.221	13.299	0.001
T_2 Left	$43.30 + 4.62 \times T_2L$	4.28	0.194	11.297	0.002
T_3 Right	$36.95 + 4.28 \times T_3R$	4.29	0.185	10.682	0.002
T_3 Left	$32.91 + 4.35 \times T_3L$	4.25	0.197	11.514	0.001
T_4 Right	$41.42 + 4.62 \times T_4R$	4.11	0.219	13.152	0.001
T_4 Left	$36.99 + 4.75 \times T_4L$	4.19	0.200	11.769	0.001
T_5 Right	$35.29 + 4.58 \times T_5R$	4.16	0.211	12.540	0.001
T_5 Left	$34.60 + 5.03 \times T_5L$	4.17	0.209	12.386	0.001
BTB Right	$42.43 + 5.26 \times BTBR$	5.11	0.108	5.703	0.021
BTB Left	$32.89 + 14.16 \times BTBL$	5.05	0.109	5.33	0.021
BTL Right	$28.87 + 15.05 \times BTLR$	5.28	0.010	0.458	0.502
BTL Left	$59.80 + 3.03 \times BTLL$	5.22	0.019	0.916	0.343
FBB Right	$57.77 + 3.49 \times FBBR$	4.08	0.234	14.342	0.000
FBB Left	$11.53 + 9.05 \times FBBL$	4.05	0.266	17.055	0.000
FBH Right	$5.12 + 5.34 \times FBHR$	5.14	0.067	3.387	0.072
FBH Left	$30.54 + 8.01 \times FBHL$	4.63	0.163	9.183	0.004

SEE, standard error of estimate; T_1, length from anterior part of 1st toe to mid-rear heel point; T_2, length from anterior part of 2nd toe to mid-rear heel point; T_3, length from anterior part of 3rd toe to mid-rear heel point; T_4, length from anterior part of 4th toe to mid-rear heel point; T_5, length from anterior part of 5th toe to mid-rear heel point; BTB, big toe breadth; BTL, big toe length; FBB, foot breadth at ball; FBH, foot breadth at heel.

$p < 0.05$ is significant.

TABLE 5—*Correlation coefficients between stature and footprint measurements.*

Variables	Value of r
T_1 Right	0.54**
T_1 Left	0.54**
T_2 Right	0.40**
T_2 Left	0.43**
T_3 Right	0.42**
T_3 Left	0.40**
T_4 Right	0.44**
T_4 Left	0.41**
T_5 Right	0.58**
T_5 Left	0.52**
BTB Right	0.04
BTB Left	0.21
BTL Right	0.14
BTL Left	0.22
FBB Right	0.31*
FBB Left	0.25
FBH Right	0.14
FBH Left	0.16

r, correlation; T_1, length from anterior part of 1st toe to mid-rear heel point; T_2, length from anterior part of 2nd toe to mid-rear heel point; T_3, length from anterior part of 3rd toe to mid-rear heel point; T_4, length from anterior part of 4th toe to mid-rear heel point; T_5, length from anterior part of 5th toe to mid-rear heel point; FBB: foot breadth at ball; BTB, big toe breadth; BTL, big toe length; FBH, foot breadth at heel.

*p Value significant; **p value highly significant.

TABLE 6—*Correlation coefficients between body weight and footprint measurements.*

Variables	Value of r
T_1 Right	0.41**
T_1 Left	0.40**
T_2 Right	0.47**
T_2 Left	0.44**
T_3 Right	0.43**
T_3 Left	0.44**
T_4 Right	0.47**
T_4 Left	0.45**
T_5 Right	0.46**
T_5 Left	0.46**
BTB Right	0.33*
BTB Left	0.33*
BTL Right	0.10
BTL Left	0.14
FBB Right	0.49**
FBB Left	0.52**
FBH Right	0.26*
FBH Left	0.40**

r, correlation; T_1, length from anterior part of 1st toe to mid-rear heel point; T_2, length from anterior part of 2nd toe to mid-rear heel point; T_3, length from anterior part of 3rd toe to mid-rear heel point; T_4, length from anterior part of 4th toe to mid-rear heel point; T_5, length from anterior part of 5th toe to mid-rear heel point; BTB, big toe breadth; BTL, big toe length; FBB, foot breadth at ball; FBH, foot breadth at heel.

*p Value significant; **p value highly significant.

they showed statistically positive significant correlation coefficients with stature except foot breadth at ball on left side, foot breadth at heel, and big toe breadth on both sides, while all breadth measurements show statistically significant correlation coefficients with body weight.

Discussion

Footprints have a considerable value in forensic science, and they can be collected from the crime scene and can be utilized as a kind of evidence for estimation of body size, i.e., stature and body weight (9).

This study provides specific formulae or regression equations which can help in prediction of Egyptians' stature and body weight in men from footprint measurements. The range of age (18–25) was chosen because the average adult length of foot is attained by the age of 16 years in men (10). Also, Roche (11) stated that generally stature at 18 years is accepted as an adult, although there are small increments in stature after this age. The median age for attaining height in men is 21.2 years with growth continuing in 10% of men until 23.5 years (12). Also, this range of age was chosen because it is the range of age of medical students in El-Minia University.

Footprint measurements of Egyptians can be compared with the studies by Philip (13) on south Indian subjects, Jasuja et al. (8) on Jat Sikhs of Punjab, and Krishan (14) on Gujjars of north India. Descriptive information emerged showing smaller footprint dimensions when compared to those subjects.

The results of this work revealed significant bilateral asymmetry in footprint measurements ($p < 0.001$) except foot breadth at ball with greater values on left side. The bilateral asymmetry can be explained on the basis of the fact that majority of persons put greater strain on the left lower limb than on the right in walking and in weight bearing (15). This is supported by a study on limb bilateral asymmetry in which the author found a significant left-sided asymmetry in the lower limb of the Gujjar population, (16) while

Krishan (14) found that only T_2 length and T_5 length showed statistically significant asymmetry among male Gujjars of north India. On the other hand, this result contrasts with that of Robbins (6) who did not find significant bilateral asymmetry in various measurements of the feet of the U.S. population. Similar views are expressed by Philip (13) who did not find any significant asymmetry while working on the footprints of a south Indian population. Jasuja et al. (8) successfully tried to estimate stature from left and right side measurements of the foot but their study did not find any significant asymmetry in the measurements of the left and right feet.

The regression equations obtained from footprint measurements were better indicators of stature because of lower SEE (3.52–4.69) than for body weight determination because of slightly higher SEE (4.05–5.28). Krishan and Sharma (17) revealed that the low SEE (about 2–6 cm) indicated the high reliability and accuracy in estimating stature by regression analysis in a north Indian population. According to Sen and Ghosh (18), the prediction of stature is more accurate using regressions among various Indian populations. In contrast to this result, Zeybek et al. (19) found that stature estimation formula allow 9- to 10-cm errors in subjects who are attending the Medical Faculty of Dokuz Eylul University and School of Physical Therapy and Rehabilitation in Turkey.

In this study, the estimation of stature from T_{1-5} length measurements in a footprint has more reliability of prediction (SEE = 3.52–4.03) than from other measurements like breadth at ball/heel and big toe breadth/length (SEE = 4.14–4.69). These findings are consistent with that of Jasuja et al. (8) who revealed that the mean error was minimum when stature was estimated from foot length measurements and maximum when estimated from foot breadth measurements in Jat Sikhs. Also, Krishan (14) detected that the mean errors were minimum (2.12–2.35) when stature was estimated from foot length measurements, and the estimates are close to the actual stature.

The coefficient of determination (R_2) or the predictive value was best for stature from toe 5 length on right side ($R_2 = 0.332$). Sen and Ghosh (18) in similar research found that R_2 was the best fit for foot length (both right and left) for male Rajbanshis in linear regression analysis ($R_2 = 0.392$), while Kanchan et al. (20) revealed that R_2 was best for left foot length ($R_2 = 0.583$).

In this investigation, the higher positive and significant correlation coefficients of T_{1-5} length measurements with stature ($r = 0.40$–0.58) suggest a moderately strong relationship between these measurements and stature. This is in accordance with Sen and Ghosh (18), who detected that stature and foot length are positively and significantly correlated with each other ($r = 0.50$–0.62). Similar results ($r = 0.75$–0.76) were obtained by Kanchan et al. (20) on Gujjars, a north Indian endogamous group. Krishan (15) observed that the correlation of stature with various toe length measurements is extremely high ($r = 0.82$–0.87). Similar results ($r = 0.80$–0.85) were obtained by Robbins (6), Ozden et al. (21) on a Turkish sample ($r = 0.46$–0.59), and Singh and Phookan (22) on four different ethnic groups of India.

Many studies have been conducted on estimation of stature from foot dimensions. Unfortunately, a few studies have reported the relationship of body weight with footprints. However, some authors have worked on the relationships of anthropometric dimensions of the foot and body weight (23–25). El-Meligy et al. (26) revealed that the regression equation formulae are helpful in the estimation of weight of the deceased from tibial length and bimalleolar breadth. In this work, the regression equation for determination of the body weight from footprint measurements gave slightly higher SEE (4.05–5.28) in agreement with Krishman (9) who detected higher error in estimating weight (3.05–4.10).

All breadth and length measurements in this work showed statistically positive significant correlation with body weight ($r = 0.26–0.52$) except big toe length on right and left sides. The significant correlation between body weight and these measurements is not as high as the correlation of these measurements with stature. This is consistent with Krishman (9) who found positive and significant correlation coefficient between body weight and measurements of footprints ($r = 0.38–0.75$), although their values are not as high as the correlation of those measurements with stature.

In conclusion, body size, i.e., stature and body weight is an important parameter in determining the identity of unidentified bodies and dismembered remains. Footprint is one of the means to establish the body size of the person. Its significance lies in the simplicity in measurement, applicability, and accuracy in prediction. The regression equation formulas which are derived from this study are helpful in estimating stature and body weight among the Egyptian population from various measurements of footprints with defined SEE. And this is of utmost importance in forensic examination to establish individuality of an unidentified body. So this study would be helpful in the interpretation and analysis of footprints in criminal cases.

An important point to remember is that the people from different regions of the world have different morphological features depending on their geographical distribution and racial characteristics, so a single formula cannot represent all parts of the world. It is therefore suggested that similar studies should be conducted in different parts of the world so that the effect of genetic and environment can be investigated in forensic terms.

References

1. Krishan K. Individualizing characteristics of footprints in Gujjars of North India—forensic aspects. Forensic Sci Int 2007;169(2–3):137–44.
2. Sharma BR. Forensic science in criminal investigation. Allahabad, India: Centra Law Agency, 1990.
3. Qamra SR, Sharma BR, Kaila P. Naked foot marks—a preliminary study of identification factors. Forensic Sci Int 1980;16:145–52.
4. Kennedy RB, Chen S, Pressman IS, Yamashita AB, Pressman AE. A large scale statistical analysis of bare foot impressions. J Forensic Sci 2005;50:1071–80.
5. Robbins LM. Estimating height and weight from size of footprints. J Forensic Sci 1986;31:143–52.
6. Robbins LM. Footprints collection, analysis and interpretation. Springfield, IL: Charles C. Thomas, 1985.
7. Lohman T, Roche A, Martorell R. Anthropometric standardization reference manual. Washington, DC: Library of Congress, 1988; 3–26.
8. Jasuja OP, Singh J, Jain M. Estimation of stature from foot and shoe measurements by multiplication factors: a revised attempt. Forensic Sci Int 1991;50:203–15.
9. Krishan K. Establishing correlation of footprints with body weight—forensic aspects. Forensic Sci Int 2008;179(1):63–9.
10. Blais MM, Green WT, Anderson M. Lengths of the growing foot. J Bone Joint Surg 1956;8:998–1000.
11. Roche AF. Bone growth and maturation. In: Falkner F, Tanner JM, editors. Human growth. New York and London: Plenum Press, 1986;25–60.
12. Roch AF, Davila GH. Late adolescent growth in stature. Pediatrics 1972;50:874–80.
13. Philip TA. Formula for estimating stature from foot size by regression method. J Ind Acad Forensic Med 1990;12:57–62.
14. Krishan K. Estimation of stature from footprint and foot outline dimensions in Gujjars of North India. Forensic Sci Int 2008;175(2–3):93–101.
15. Singh I. Functional asymmetry in the lower limbs. Acta Anat 1970;77:131–8.
16. Krishan K. Limb bilateral asymmetry and footprints of male adult Gujjar population in parts of Punjab and Haryana (dissertation). Chandigarh, India: University of Punjab, 2002.
17. Krishan K, Sharma A. Estimation of stature from dimensions of hands and feet in north Indian population. J Forensic Leg Med 2007;14(6):327–32.
18. Sen J, Ghosh S. Estimation of stature from foot breadth among the Rajbanshi. Forensic Sci Int 2008;181(55):55.e1–e6.
19. Zeybek G, Ergur I, Demiroglu Z. Stature and gender estimation using foot measurements. Forensic Sci Int 2008;181(54):54.e1–e5.
20. Kanchan T, Menezes RG, Moudil R, Kaur R, Kotian MS, Grag RK. Stature estimation from foot dimensions. Forensic Sci Int 2008;179(2–3):241–5.
21. Ozden H, Balci Y, Demiruslu A, Turgut A, Griugrul M. Stature and sex estimate using foot and shoe dimensions. Forensic Sci Int 2005;147:181–4.
22. Singh TS, Phookan MN. Stature and foot size in four Thai communities of Assam, India. Anthropol Anz 1993;51:349–55.
23. Gorden CC, Buikstra JE, Walker RA. Forensic identification of height and weight using applied anthropometry of the foot. Am J Phys Anthropol 1989;78:230.
24. Cheng JC, Leung SS, Leung AK, Guo X, Sher A, Mak AF. Change of foot size with weight bearing—a study of 2829 children 3–18 years of age. Clin Orthop Relat Res 1997;342:123–31.
25. Anil A, Peker T, Turgut HB, Ulukent SC. An examination of the relationship between foot length, foot breadth, ball girth, height and weight of Turkish university students aged between 17 to 25. Anthropol Anz 1997;55:79–87.
26. El-Meligy MM, Abdel Hady RH, Abdel Maaboud RM, Mohamed ZT. Estimation of human body built in Egyptians. Forensic Sci Int 2006;159(1):27–31.

Additional information and reprint requests:
Irene Atef Fawzy, M.D.
Faculty of Medicine
Forensic Medicine and Toxicology Department
El-Minia University
Minia
Egypt
E-mail: ireneatef@yahoo.com

PAPER

PHYSICAL ANTHROPOLOGY; PATHOLOGY AND BIOLOGY

Tal Simmons,[1] *Ph.D.; Peter A. Cross,*[1] *M.Sc.; Rachel E. Adlam,*[1] *M.Sc.; and Colin Moffatt,*[1] *Ph.D.*

The Influence of Insects on Decomposition Rate in Buried and Surface Remains

ABSTRACT: This article reports results of a comparative study of decomposition rates of wild rabbit (*Oryctolagus cuniculus*) which were either (i) buried after exposure to insect activity, (ii) buried without exposure, (iii) kept above ground behind an insect screen, or (iv) continuously exposed above ground in a field experiment. Results showed that dipteran oviposition occurred consistently in groups i and iv only. Decomposition rates (measured by Total Body Score every *c.* 50 accumulated degree days [ADD]) of rabbits kept behind the screen and those buried without exposure showed no difference ($p = 0.450$). This was significantly slower than those buried after exposure ($p = 0.0016$) which was in turn significantly slower than those continuously exposed ($p \ll 0.001$). Temperatures collected from animals showed the presence of feeding larvae increased intra-abdominal temperatures to $>5°C$ above ambient. The findings support the assertion that insect presence is the primary agent affecting decomposition rate via tissue consumption and also the heat they generate.

KEYWORDS: forensic science, taphonomy, insects, decomposition, degree-days, terrestrial, burial

Although long utilized by entomologists and introduced to human decomposition research by Vass et al. in 1992 (1), forensic anthropologists have been slow to accept the standardization of time/temperature (accumulated degree days [ADD]) as the *x*-axis event timeline for decomposition; even some of the most current publications indicate that it has not been incorporated into research designs or data analysis (2–4) and reluctance to accept the implications of results generated by its use appears to be common. An analogy to an understanding of ADD is quite simple: for example, a pot of water on the stove at low heat may take 15 min to boil, whereas the same volume of water in the same pot will boil in 5 min if the heat is turned up. The same endpoint is reached, the pot boils—at the same accumulation of heat energy; the amount of chronological time it takes the pot to boil is incidental, as long as the right amount of heat energy accumulates.

A slightly more complex analogy is to basic chemistry, where the input of thermal energy increases the speed of a chemical reaction. Hence in decomposition studies, ADD records the input of thermal energy (temperature multiplied by time) and measures the reaction (decay, measured in total body score [TBS]) to that input. Thus, the number of days it takes a corpse to skeletonize is incidental to the amount of thermal energy required to reach skeletonization. To continue the analogy, the experiment discussed in this article explores the use of a catalyst, to further increase the rate of decomposition in relation to ADD. In this experiment, the catalyst is insect access and, more specifically, the action of maggots.

Recently published data (5) indicate that, when the accumulation of temperature over time is standardized by using ADD, insects influence the rate of decomposition more than any other variable. Simmons and colleagues (5) showed that there was no significant difference in the decomposition rate of carcasses submerged in water, buried, or left indoors. This held true regardless of carcass species, carcass size and depositional environment. In those cases where insects were excluded, decomposition progressed at a much slower rate. When this insect-excluded group was contrasted to remains deposited terrestrially where insects could access the remains, the latter group decomposed significantly faster. Carcass size has been demonstrated to be of secondary importance (5) to decomposition rate; smaller carcasses decompose faster than larger ones. This is mainly because of the proportionally greater volume of the dipteran larval mass on a relatively smaller carcass, enabling faster consumption. The larval mass also generates proportionally the same amount of heat (6,7), distributed throughout a relatively smaller body mass. Carcass disturbance also affected decomposition, particularly when measured via percent weight lost, as movement of a carcass disrupts the insect feeding activity and delays skeletalization (8). This article further explores the effect of insect exclusion on decomposition rate, but in a direct and controlled comparison of insect-accessed and insect-excluded carcasses of the same size class, using a wild rabbit (*Oryctolagus cuniculus*) model. These experiments were designed to explicitly test the effect of insect exclusion, which had been revealed in the previous study (5).

Methods

Rabbits in this experiment were divided into insect-accessed (A) and insect-excluded (X) which were either buried (B) or left on the surface of the ground (S). Numbers of animals were: AS = 24 (8), AB = 30, XS = 30 and XB = 30 (9) for a total of $n = 114$. All rabbits were killed as part of annual culling, and their use was

approved by university animal research ethics committees and official veterinarians of the UK Meat Hygiene Service. With the exception of the AS group (8), the studies were conducted primarily in the northwest of England during the months May, June, and July of 2008.

Insect-excluded rabbits (XB and XS) were placed in plastic bags immediately following culling, thus effectively precluding oviposition by blow flies. XB rabbits were placed into graves without delay after recording their weight, and thermocouples were placed under the trunk of each rabbit to record changes in carcass temperature during underground decomposition. The burial depth for all rabbits was 30–35 cm, thus mimicking a shallow, clandestine grave. Similar protocols were followed after culling with regard to the XS rabbits, which were placed into raised cages screened with 1 mm aluminum mesh. Thermocouples were inserted subcutaneously in the abdomens of seven randomly distributed carcasses (Fig. 1). XS rabbit carcasses were inspected every 3 days for evidence of insect colonization and where it was suspected, those animals were discarded from the experiment and no data collected subsequently. Appropriate control measures—including the application of pyrethrin powder and fly spray to the cages—were taken to prevent further spread. Maggots are known to remain on their food source until they finish feeding, so adjacent carcasses were unlikely to be colonized by wandering larvae (10).

For the AB rabbits (9), carcasses were left on the ground surface near the graves and exposed to normal insect activity for c. 5 hours prior to burial from 10:00 a.m. AS rabbits (8) were placed on the ground surface under chicken wire fencing to prevent scavenging for the duration of the experiment.

The data collection protocol for both the buried and surface rabbits was carried out approximately every 50 ADD for three rabbits in the AB, XB, and AS groups; all rabbits in the XS group were scored every time. This included weighing the rabbits, assigning a TBS on the scale for rabbits (8); and, for all groups except the XS group, taking soil samples for pH measurement. A thermocouple with a data logger was used to continuously record ambient temperature at the site. Two thermocouples were used to record the ground temperature external to the graves. Because disturbance of carcasses has been shown to alter the decay rate (9,11), data were collected from both groups of buried rabbits (AB and XB) and the AS rabbits only once, and then these rabbits were discarded from the experiment. Other rabbits, which remained undisturbed for successively longer periods of time, were utilized as the subsequent data points.

Decomposition score was plotted against log ADD for all four groups. This produced linear graphs, allowing comparison of the rate of decomposition across multiple variables using analysis of variance (ANOVA) and analysis of co-variance (ANCOVA). Diagnostics were used to verify the fit of the ANCOVA model. Statistical analyses were carried out using the free statistical software package "R" (12).

Results

The results indicate that oviposition occurred successfully on both AB and AS groups and was prevented (or delayed substantially) in the XB and XS groups. A representative sample of maggots from all groups were reared to adulthood and identified as *Calliphora vicina* and *Calliphora vomitoria*. In addition, one or two individuals of the families Muscidae and Anthomyiidae were also present on carcasses, but unidentified to the species level.

Figure 2 shows the regression slopes of decomposition measured by TBS on ADD for the four treatment groups. The decomposition rate of the insect-accessed before burial (AB) group clearly showed a leveling off toward the end of the experimental period. Thus, only the data before 450 ADD were used in the ANCOVA model for this group.

The rate of decomposition in rabbits where insects were excluded either by burial (XB) or by the screen (XS) showed no difference (t = 0.756, p = 0.450, df = 276), so the ANCOVA model was simplified to three slopes. The surface-exposed carcasses (AS) showed a significantly faster rate of decomposition than those which were temporarily exposed before burial (AB) (t = 3.19, p = 0.0016, df = 278), and these in turn decomposed more rapidly than where insects were excluded (XB and XS)

FIG. 1—*Rabbits in the insect excluded surface group. Two rabbits in this group are seen to contain subcutaneous intra-abdominal thermocouples with data loggers, protected from the elements by plastic bags.*
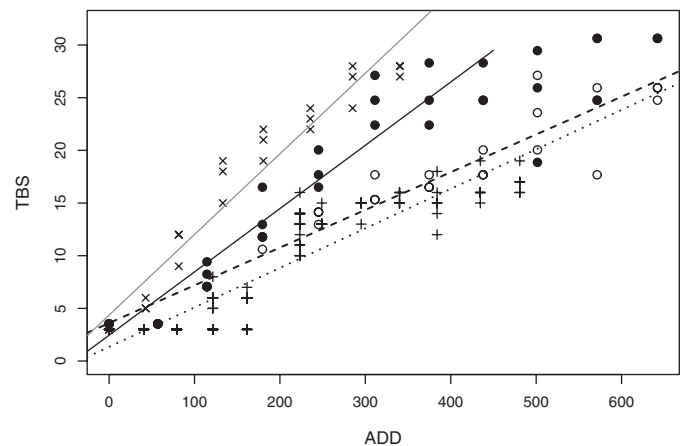

FIG. 2—*Regression slopes of total body score (TBS) versus accumulated degree days (ADD) for all groups. Insect Accessed Surface (AS) = solid gray line and "x"s (4.36 + 0.077 ADD); Insect Accessed Buried (AB) = solid black line and closed circles (2.46 + 0.060 ADD), drawn only for data used in ANCOVA analysis (ADD < 450); Insect Excluded Buried (XB) = dashed line and open circles (3.60 + 0.036 ADD); Insect Excluded Surface (XS) = dotted line and "+"s (1.34 + 0.037 ADD).*

(t = 6.50, $p < 0.001$). This untransformed ANCOVA model was a better fit than when TBS was plotted against the logarithm of ADD in terms of both the coefficient of determination ($r^2 = 0.911$ and 0.889, respectively) and when the diagnostic plots were examined.

Flies were quite persistent in attempting to lay eggs on the rabbits enclosed in the raised screen cages and several did eventually become colonized by maggots. As a by-product of the main experiment described earlier, rabbits which became infested were allowed to decompose some 5 m away from the experimental area on the ground surface. Although these rabbits were discarded from any of the experimental data presented earlier, subcutaneous intra-abdominal thermocouples in rabbits on which flies successfully laid eggs continued to record the hourly temperature over the duration of the experiment. When compared to the insect-free rabbits (XS), these temperature readings unequivocally reveal the catalytic thermal contribution of the maggot mass to the decay process.

Figure 3 shows the ambient temperature over a 10-day experimental interval (20 June to 1 July) of approximately 155 ADD, plotted against the intra-abdominal temperatures for an insect-free carcass (XS data logger 2) and an insect-accessed carcass (AS data logger 5). From 20 June (59.6 ADD) to 25 June (132 ADD), the internal temperatures of both rabbits deviated little from the ambient temperature, tracking the diurnal temperature cycle. On 25 June (132 ADD), however, data logger 5 suddenly jumped first 5°C, then rose to 10°C above both ambient and the insect-free carcass temperatures. From this point, data logger 5 recorded the temperature of the subcutaneous maggot mass in the abdomen of this rabbit, which can be seen to remain at $c.$ 5°C above ambient for the next 4 days. Throughout this period, data logger 2 within the abdomen of the insect-free carcass recorded its temperature as equivalent to that of ambient temperature. On 30 June (195.9 ADD), the subcutaneous intra-abdominal temperature recorded by data logger 5 again rose some 10°C above both ambient temperature and that of the insect-free carcass. This is consistent with the appearance of a second wave of insect colonization, with a new maggot mass reaching maturity at this point.

It is interesting to note how the maggot mass temperature cycle shows a lag from that of the ambient temperature as well as being above it, with the first departure above ambient being one of the biggest temperature spikes recorded. There is some variation in when the peaks of both ambient and "maggot" temperatures occur; the lower three maggot temperature peaks are within 30 min of the corresponding ambient peaks, but the two peaks of greatest magnitude both occur at 19:00 h and were 90 and 180 min after the ambient peak. These peaks are 64 ADD (ambient temperature) apart, which corresponds to about the length of time that larvae of *Lucilia*, which were present when oviposition occurred, will spend feeding at the recorded temperatures (calculated from [13]). Thus, the second large peak (at 196 ADD) is consistent with the second cohort of maggots.

Discussion

The results reported support earlier findings that the most important influence on decomposition rate is whether insects have access (5). Insect exclusion was achieved in two ways here, by burial and by a screen. An intermediate situation where insect colonization was allowed only a time window of a few hours, resulted in an intermediate rate of decomposition, presumably because subsequent waves of insects are required to maintain the fastest decomposition rate. It is suggested that this occurs not only because of the loss of the maggot mass after a few days, resulting in the lack of continued tissue consumption, but as importantly the reduction of the thermal energy input to the site of decomposition. The thermocouple recording the intra-abdominal temperature was not actually placed within the maggot mass itself (as these were located more in the groin area), hence the actual maggot mass temperature could be substantially higher. Although other researchers (e.g., [14]) have demonstrated the thermal effect of a maggot mass, this research has been conducted in entirely artificial conditions.

Although the calculation of postmortem interval (PMI) using ADD of ambient temperature (and a lower developmental threshold temperature cut-off) has been the norm, the results of this experiment call that practice into question, as the intra-abdominal maggot mass temperatures are minimally 5°C above that of ambient. It has certainly been shown both experimentally and anecdotally that maggot masses of a certain size can survive refrigeration and more importantly not become delayed in their development because of the fact that the maggot mass has its own higher temperature. However, this knowledge has yet to be applied to decomposition rate calculations. More research on maggot mass thermodynamics is needed to fully appreciate their influence on decomposition with respect to ambient temperature.

Another area for further investigation suggested by the results of this experiment is related to diurnal timings and temperature
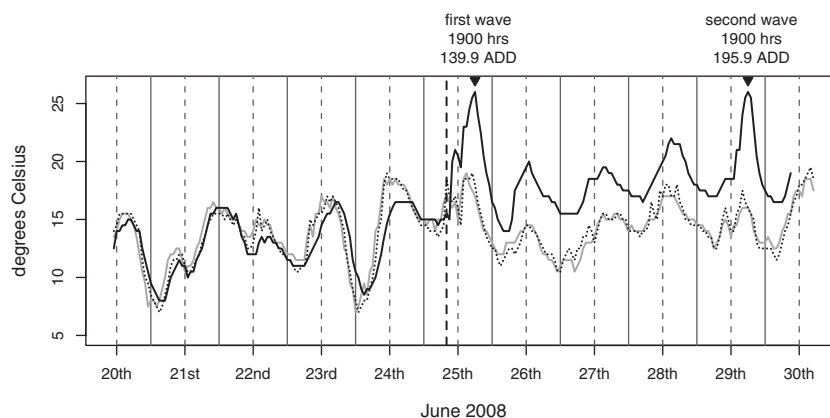

FIG. 3—*Graph of hourly ambient temperature (gray line) over a 10-day experimental interval. Intra-abdominal temperature in both an Insect Excluded Surface (XS) rabbit (dotted black line, data logger 2) and an Insect Accessed Surface (AS) rabbit (solid black line, data logger 5) are also depicted. Light gray vertical solid lines indicate 00:00 h, light gray vertical dashed lines indicate 12:00, and the dashed black vertical line indicates 132 accumulated degree days (ADD). Real time and ADD for the peaks of greatest magnitude are highlighted.*

cycling of calliphorid maggot behavioral phenomena. Although circadian rhythms have been well demonstrated in other flies (e.g., *Drosophila*), and the relative timings of some behavioral circadian characteristics (e.g., locomotor activity rhythms [15,16], oviposition [17,18], eclosion rhythm [19]), photoperiod developmental variation (20), and larval diapause (21) of various blow flies (Calliphoridae) have been noted, this has not been carried out with respect to maggot behavior, *per se*.

Conclusion

The importance of the experiment is twofold. First, the data presented show that, when time/ambient temperature is standardized as ADD, it is the presence of insects alone, via the consumption of the carcass by maggots, which accelerates decomposition. This holds true for both surface-deposited remains and for buried remains. Burial is thus a factor in decomposition solely because it normally excludes (or reduces, depending on the duration of the exposure interval between death and burial) oviposition by Diptera species. Hence, ADD can still be used to estimate the PMI in burials in the same manner as it is employed for surface depositions (22), but an insect exclusion model must be used.

Secondly, the experiment represents one of the first field-based studies that captures the intra-abdominal temperature changes associated with a maggot mass. This may have, to paraphrase an adage, opened up a can of maggots. Further work is needed to address the quantification of this phenomenon comprehensively and to investigate the implications of circadian rhythms, and the relevance of both to the estimation of the PMI.

References

1. Vass AA, Bass WM, Wolt JD, Foss JE, Ammons JT. Time since death determinations of human cadavers using soil solution. J Forensic Sci 1992;37(5):1236–53.
2. Bunch A. The impact of cold climate on the decomposition process. J Forensic Identif 2009;59(1):26–44.
3. Sharanowski BJ, Walker EG, Anderson GS. Insect succession and decomposition patterns on shaded and sunlit carrion in Saskatchewan in three different seasons. Forensic Sci Int 2008;179(2–3):219–40.
4. Magnanti BL, Williams A. Decomposition and post-mortem interval: a critical analysis of British medico-legal investigation and trends in South Yorkshire, 1995–2002. Proc Am Acad Forensic Sci 2008;14: 318.
5. Simmons T, Adlam R, Moffatt C. Debugging decomposition data—comparative taphonomic studies and the influence of insects and carcass size on decomposition rate. J Forensic Sci 2010;55(1):8–13.
6. Slone DH, Gruner SV. Thermoregulation in larval aggregations of carrion-feeding blow flies (Diptera: Calliphoridae). J Med Entomol 2007;44(3):516–23.
7. Webster R. Study of population size and species proportions of larval feeding aggregations and their effects on PMI estimations: an experimental quantitative study [unpublished MSc thesis]. Preston (UK): School of Forensic and Investigative Sciences, University of Central Lancashire, 2007.
8. Adlam RE, Simmons T. The effect of repeated disturbance on soft tissue decomposition—are taphonomic studies an accurate reflection of decomposition? J Forensic Sci 2007;52(5):1007–14.
9. Bachmann J, Simmons T. The influence of preburial insect access on the decomposition rate. J Forensic Sci 2010;55(4):doi: 10.1111/j.1556-4029.2010.01403.x
10. Byrd JH. Laboratory rearing of forensic insects. In: Byrd JH, Castner JL, editors. Forensic entomology: the utility of arthropods in legal investigations. Boca Raton, FL: CRC Press, 2001;121–40.
11. Cross P, Simmons T. The influence of penetrative trauma on the rate of decomposition. J Forensic Sci 2010;55(2):295–301.
12. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2007.
13. Grassberger M, Reiter C. Effect of temperature on *Lucilia sericata* (Diptera: Calliphoridae) development with special reference to the isomegalen- and isomorphen-diagram. Forensic Sci Int 2001;120(1–2):32–6.
14. Greenberg B. Flies as forensic indicators. J Med Entomol 1991;28(5): 565–77.
15. Saunders DS. The blow fly *Callipora vicina*: a "clock-work" insect. In: Denlinger DL, Giebultowicz JM, Saunder DL, editors. Insect timing: circadian rhythmicity to seasonality. Amsterdam: Elsevier, 2001;1–14.
16. Hong SF, Saunders DS. Internal desynchronisation in blow fly (*Calliphora vicina*) locomotor activity rhythms: evidence for a complex circadian pacemaker. Bio Rhythm Res 1998;29:387–96.
17. Amendt J, Zehner R, Reckel F. The nocturnal oviposition behaviour of blowflies (Diptera: Calliphoridae) in central Europe and its forensic implications. Forensic Sci Int 2008;175:61–4.
18. Baldridge RS, Wallace SG, Kirkpatrick R. Investigation of nocturnal oviposition by necrophilous flies in central Texas. J Forensic Sci 2006;51(1):125–6.
19. Roberts RM, Northover JM, Lewis RD. Circadian clock control of the eclosion rhythm of the brown blowfly *Calliphora stygia* (Diptera: Calliphoridae). N Z Entomol 1983;7(4):424–31.
20. Nabity PD, Higley LG, Heng-Moss TM. Light-induced variability in development of forensically important blow fly *Phormia regina* (Diptera: Callipohoridae). J Med Entomol 2007;44(2):351–8.
21. Saunders DL. Geographical strains and selection for the diapause trait in *Calliphora vicina*. In: Denlinger DL, Giebultowicz JM, Saunder DL, editors. Insect timing: circadian rhythmicity to seasonality. Amsterdam: Elsevier, 2001;113–21.
22. Megyesi MS, Nawrocki SP, Haskell NH. Using accumulated degree-days to estimate the postmortem interval from decomposed human remains. J Forensic Sci 2005;50(3):618–26.

Additional information and reprint requests:
Tal Simmons, Ph.D.
Maudland Building
School of Forensic and Investigative Sciences
University of Central Lancashire
Preston, PR1 2HE
U.K.
E-mail: tlisimmons@uclan.ac.uk

PAPER

PHYSICAL ANTHROPOLOGY; PATHOLOGY AND BIOLOGY

Jutta Bachmann,[1] *M.D., M.Sc. and Tal Simmons,*[2] *Ph.D.*

The Influence of Preburial Insect Access on the Decomposition Rate

ABSTRACT: This study compared total body score (TBS) in buried remains (35 cm depth) with and without insect access prior to burial. Sixty rabbit carcasses were exhumed at 50 accumulated degree day (ADD) intervals. Weight loss, TBS, intra-abdominal decomposition, carcass/soil interface temperature, and below-carcass soil pH were recorded and analyzed. Results showed significant differences ($p \ll 0.001$) in decomposition rates between carcasses with and without insect access prior to burial. An approximately 30% enhanced decomposition rate with insects was observed. TBS was the most valid tool in postmortem interval (PMI) estimation. All other variables showed only weak relationships to decomposition stages, adding little value to PMI estimation. Although progress in estimating the PMI for surface remains has been made, no previous studies have accomplished this for buried remains. This study builds a framework to which further comparable studies can contribute, to produce predictive models for PMI estimation in buried human remains.

KEYWORDS: forensic science, burial, decomposition, insects, postmortem interval, accumulated degree days

In forensic investigations of individual homicides as well as in contexts of conflict or war, the burial of victims as an attempt to conceal the crime and to remove traces of evidence is common. Knowledge of the decomposition patterns in single or mass graves is based mainly on anecdotal and descriptive information, producing qualitative, but no quantifiable information for precise postmortem interval (PMI) estimation in burials. This research focuses on the collection of fundamental data about decomposition in burials, with and without insect access prior to burial, and uses accumulated degree days (ADD) and total body score (TBS) (1) to derive regression formulae for the estimation of PMI. The standardized methodology presented herein also facilitates the comparability of future studies.

A wide variety of factors relevant in decomposition have been previously described, including temperature (2,3), insect activity (4–6), carnivore/rodent activity (7–11), humidity/rain (2,12,13), trauma (12,14), soil composition (15–17), body size and weight (18,19), clothing (13,20,21), the effect of disturbance (19), and embalming (22); however, the majority of these studies relate to surface decomposition. The dependence of insect activity on temperature has been well established (4,23), as has the capacity of larval masses to produce heat on their own (5), resulting in higher carcass body temperature (4,14,18,19,24–26), and thus enhanced decomposition. The importance of insects as precise PMI indicators during their first generation lifecycle, and with less precision in later decomposition stages by insect succession pattern, has been described by multiple authors (4,27–31).

Burials are characterized by soil temperatures that generally range below ambient temperature and display diurnal fluctuations of lesser magnitude. Additionally, insect access is restricted below

the ground surface and species present can differ fundamentally from those found on surface carcasses (32–34). Decreased gaseous diffusion in burials reduces the presence of micro- and macro-organisms and increases the carbon dioxide content in the soil, resulting in an anaerobic environment (15). Anaerobic and humid conditions are known to enhance adipocere formation, inhibiting further decomposition (35,36). Bodies in shallow burials are subjected to temperatures similar to ambient temperature with some diurnal fluctuations. Thermal stabilization with only seasonal variations was reported to occur at soil depths of 60 cm or deeper (25). Thus, decomposition rates decrease with burial depth because of lower temperatures, less insect access (25,32,34,37,38), and absence of scavengers (25,39). For example, White (40) observed that skeletonization in buried rabbits was not yet reached at 1050 ADD, when Vass et al.'s (15) weight-adapted formula would predict this to occur at 214.20 ADD. Therefore, burial decomposition in White's study was approximately five times prolonged compared to surface decomposition.

Lack of insect access to the carcass is considered the most important factor in the delayed decomposition of buried remains (25,34). Multiple, often interrelated factors are relevant in burial decomposition, the most fundamental of which are soil temperature and insect access (Table 1). Blow fly colonization in forensic settings at depths greater than 30 cm has not been observed (25,33), although Muscid species may occur at shallow depths of 10 cm and Phoridae species between 25 cm and 50 cm (32). Reports differ as to elevation of carcass temperature because of insect activity within 30-cm depth burials; some authors indicate temperature differences of 7–10°C between carcass and soil temperature at 30 cm depth (25), whereas others state absence of elevation of carcass temperature at that depth (33). However, remains buried directly after death did not exhibit signs of colonization, whereas those buried 48-h postmortem displayed colonization of Calliphoridae larvae 2 weeks after burial (33,39).

[1]Postweg 2, 70736 Fellbach, Germany.
[2]School of Forensic and Investigative Sciences, University of Central Lancashire, Preston, PR1 2HE, United Kingdom.

TABLE 1—*Factors influencing burial decomposition.*

Factors Influencing Burial Decomposition	Sources
Soil temperature	(33,39)
Postmortem interval before burial	(33,39)
Burial depth	(32,41)
Insect species	e.g., initially, Motter (27)
Insect density	(32)
Insect size	(32)
Scavengers	(7–10,17,39)
Soil pH	(35,41)
Soil compaction	(41)
Soil humidity	(35)
Soil type	(32,33,35,39)
Season	(32)
Habitat	(32,33,39)

Apart from quantified decomposition scores, weight loss in surface remains has been used as a tool in the assessment of PMI. Weight loss in surface decomposition is strongly related to maggot mass activity, following an s-shaped curve if insects are present, and a more linear model in the absence of arthropods (4). In buried remains, a significantly reduced loss of biomass was observed (33,39,40). No change in loss of biomass was reported between carcasses buried directly after death and those buried 48-h postmortem (39). At approximately 1080 ADD (conversion of time/temperature data from VanLaerhoven and Anderson [39]), both immediate burial and delayed burial carcasses had lost roughly 20% of their initial body weight. At approximately 6000 ADD, weight loss varied between 30% and 70%, respectively, depending on biogeoclimatic zone (33). A less than expected weight loss in buried rabbit carcasses ranging between 18.4% and 14.7% at 1072 ADD was reported because of soil adherence (40).

Weight loss in pig coffin burials at a depth of 90 cm amounted to 40% after 3 weeks' burial and to 20% of initial weight after 2 months' burial time (42). These data do not provide daily average temperatures for ADD calculation, however; therefore, no comparison is possible. Hence, considerable differences are exhibited when weight loss is used as a measure of the PMI in buried remains.

Materials and Methods

The experimental site was located at Myerscough College, Lancashire, UK, GB National Grid Reference: SD 481 408 (43). The field comprised 30 m² of semi-shaded forest area with mainly Pedunculate Oak (*Quercus robur*) and few Horse Chestnut (*Aesculus hippocastanum*), with thick undergrowth of Common Hawthorn (*Crataegus monogyna*) and Stinging Nettle (*Urtica dioica*). The soil consisted of naturally wet, loamey, and clayey floodplain soils with naturally high groundwater.

The 60 individual graves (50 cm × 30 cm × 35 cm depth) were distributed at random with a minimum of 1.5 m distance between them. A larger inter-grave distance was not feasible because of impenetrable dense tree roots and remnants of previous buildings (cement blocks, bricks) 20 cm below soil surface. The burial depth of 35 cm was chosen to represent a shallow burial, as predominant in forensic cases, and to reduce the risk of scavenger access and its impact in more superficial graves (7–10,25,43). Single burials were used to assess individual decomposition rates of rabbit carcasses over time in a large population and to avoid confounding factors such as possible interactions between adjacent carcasses. The 60 wild rabbits, *Oryctolagus cuniculus* (mean rabbit weight 1590 g, SD = 269 g), had been killed as part of standard pest control procedures. Rabbits are generally culled with a shotgun, however

resulting trauma presented no consistent pattern, with some being gut shot and others evincing shotgun trauma to the head or chest. Veterinary and ethics approval for use of the rabbits were granted from the Meat Hygiene Service, Central England Regional Office Wolverhampton (17 April 2008), and the University of Central Lancashire Health and Safety & Ethics Committee (1 May 2008).

All rabbits were buried on the same day. They had been placed in plastic bags immediately after death to avoid insect access and been cooled during storage (+4°C). The 30 rabbits of the non-insect group were buried immediately after removal from the bags and weighing. The 30 rabbits of the insect access group were exposed for 5 h to insects on the forest surface soil, prior to burial. Presence of bluebottle flies (*Calliphora vomitoria, Calliphora vicina*) and muscid flies on the carcasses was documented during this time.

Data from previous work indicated that the maximum slope of the linear phase of the decomposition curve for prediction of further decomposition until skeletonization is attained at approximately 500 ADD (40). At an expected average soil temperature of 10°C, this was calculated to be reached after approximately 50 calendar days. Therefore, the study time was defined to be 50 days. Data were obtained during 10 exhumations at 5-day intervals with three rabbits per group assessed. Reference soil temperature was obtained by a data logger (Thermocouple Data Logger, Lascar Electronics) at 35-cm soil depth on the field site, measuring temperature at 6-h intervals for the study duration.

To assess individual interface temperature immediately prior to exhumation, all rabbits were positioned with a thermocouple lead underneath the torso. Additionally, one rabbit of each second generation (Gen 2, 4, 6, 8) and all rabbits of the last generation (Gen 10) were provided with data loggers positioned below the trunk to assess interface temperature every 6 h until exhumation.

Rabbits were weighed by a spring balance (precision of 25 g) before burial and on exhumation. On exhumation, soil adherent to the carcass was removed manually as far as possible, to reduce potential weight bias. Initial reference soil pH samples from various locations on the site differed between pH 5.0 and pH 6.7. A *c*. 10 g soil sample was taken from under the torso of each rabbit at exhumation. Samples were returned to the laboratory and stored at 4°C until analyzed, within 48 h. Soil pH was determined using a 1:2.5 weight to volume solution in distilled water (16). The mixture was shaken and allowed to settle, and pH of the resulting liquid was measured using a digital pH meter.

The decomposition scoring system used was derived from a previous study (40) for rabbit burial decomposition (adapted from Megyesi et al.'s [1] TBS score). For assessment of intra-abdominal decomposition, all rabbits were systematically laparotomied on exhumation, unless spontaneous perforation had already occurred. Stages of intra-abdominal organ decay (remaining organ structure, partial liquefaction with still some recognizable organ structure, and liquefaction with absence of organ structure), in addition to intra-abdominal insect infestation, were documented. Presence of all developmental stages of insects and their location on exhumation were noted. Postfeeding larvae and pupae were collected, both directly from the carcasses and from systematic sifting of all grave soil, for further rearing and species identification.

Data were processed by use of the Statistical Package R Version R-2.6.0 (44). Analysis of covariance (ANCOVA) was used to assess differences in TBS, weight loss, interface temperature, below-carcass soil pH, and insect quantity in both groups in relation to both ADD and log ADD. The log transformation of ADD has been used previously (1) to assess decomposition rate and assists in producing a linear relationship. Correlation tests (nonparametric Spearman rank correlation, Pearson product correlation)

assessed correlations between TBS and the previously mentioned parameters separately.

Results

Decomposition

The rabbit carcasses exposed for 5 h to insects prior to burial attracted the presence of bluebottle and muscid flies. Decomposition in both groups during the early decomposition stage (until 114.5 ADD) followed the same pattern, characterized by fur loss and green discoloration of the skin, but from this point the two groups began to diverge. Purging of decomposition fluids in the head and abdominal area began in the insect group from 114.5 ADD and not until 179.5 ADD in the non-insect group. First feeding larvae appeared on carcasses (insect group #21, #24) and within the grave fill at 179.5 ADD. Bloating was not visible on any of the rabbits in either the insect or non-insect groups. From the onset of advanced decomposition in the insect group (245.1 ADD) and onwards, increasing differences between both groups became apparent (Table 2). Abdominal collapse and wet decay were initially observed at 179.5 ADD in the insect group and at 245.1 ADD in the non-insect group; however, this phase persisted much longer in the non-insect group. In the insect group, wet decay lasted from 179 to 245.1 ADD (lasting for 66 ADD), whereas in the non-insect group it persisted from 245.1 to 437.7 ADD (lasting for 192.5 ADD). Adipocere, visible predominantly as a greasy texture on underside abdominal surfaces, also became evident at 245.1 ADD in both groups. The presence of adipocere persisted in copious amounts until 374.6 ADD and in gradually diminishing quantities thereafter until the end of the experimental period at 641.8 ADD. Larval infestation was observed most frequently in the genital area, followed by infestation of the oral region, reflecting preferred observed oviposition sites before burial. Correspondingly, hind legs showed consistently enhanced decomposition in both the insect and non-insect groups. Bone exposure in insect group carcasses was apparent from 245.1 ADD, and skeletonization occurred at 501.37 ADD. For the non-insect group, bone exposure occurred first at 437.7 ADD, and skeletonization was not observed until 641.8 ADD.

The following general skeletonization sequence was observed:

Hindlegs ► forelegs and/or head (maxilla/mandible) ► horax (ribcage/scapulae) ► pelvis

Similarly, disarticulation started on hind and forelegs, followed by head and trunk areas, occurring at 311.12 ADD in the insect,

and 501.62 ADD in the non-insect group. During late advanced decomposition (641.87 ADD) at the end of the field study, the final decomposition stages were characterized by generalized desiccated black skin over disarticulated, skeletonized (>50% of scored area) bones, and complete liquefaction of internal organs. The non-insect group carcasses showed a consistent lesser degree of disarticulation and skeletonization than insect group carcasses.

Intra-Abdominal Decomposition

The following decomposition stages of intra-abdominal organs were observed:

- Intra-abdominal organ structure retained
- Partial liquefaction with some recognizable organ structures
- Liquefaction with absence of organ structure

Partial liquefaction occurred in the insect group from 311.12 ADD onwards, in the non-insect group from 437.78 ADD onwards, starting with dissolution of intestinal walls. Complete liquefaction (loss of identifiable organ structures) was achieved at 501.62 ADD in the insect and 571.37 ADD in the non-insect group. Intra-abdominal decomposition stages were integrated into the existing TBS (1,40). Partial liquefaction in both groups was mainly observed at a score of 7 for the abdominal area, and complete liquefaction at an abdominal decomposition score of 8 during advanced decomposition.

Total Body Score

Decomposition scores of insect and non-insect groups began to diverge after 57 ADD (Fig. 1). After 311.12 ADD, decomposition plateaus in both groups, analogous to the appearance of 1st/2nd instar larvae on a number of both insect and non-insect group carcasses, along with first insect cycle postfeeding larvae and pupae. Presence of insect larvae in the non-insect carcasses will be discussed later. Both the insect and non-insect groups showed signs of early decomposition from 114.5 ADD onwards. Insect group carcasses entered into advanced decomposition as defined by White (40) for rabbits at 245.12 ADD, whereas the non-insect group came into that stage predominantly at 374.62 ADD.

Statistical analysis (ANCOVA) showed a significant difference in decomposition rates between insect and non-insect groups ($F_{1,56} = 67.2$, $p \ll 0.001$) in relation to log ADD (Fig. 2). The insect access group showed faster decomposition, indicated by the

TABLE 2—*Summary table of observed decomposition stages in relation to ADD for insect and non-insect groups. The values provided for both ADD and the corresponding exhumation group reflect the last time at which the stage was observed discretely.*

| Decomposition Stage	Observed Stage	ADD		Exhumation Group	
		Insect Group	Non-Insect Group	Insect Group	Non-Insect Group
Fresh stage	Fresh	0	0	0	0
Early decomposition	Fur loss	57	179.5	1	1–3
	Discoloration of skin	57	179.5	1	1–3
	Purging	114.5	179.5	2	1–3
	Wet decay	179.5	374.6	3	4–6
	Abdominal collapse	179.5	374.6	3	4–6
Advanced decomposition	First bone exposure	245.1	437.7	4	7
	Partial liquefaction	311.1	437.7	5	7
	Disarticulation	374.6	437.7	6	7
	Black skin on all areas	437.7	571.3	7	8–9
	Complete liquefaction	501.6	641.8	8–10	10
	Disarticulation and skeletonization	501.6	641.8	8–10	10
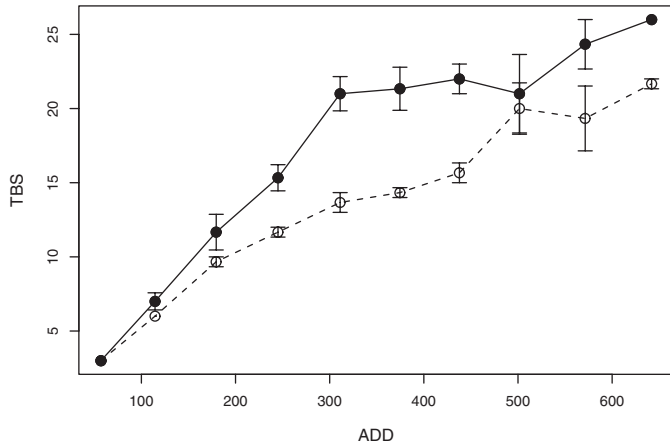
FIG. 1—*Mean total body score values of insect and non-insect groups with respective standard errors in relation to accumulated degree day (open circles and dashed line = non-insect; dots and solid line = insect).*

significant difference in the regression line slopes between both groups ($t = 3.68$, $p < 0.001$, df = 56).

The following regression formulae resulted for the two groups:

Insect group: TBS = $-39.4 + 23.38 \times$ log ADD
Non-insect group: TBS = $-29.49 + 17.63 \times$ log ADD

Insects

A total of 76.6% of all insect group rabbits and 10% of the non-insect group showed presence of first generation arthropods on the carcass or in the surrounding grave fill on exhumation. Overall numbers of insects present on exhumation in both groups were very limited, and true larval masses were never observed. Feeding larvae were apparent for the first time in the insect group at 179.5 ADD and in the non-insect group at 245.12 ADD. From 374.62 ADD
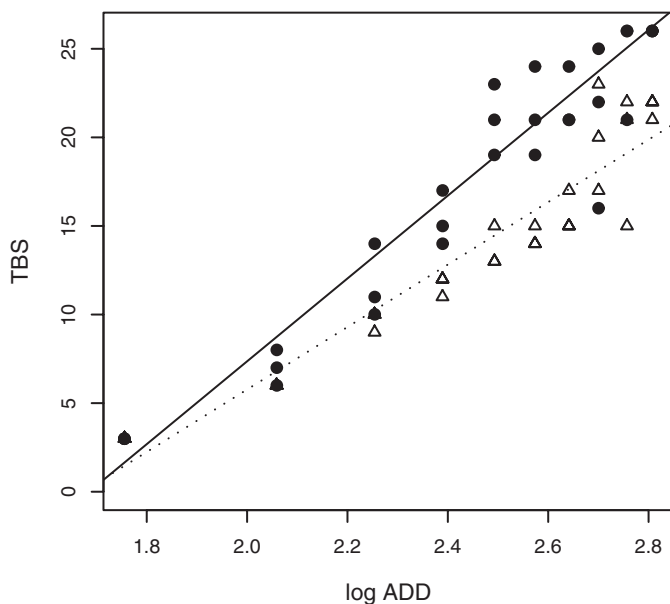
FIG. 2—*Regression lines of total body score against log accumulated degree day of insect (indicated by circles and solid line) and non-insect group (indicated by triangles and dashed line), showing enhanced decomposition in the insect access group.*

TABLE 3—*Contingency table of insect infested-carcasses of both groups.*

	Insect Group	Non-Insect Group
Insect presence on carcass/grave fill on exhumation	23/30 Total 9 of those additional late infestation	9/30 Total 3 exclusively early infestation 6 exclusively late infestation
No insects on exhumation	7/30 Total (6 of those during initial phase until 179.5 ADD)	21/30 Total (6 of those during initial phase until 179.5 ADD)

ADD, Accumulated Degree Day.

onwards, small-sized feeding larvae (1st or 2nd instar) in minute quantities were observed on the carcass, often simultaneously with more advanced larval stages and/or pupae colonizing both carcass groups (Table 3). Adult Coleoptera species were not observed on the carcasses, and only very few adult members of the Carabidae family were found in the grave fill.

All postfeeding larvae and pupae stages identified on carcasses or in grave fill were collected (between 245.12 and 571.37 ADD) and reared for species identification (Table 4).

Interface Temperature

Throughout the experiment reference soil temperature (at 35 cm depth) ranged between 11.5 and 13.5°C; mean interface temperatures from the insect group ranged between 12°C and 15.5°C and those of non-insect groups between 12.3 and 16.3°C (Fig. 3). ANCOVA showed no significant differences between reference soil temperature and interface temperatures of both insect ($F_{1,36} = 3.38$, $p = 0.07$) and non-insect groups ($F_{1,36} = 1.17$, $p = 0.28$) in relation to log ADD. In addition, ANCOVA indicated no significant difference in the interface temperatures between insect and non-insect groups ($F_{1,16} = 2.75$, $p = 6.90$) in relation to log ADD.

Weight Loss

Exhumed carcasses showed extensive soil adherence to the body surface, especially at later decomposition stages. Weight loss occurred earlier and at an initially faster rate in the insect group, and the final remaining weight in the insect group was slightly lower (77.01%) than in the non-insect group (79.2%) (Fig. 4). ANCOVA showed that differences between the groups were statistically significant ($F_{1,55} = 4.26$, $p = 0.043$), with higher weight loss in the insect-access group in relation to log ADD.

Soil pH

Soil pH mean values showed a maximum increase of 1.95 pH levels in the insect group and 1.66 in the non-insect group over the total study time. Peak alkalinity in both groups was reached at 311.1 ADD with pH values of 8.89 in the insect and 8.74 in the non-insect group, respectively. ANCOVA showed no significant difference in below-carcass soil pH values between insect and non-insect groups ($F_{1,56} = 0.13$, $p = 0.71$).

Discussion

Decomposition

The short bloating stage might have been missed in this study because of timing of the first exhumation at 57 ADD, as others

TABLE 4—*Identified adult species from collected and reared postfeeding and pupal insect stages (I = Insect group; NI = Non-insect group).*

Order	Family	Species	Rabbit Group	Collected at Accumulated Degree Day
Diptera	Calliphoridae	*Calliphora vicina* spp	I	245.12
				374.62
				437.78
				571.37
		Calliphora vomitoria spp	I	374.62
				501.62
	Muscidae		I	245.12
	Anthomyiidae		NI	501.62
Coleoptera	Carabidae	*Nebria brevicollis* spp	NI	501.62
		Abax parallelopipedus spp	I	501.62
		Leistus (genus)	I	571.37
		Loricera pilicornis spp	I	571.37

FIG. 4—*Percent remaining mean weight of insect and non-insect group with respective standard errors in relation to accumulated degree day (open circles and dashed line = non-insect; dots and solid line = insect).*

FIG. 3—*Comparison between reference soil temperature (35 cm) and interface temperatures of insect and non-insect group (with respective standard errors) in relation to accumulated degree day (crosses and grey line = reference temperature; open circles and dashed line = non-insect interface temperature; dots and solid line = insect interface temperature).*

(26) have reported bloating in surface rabbits already at 42.8 ADD. During early phases, decomposition pattern and rates were comparable in both groups, differences becoming more marked after 179.5 ADD. The distinct differences in decomposition patterns and rates between both groups after 179.5 ADD coincided with the observed presence of developing feeding larvae of the first insect cycle on the carcasses. For TBS, the differences of greatest magnitude between groups were evident at 311.12 ADD; this was related to the simultaneous emergence of 1st and 2nd instar feeding larvae, emphasizing the catalytic impact of insect activity on decomposition rates. After this point, decomposition rates in both groups became parallel (Fig. 1), and this pattern persisted with minor variations thereafter.

Advanced decomposition was characterized by increasing bone exposure, disarticulation, and liquefaction, the insect group preceding the non-insect group in decomposition by ca. 200 ADD between TBS 15 and 20 (Table 2 and Fig. 1). This corresponds to an approximately 30% retarded decomposition process in the absence of insect exposure for the duration of the study (641.87 ADD). The greatest divergence in decomposition sequence pattern
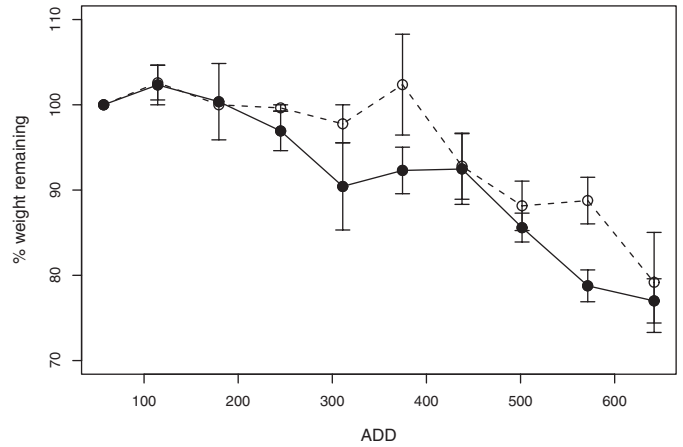
between groups was observed in the appearance of bone exposure (Table 2). On average, non-insect carcasses showed signs of skeletonization on the hind limbs 300 ADD later than in the insect group, emphasizing the importance of localized insect infestation on the advancement of skeletonization. As mentioned earlier, despite the fact that rabbits were dispatched with shotguns, there was no consistent trauma pattern. Thus, preferential insect presence in the genital region cannot be seen to relate to the gunshot trauma.

The observed differences in decomposition rates between groups were confirmed by statistical analysis, indicating highly significant differences of decomposition rates in relation to ADD between insect and non-insect groups ($p \ll 0.001$). This is in line with previous research (45) stating that decomposition rates in various environments follow—in relation to ADD—the same pattern and differ solely in regard to presence or absence of insects.

Intra-Abdominal Decomposition

Liquefaction, previously classified as characteristic of early decomposition, was shown to begin only during advanced decomposition, at 311.12 ADD and 437.78 ADD in the insect and non-insect groups, respectively. Complete liquefaction was achieved during late advanced decomposition stages after 500 ADD, as carcasses were approaching the skeletonization stage. The earlier liquefaction in the insect group may be linked to intrusion of feeding larvae into the abdominal area from their initial genital location. Thus, an early larval perforation of intestines may have resulted in the liberation of intestinal bacteria and enzymes (46), further enhancing intra-abdominal decomposition.

Insects

One of the prime observations of this study included the low quantities of insects found on exhumed carcasses and in grave fill during all exhumations. There was absence of true larval masses, confirming previous observations (39). Only a minor percentage (10%) of non-insect carcasses displayed presence of first generation insects on exhumation as opposed to 76.6% of the insect group. This is in line with reports about absence of insect colonization in carcasses buried directly after death; in contrast, cadavers buried 48-h postmortem exhibited Calliphoridae larvae after 2 weeks burial time (33,39). Insect species recovered during exhumations were mainly members of the Calliphoridae family, namely *Calliphora*

vicina and *Calliphora vomitoria*, followed by few of the family Muscidae, reflecting predominant necrophagous species identified by former research (32). The only adult Coleoptera species sporadically found on exhumation were members of the Carabidae family; the families Staphilidae and Silphidae were not present, which corresponded to earlier published observations (32,47). Previously reported presence of smaller-sized Staphylinidae in a burial depth of 30 cm (39) could not be confirmed.

The feeding larvae observed at 179.5 ADD correspond to first generation development stages of bluebottle flies (at given soil temperatures/ADD) that were evident on the carcasses before burial. Thus, oviposition did occur within the 5 hours of insect exposition of carcasses before burial. The origin of small feeding larvae (1st or 2nd instar) observed from 374.62 ADD onwards on insect and as well several non-insect group carcasses, remains to be clarified. Considering the timeframe of 38.8 calendar days at 12°C indicated by previous research (48), or 465.6 ADD, for time from oviposition until adult emergence of *Calliphora vicina*, it is unlikely that the small feeding larvae seen at 374.62 ADD (corresponding to approximately 30 calendar days at average temperatures of 12.4°C) represent the second generation cycle of insects on the carcasses. This would indicate emergence at 96 ADD earlier than expected within these thresholds. In addition, empty pupae shells evident in the graves as late as 571.37 ADD imply first generation adult emergence at a point later than 374.62 ADD. Therefore, feeding larvae may have originated from other species, e.g. members of the Diptera order/Phoridae family or Coleoptera order (32,39) that had gained secondary access to the buried remains. However, because of the lack of appropriate facilities, further rearing and species identification of these larvae could not be realized during the course of this study to clarify this issue. Alternatively, the developmental cycle of Calliphoridae species might have been affected by unidentified factors other than ADD, by virtue of it occurring underground.

Interface Temperature

Mean interface temperatures of both insect and non-insect groups were not statistically different to reference soil temperature on site. Interface temperatures of insect and non-insect groups were also not statistically different, although the peak of interface temperature in the insect group occurred at 311.1 ADD, the same peak as for the maximum divergence in TBS score between insect and non-insect groups. Carcass temperature increases with larval mass size (49,50), thus absence of an interface temperature increase would indicate absence of relevant larval mass activity of significant size in both groups, which corresponds to the visual observations during the study. Therefore, this study confirms former findings (33), where no elevation of carcass temperature in buried carcasses in 30-cm depth was observed.

Weight Loss

After an initial steady state, weight loss curves in the non-insect group followed a rather linear model as described for weight loss in absence of arthropods (4). The weight loss curve in the insect group appeared as a mixture of a sigmoid curve, as mentioned for surface decomposition with insect access (4), and the previously mentioned linear model in the absence of insects. The relatively constant weight loss in the insect group may have resulted from low quantities of feeding larvae over the total study duration, leading to a constant rate of tissue consumption. Thus, the absence of an exponential rate of weight loss is possibly because of the

absence of re-infestation of ovipositing insects and the subsequent continued presence of new insect feeding colonies. In the non-insect group, a rapid decline in body weight to 79.2% of initial weight occurred during the period from 374.62 ADD to 641.87 ADD. This weight loss may be linked to the emerging 1st and 2nd instar feeding larvae on the carcasses at that time, resulting in increased soft tissue consumption. The small difference in remaining weight in both groups is similar to prior findings (39), which displayed no difference in weight loss between carcasses buried directly postmortem, and delayed burials after 48 h. However, after conversion of those data into ADD, buried carcasses had lost 20% of their initial weight at approximately 1080 ADD, thus decomposition in that study occurred at a slower rate than the findings of this study suggest.

Nevertheless, weight measurement on buried carcasses is subject to various biases, including soil adherence to the carcass surface, moisture content of the carcass and its adherent soil, and loss of superficial tissue by the cleaning procedure. Plastic netting used to facilitate recovery of carcasses on exhumation, resulted similarly in some tissue loss because of net adherence to the carcass, especially in advanced decomposition stages. Hence, TBS is preferable to per cent weight loss as a measure of decomposition. A possible bias from carcass maggot (mass) weight was irrelevant in this study because of their small numbers.

Soil pH

Beneath-carcass soil pH was shown to increase in alkalinity during phases of soft tissue decay, with peaks at 311.12 ADD in both groups. Overall, below-carcass pH values were not strongly correlated ($p = 0.003*$ in the insect group and $p = 0.013*$ in the non-insect group)[1] with decomposition scores in both groups, thus their utility for assessment of the decomposition stage could not be confirmed. In addition, the presence of construction materials in the soil possibly influenced individual graves to a varying degree, thus adding bias to soil pH measurement. This also suggests that soil pH in buried remains in forensic cases may be subject to the same bias, depending on the context.

Conclusion

This research project assessed fundamental parameters in subsurface decomposition with or without insect access prior to burial. A comparison with previously published results was difficult because of *de facto* absence of comparable data; therefore, the findings of this study introduce data in a format enabling future comparisons to be made. The consistent use of ADD in decomposition studies is crucial to standardize data collection and enables comparison between different geographic locations, climate zones, seasons, or merely populations in the same study location. It will also promote a better understanding and use of existing ADD-based entomological data for both buried and surface remains.

In conclusion, TBS is the variable most closely related to assessing the progress of decomposition in relation to ADD, thus proving to be the most valid tool in PMI estimation. TBS quantifies previously described visual stages in decomposition. Other assessed parameters such as weight loss, below-carcass soil pH, and interface temperature were shown to be much less closely related to the

---

[1]The * marked numbers indicate correlations categorized as significant, but where strict use of the Bonferroni correction would make them nonsignificant.

various decomposition stages and hence, were of little added value for PMI estimation in burials. The highly significant differences in decomposition rates between the insect and the non-insect groups emphasize the need to take postmortem time prior to burial, i.e., insect exposure, into account for calculation of PMI in forensic burial cases. This may necessitate the collection of larvae and pupae from both carcass and grave soil for assessment of insect presence, development stage, and species identification for PMI calculation. The mere presence of pupal cases suggests that insects were a factor in decomposition within a burial and therefore should be taken into account for PMI estimation, as a strict non-insect model would not be appropriate. Insect access in buried remains is as decisive a factor in influencing decomposition rate and pattern as it is in surface remains. Definitive equations for calculating PMI from the decomposition stages of clandestine buried human remains cannot be produced from this work. Retrospective studies considering the TBS of human remains recovered from clandestine burials, in reference to ADD, conducted along the same parameters as those in surface remains (1) would assist in delivery of this outcome.

References

1. Megyesi MS, Nawrocki SP, Haskell NH. Using accumulated degree-days to estimate the postmortem interval from decomposed human remains. J Forensic Sci 2005;50(3):618–26.
2. Archer MS. Rainfall and temperature effects on the decomposition rate of exposed neonatal remains. Sci Justice 2004;44:35–41.
3. Komar DA. Decay rates in a cold climate region: a review of cases involving advanced decomposition from the Medical Examiner's office in Edmonton, Alberta. J Forensic Sci 1998;43(1):57–61.
4. Payne JA. A summer carrion study of the baby pig *Sus scrofa* Linnaeus. Ecology 1965;46(5):592–602.
5. Haskell N, Hall R, Cervenka V, Clark M. On the body: insects' life stage presence and their postmortem artifacts. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press, 1997;415–41.
6. Bourel B, Martin-Bouyer L, Hedouin V, Cailliez JC, Derout D, Gosset D. Necrophilous insect succession on rabbit carrion in sand dune habitats in northern France. J Med Entomol 1999;36:420–5.
7. Haglund W. Dogs and coyotes: post-mortem involvement with human remains. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press, 1997;367–82.
8. Haglund W. Rodents and human remains. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press, 1997;405–14.
9. Haglund WD, Reay DT, Swindler DR. Tooth mark artifacts and survival of bones in animal scavenged human skeletons. J Forensic Sci 1988;33(4):985–97.
10. Haglund WD, Reay DT, Swindler DR. Canid scavenging/disarticulation sequence of human remains in the Pacific Northwest. J Forensic Sci 1989;34(3):587–606.
11. Murmann DC, Brumit PC, Schrader PA, Senn DR. A comparison of animal jaws and bite mark patterns. J Forensic Sci 2006;51(4):846–60.
12. Micozzi MS. Experimental-study of postmortem change under field conditions—effects of freezing, thawing, and mechanical injury. J Forensic Sci 1986;31(3):953–61.
13. Spennemann DHR, Franke B. Decomposition of buried human bodies and associated death scene materials on Coral Atolls in the Tropical Pacific. J Forensic Sci 1995;40(3):356–67.
14. Cross P, Simmons T. The influence of penetrative trauma on the rate of decomposition. J Forensic Sci 2010;55(2):295–301.
15. Vass AA, Bass WM, Wolt JD, Foss JE, Ammons JT. Time since death determinations of human cadavers using soil solution. J Forensic Sci 1992;37(5):1236–53.
16. Hopkins DW, Wiltshire PEJ, Turner BD. Microbial characteristics of soils from graves: arm investigation at the interface of soil microbiology and forensic science. Appl Soil Ecol 2000;14:283–8.
17. Wilson AS, Janaway RC, Holland AD, Dodoson HI, Baran E, Pollard AM, et al. Modelling the buried human body environment in upland climes using three contrasting field sites. Forensic Sci Int 2007;169:6–18.
18. Hewadikaram KA, Goff ML. Effect of carcass size on rate of decomposition and arthropod succession patterns. Am J Forensic Med Pathol 1991;12(3):235–40.
19. Adlam RE, Simmons T. The effect of repeated physical disturbance on soft tissue decomposition—are taphonomic studies an accurate reflection of decomposition? J Forensic Sci 2007;52(5):1007–14.
20. Aturaliya S, Lukasewycz A. Experimental forensic and bioanthropological aspects of soft tissue taphonomy: 1. Factors influencing postmortem tissue desiccation rate. J Forensic Sci 1999;44(5):893–6.
21. Weitzel MA. A report of decomposition rates of a special burial type in Edmonton, Alberta from an experimental field study. J Forensic Sci 2005;50(3):641–7.
22. Sledzik PS, Micozzi MS. Autopsied, embalmed, and preserved human remains: distinguishing features in forensic and historic contexts. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press, 1997;483–95.
23. Davidson J. On the relationship between temperature and rate of development of insects at constant temperatures. J Anim Ecol 1944;13(1):26–38.
24. Mann RW, Bass WM, Meadows L. Time since death and decomposition of the human body: variables and observations in case and experimental field studies. J Forensic Sci 1990;35(1):103–11.
25. Rodriguez WC, Bass WM. Decomposition of buried bodies and methods that may aid in their location. J Forensic Sci 1985;30(3):836–52.
26. Adlam RE, Simmons T. Assessing the effect of repeated physical disturbance associated with data collection in experimental decomposition studies. In: Proceedings of the 58th Annual Meeting of the American Academy of Forensic Sciences; 2008 Feb 18–23; Seattle, WA. Colorado Springs, CO: American Academy of Forensic Sciences, 2006.
27. Motter MG. A contribution to the study of the fauna of the grave. A study of one hundred and fifty disinternments with some additional experimental observations. NY Entomol Soc 1898; VI(4):201–31.
28. Reed HB. A study of dog carcass communities in Tennessee with special reference to the insects. Am Midl Nat 1958;59(1):213–45.
29. Johnson MD. Seasonal and microseral variations in the insect population on carrion. Am Midl Nat 1975;93(1):79–90.
30. Rodriguez WC, Bass WM. Insect activity and its relationship to decay rates of human cadavers in east Tennessee. J Forensic Sci 1983;28(2):423–32.
31. Grassberger M, Frank C. Initial study of arthropod succession on pig carrion in a Central European urban habitat. J Med Entomol 2004;41(3):511–23.
32. Lundt H. Ökologische Untersuchungen über die tierische Besiedlung von Aas im Boden. Pedobiologia 1964;4:158–80.
33. VanLaerhoven SL, Anderson G. Insect succession on buried carrion in two biogeoclimatic zones of British Columbia. J Forensic Sci 1999;44(1):32–41.
34. Campobasso CP, Di Vella G, Introna F. Factors affecting decomposition and Diptera colonization. Forensic Sci Int 2001;120:18–27.
35. Forbes SL, Dent BB, Stuart BH. The effect of soil type on adipocere formation. Forensic Sci Int 2005;154:35–43.
36. Forbes SL, Stuart BH, Dent BB. The effect of the burial environment on adipocere formation. Forensic Sci Int 2005;154:24–34.
37. Rhine S, Dawson JE. Estimation of time since death in the Southwestern United States. In: Reichs KJ, editor. Forensic osteology. Advances in the identification of human remains. Springfield, IL: Charles C. Thomas, 1998;145–60.
38. Sledzik PS. Forensic taphonomy: postmortem decomposition and decay. In: Reichs KJ, editor. Forensic osteology. Advances in the identification of human remains. Springfield, IL: Charles C. Thomas, 1998;109–19.
39. VanLaerhoven SL, Anderson G Forensic entomology: determining time of death in buried homicide victims using insect succession, Technical Report TR 02-96. Ottawa, Ontario: Canadian Police Research Center, 1996.

40. White R. Decomposition within a mass burial and the implications for post-mortem interval [unpublished Master's dissertation]. Preston, UK: University of Central Lancashire, 2006.

41. Rodriguez WC. Decomposition of buried and submerged bodies. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press, 1997;459–67.

42. Payne JA, King EW, Beinhart G. Arthropod succession and decomposition of buried pigs. Nature 1968;219:1180–1.

43. Ordnance Survey, http://www.ordnancesurvey.co.uk/oswebsite/2008 (accessed August 20, 2008).

44. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2007.

45. Simmons T, Adlam RE, Moffatt C. Debugging decomposition data—comparative taphonomic studies and the influence of insects and carcass size on decomposition rate. J Forensic Sci 2010;55(1):8–13.

46. Dent BB, Forbes SL, Stuart BH. Review of human decomposition processes in soil. Environ Geol 2004;45:576–85.

47. Shubeck PP, Blank DL. Carrion beetle attraction to buried fetal pig carrion (Coleoptera: Silphidae). Coleopt Bull 1982;36(2):240–5.

48. Marchenko MI. Medicolegal relevance of cadaver entomofauna for the determination of the time of death. Forensic Sci Int 2001;120:89–109.

49. Webster R. Study of population size and species proportion of larval feeding aggregations and their effects on PMI estimations: an experimental quantitative study [unpublished Master's dissertation]. Preston, UK: University of Central Lancashire, 2007.

50. Slone DH, Gruner SV. Thermoregulation in larval aggregations of carrion-feeding blow flies (Diptera: Calliphoridae). J Med Entomol 2007;44(3):516–23.

Additional information and reprint requests:
Tal Simmons, Ph.D.
Maudland Building
School of Forensic and Investigative Sciences
University of Central Lancashire
Preston, PR1 2HE
U.K.
E-mail: tlsimmons@uclan.ac.uk

PAPER

CRIMINALISTICS

Zhaoshu Zeng,[1] *Ph.D.; Hongtao Yan,*[1] *Ph.D.; Li Wang,*[2] *M.D.; Enwu Yuan,*[3] *M.D.;*
Weihong Yang,[1] *M.D.; Zhigang Liao,*[4,5] *M.D.; and Ziming Dong,*[1] *M.D., Ph.D.*

Genome-wide Screen for Individual Identification SNPs (IISNPs) and the Confirmation of Six of Them in Han Chinese with Pyrosequencing Technology*

ABSTRACT: Single nucleotide polymorphisms (SNPs) offer promise to forensic DNA analysts, but it remains uncertain whether a panel of individual identification SNPs can be as informative as the Combined DNA Index System short tandem repeats. Based on the highly accurate and publicly available HapMap SNP database (r21a) and a minor allele frequency cutoff of ≥ 0.45, we completed a genome-wide screen through 3,905,819 SNPs with internally modified computer programs and identified 1439 SNPs with high heterozygosity and low F_{st} values among four populations (Utah Caucasian, Han Chinese, Tokyo Japanese, and Nigerian Yoruba). Using pyrosequencing technology, we studied six loci in a relatively large group of samples to determine whether these loci were as informative as the HapMap data suggest. These SNPs performed as expected in the Han Chinese in terms of heterozygosity and F_{st}. The 1439 identified SNPs should provide a comprehensive and reliable set of loci for identity and relationship testing.

KEYWORDS: forensic sciences, individual identification, SNP, HapMap database, minor allele frequency (MAF), Han Chinese

In the field of forensic sciences, PCR-based short tandem repeat (STR) polymorphism analyses have been successfully used for human identification and parentage testing for more than 10 years (1,2). The 13 loci recommended by the Federal Bureau of Investigation, called the Combined DNA Index System (CODIS) core STRs (3), have been applied universally because of their high polymorphism among various populations, straightforward interpretation to digital data records, and sophisticated typing technology. However, with recent technological advances in DNA sequencing that allow the detection of single-base differences, we are approaching a time when all differences in DNA sequence among individuals can be identified. Single nucleotide polymorphisms (SNPs) have been by far the most prevalent of all DNA sequence variations and will be the markers of choice in complex disease mapping. These will thus be the focus of the next phase of the human genome project (4,5).

Both the advantages and disadvantages of SNP applications have been discussed in the forensic community (6,7). Two main advantages of applying SNPs to identity and relationship testing are obvious: a higher recovery of information from degraded DNA samples and a lower mutation rate, *c.* 100,000 times lower than that of STRs (10^{-8} vs. 10^{-3}) (8). However, limitations such as lower polymorphic information content (PIC), mixture resolution issues/interpretation, the large amount of DNA required for multiplexing assays, and the lack of widely established core loci must still be overcome before SNPs can be used successfully. It is commonly thought that SNPs may serve as an adjunct to STRs for solving special problems in forensic genetics (9).

However, this does not mean that SNP-based assays should not be explored. In fact, explorations of SNPs in forensics have been ongoing for several years. Among the most recent and most successful, Pakstis et al. (10) reported 40 SNPs with extremely low allele frequency variation among 40 populations; Sanchez et al. (11,12) recommended 52 SNPs for human identification; Lee et al. (13) published 24 highly informative SNP markers in Koreans; and Vallone et al. (14) investigated 70 SNP markers among 189 U.S. Caucasian, African-American, and Hispanic samples. In addition, many other reports have focused on a variety of forensic interests (15).

Despite these studies, a large amount of recently generated genotype data, which could potentially reveal additional "individual identification SNPs" (IISNPs), remains unexplored. At the end of November 2008, *c.* 4,029,145 nonredundant SNPs had been genotyped and recorded in the main public databases (HapMap data release #26, http://www.hapmap.org/; or NCBI SNP Build36 [dbSNP b126], http://www.ncbi.nlm.nih.gov/snp), while fewer than 1.5 million SNPs had been screened for forensic purposes (11,12). Identifying all (or most) of the potential IISNPs in the human

[1]Center of Forensic Identification, School of Basic Medical Sciences, Zhengzhou University, Zhengzhou, Henan 450052, China.
[2]Department of Parasitology, School of Basic Medical Sciences, Zhengzhou University, Zhengzhou, Henan 450052, China.
[3]Department of Clinical Laboratories, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China.
[4]Department of Forensic Pathology, College of Pre-clinical and Forensic Medicine, West China Medical Center, Sichuan University, Chengdu, Sichuan 610044, China.
[5]Posthumous.

genome is a key step toward the successful forensic application of SNPs. This kind of screening is a necessary prerequisite for the comprehensive follow-up assessments that will be required before a universal multiplex panel can be developed for forensic application.

However, such screening has proven difficult because the data volume exceeds the processing capacities of most personal computers unless specialized configurations and software are utilized. In our attempt to develop a SNP-based marker panel for forensic applications, we have determined that it is important to look for SNPs where the minor allele frequency (MAF) is between 0.45 and 0.50. We have self-edited a set of computer programs and used them to perform a whole genome search through the 3,905,819 nonredundant SNPs in the HapMap database r21a (or NCBI SNP database Build125, released on January 2007) for all the shared identity testing SNPs among four populations: U.S. residents with northern and western European ancestry at the Centre d'Etude du polymorphisme humain (CEPH or CEU), Han Chinese in Beijing (HCB), Japanese from Tokyo (JPT), and Yorubans of Ibadan, Nigeria (YRI). Here, we report the theoretical retrieval of 1439 SNPs with high heterozygosity and low coefficient of gene differentiation (F_{st}). We also report the experimental validation of six of these SNPs using a group of 144 Han Chinese using pyrosequencing technology (16–18), which confirmed that these six SNPs were equally informative in larger sample sizes, and demonstrated the potential forensic utility of these SNPs. This experimental analysis is expected to help in understanding how reliable these 1439 SNPs will be for future identity testing SNP applications.

Materials and Methods

Selection of Highly Informative SNPs

Bulk data of the whole genome-genotyped SNP allele frequencies from the HapMap Public Release #21a were downloaded from the HapMap Project website (http://www.hapmap.org/downloads/frequencies/?N=D). Files in.gz format were unzipped and transferred to the Microsoft Word-readable.doc format. Several self-edited computer programs written in Visual Basic Application (VBA) language were applied consecutively in the Microsoft Word environment with a number of different MAF criteria (0.49–0.50, 0.48–0.50, 0.47–0.50, 0.46–0.50, 0.45–0.50, 0.44–0.50) to extract SNPs from each frequency file. Qualified SNPs were automatically transferred to a new.doc file. For each chromosome, the reference SNP ID (rs#) of selected SNPs in each population was compiled into a Microsoft Excel file, and the numbers of selected SNPs as well as SNPs identified in all four populations were calculated and listed separately.

Subsequently, SNPs that displayed a MAF ≥ 0.45 in all four populations were further studied to avoid theoretical linkage (19). Six of these 1439 loci were investigated to confirm whether they were as informative as their HapMap data suggested when the number of samples (taken from the Han Chinese population) increased from 45 to 144. Because the genotyping of six SNPs was a small-scale undertaking, the selection procedures were set to be strict. The SNPs chosen all had high heterozygosity and low F_{st} among the four populations (CEU, HCB, JPT, and YRI), conformed to Hardy–Weinberg equilibrium, and had appropriate GC content (45–55%) in their flanking sequences to avoid technical challenges and to guarantee good experimental performance (20). Finally, we ensured that the chosen loci were not closely linked by selecting SNPs on different chromosomes or at distant chromosomal sites. Based on these criteria, we selected six SNPs, rs6427658, rs4607417, rs4722616, rs7907658, rs2369522, and rs749305, for experimental validation.

DNA Extraction and Quantification

Peripheral vein blood (1 mL) was collected from 72 male and 72 female unrelated Han Chinese in Henan Province with informed consent. Genomic DNA from each sample was prepared with the phenol–chloroform extraction method. OD260 and OD280 values for the extracted DNA were determined with a Hitachi Model U-2001 UV/Vis spectrophotometer (Hitachi Instruments Inc., San Jose, CA), and the OD260/OD280 ratio was calculated for the quantification of purity and concentration. K562 (Promega, Madison, WI) and 9947A (PE Applied Biosystems, Foster City, CA) standard DNAs were used as positive controls. Tris–EDTA solution (10 mmol/L Tris–HCl, 0.1 mmol/L EDTA, pH 7.5; Sigma, St. Louis, MO), which was used to dissolve DNA, was also used as a negative control during the PCR and sequencing procedures.

PCR Amplification

Sequence-specific primers were designed with the software provided by Pyrosequencing AB (Biotage, Uppsala, Sweden; http://www.biotage.com) and are listed in Table 1. One primer for each SNP was biotinylated at the 5′-end to allow later single-strand immobilization. A third sequencing primer was used for the sequencing procedures, and when annealed to the template, its 3′ end was adjacent to the 5′-end of the sequence to be analyzed (see Table 1). All primers were synthesized by Shanghai Sangon Biological Engineering & Services (Shanghai, China).

Single PCR for each SNP was performed in a 50 μL reaction volume on an ABI 9700 thermal cycler (Perkin-Elmer Applied Biosystems). For each PCR, 10–20 ng of genomic DNA was mixed with 5 μL of 10× PCR buffer (Takara Biotechnology, Dalian, China), 2 U of Hotstart Taq polymerase (Takara Biotechnology), each dNTP at 200 μM, with each forward and reverse primer at 0.20 μM, and MgCl$_2$ at 1.5 mM.

All SNPs were amplified via the same PCR procedure, which included an initial denaturation at 95°C for 3 min followed by 45 cycles of 94°C for 20 sec, 60°C for 20 sec, and 72°C for 30 sec. The final extension was carried out at 72°C for 5 min. The products were then kept at 4°C before being used for agarose gel electrophoresis.

Pyrosequencing

A Pyrosequencing PSQ96MA instrument (Biotage) was used for the real-time sequencing of PCR products and SNP analysis. Pyrosequencing was performed according to the manufacturer's instructions. For each genotyping procedure, single strand DNA (ssDNA) was isolated from the PCR using the Pyrosequencing Vacuum Prep Workstation (Biotage). Briefly, 4 μL Streptavidin Sepharose high-performance beads (Amersham Biosciences, Piscataway, NJ) were dissolved in 36 μL Binding-Washing buffer (10 mM Tris–HCl, 2 M NaCl, 1 mM EDTA, 0.1% Tween, pH 7.6; Sigma, Stockholm, Sweden). This mixture was added to 40 μL of biotinylated PCR products and agitated at room temperature for 5 min (>100×g); then, the beads with captured DNA were washed once with 70% ethanol (Sigma), transferred to 0.2 M NaOH (Sigma), and flushed once with washing buffer (10 mM Tris-acetate, 5 mM magnesium acetate, pH 7.6; Sigma), with each step lasting 15 sec. Finally, the beads were placed into a 96-well plate containing 45 μL annealing buffer (10 mM Tris-acetate, 5 mM magnesium acetate, pH 7.6; Sigma), and the specific sequencing primer for each SNP (see Table 1). Annealing of the ssDNA to the sequencing primer was accomplished by heating the samples at

TABLE 1—*Identification numbers (rs#) and related sequences for the pyrosequencing analysis of the six single nucleotide polymorphisms.*

Rs#	Chromosome		Related Sequences	Product Length, bp
rs6427658	1	Forward primer	BIO-CCAGTTTGGTCCCTAGAGTAAACA	125
		Reverse primer	TCAAGAGACTGGCCCAGATG	
		Sequencing primer	CTCCCTCCAGTACCC	
		Sequence to be analyzed	TAC A/G TCTACT ACAGCAATGC	
		Dispensing order	ESGTACGAGTC	
rs4607417	6	Forward primer	BIO-GGCCCAAGGAAAGACACCA	238
		Reverse primer	ACCTATCTGACTCCTGCTGACC	
		Sequencing primer	GCCGTGCTTCATTTACT	
		Sequence to be analyzed	A/G TCGGGCTTG GTGTCTTTCC TTGGGC	
		Dispensing order	ESCGAGTCGCT	
rs4722616	7	Forward primer	GGTTCACAGAATTGGCTGAGG	168
		Reverse primer	BIO-CCAGGCACTGTGGCATTTT	
		Sequencing primer	AGAGGTCTTCTGGAAGG	
		Sequence to be analyzed	C A/G TCCATCAG TAATGGATGA	
		Dispensing order	ESTCGAGTCAT	
rs7907658	10	Forward primer	BIO-GCACAGCTGCAAGGAAGACT	95
		Reverse primer	CCTCCCTTTCTCGGTACTAGTGC	
		Sequencing primer	AGTGCCCTCGATTTT	
		Sequence to be analyzed	T G/T GTTGTGGA TATGGGTACT	
		Dispensing order	ESCTGCTGTGA	
rs2369522	14	Forward primer	TCATGAGGGAAAACAGGAGTGTC	112
		Reverse primer	BIO-AACAAGCTTTCTCCAAGGGATAAC	
		Sequencing primer	GAGGATGCAAATTTCAA	
		Sequence to be analyzed	A/C TATCTGTTA TCCCTTGGAG AAAGCT	
		Dispensing order	ESGACGTATCT	
rs749305	15	Forward primer	TTCCATGCTGGCTGAAGTAATTTA	101
		Reverse primer	BIO-GCTGGCACCACCTATTGAAAT	
		Sequencing primer	TGTGTTCCTGGACCACT	
		Sequence to be analyzed	A A/G CCACAATA TTTCAATAGG	
		Dispensing order	ESTAGTCACAT	

80°C for 2 min and allowing them to cool to room temperature. The sample plate was then transferred to the PSQ96MA instrument for real-time dispensing and sequencing.

To facilitate the detection of the six known SNP sites, the order of nucleotide dispensations was set according to the specific dispensing protocol including nine dispensations for each SNP. The dispensation order for each genotyping reaction is presented in Table 1.

During the dispensation process, four enzymes were used: DNA polymerase, ATP sulfurylase, luciferase, and apyrase. The reaction substrates were adenosine-5′-phosphosulfate (APS) and luciferin. After these ingredients were added together into the reaction wells, one of four dNTPs (dATP S, dTTP, dCTP, dGTP) was dispensed into the wells containing the ssDNA-sequencing primer hybrids. The dispensing pressure was 650 mbar with 9 msec of open time and a 65 sec cycle time. DNA polymerase incorporated the dispensed dNTP into the 3′-end of the sequencing primer that was pre-annealed to the template ssDNA. Immediately after this single-base extension, a pyrophosphate (PPi) was released. This PPi activated ATP sulfurylase, which converted APS to ATP. Luciferase then produced light after consuming the free ATP. The light was detected by a charge coupled device camera and presented as a peak in a pyrogram. Each peak height was proportional to the number of dNTPs incorporated, allowing quantitative measurements. During this process, the unincorporated nucleotides were degraded by the enzyme apyrase in the reaction mixture as they could not bind to the single template strands. After the apyrase degraded the unincorporated dNTP and excess ATP, dNTP was sprayed into the wells again. Sequence data were collected and analyzed with PSQ96MA 2.1 software.

Genotyping Quality Control

Pyrosequencing quality was assessed using 20 randomly selected samples as duplicates for each of the six SNPs and by applying the restriction enzyme TaiI (Fermentas Inc., Hanover, MD) to digest 50 samples for rs6427658 only. The 20 DNA samples were renamed to avoid recognition by the pyrosequencing operators, and pyrosequencing was carried out again for this batch of DNA. At the same time, rs6427658 PCR products of 50 samples were digested with the TaiI restriction enzyme (which recognizes ACGT) to confirm the pyrosequencing results. The rs6427658 PCR products digested by the TaiI restriction enzyme and were detected by agarose gel electrophoresis. For TaiI digestion, 10 µL of PCR product (c. 0.1–0.5 µg of DNA), 17 µL nuclease-free water, 2 µL of TaiI (10 U/µL), and 2 µL of 10× Buffer R (10 mM Tris–HCl, pH 8.5, 10 mM MgCl$_2$, 100 mM KCl, 0.1 mg/mL BSA) were mixed gently and spun briefly to collect the sample at the bottom of the tube. The samples were incubated at 65°C for 1 h, and digestion was stopped via the addition of 1.2 µL 0.5 M EDTA (pH 8.0) to achieve a 20 mM final concentration. After a thorough mixing, 6× loading dye was added and the digested products were loaded onto an 8% polyacrylamide gel for electrophoresis and ultraviolet photography.

Statistical Analysis

The gene count method was used to calculate the observed allele frequencies and the observed genotypes in the population data for each of the six SNP loci. Subsequently, genetic data and forensic parameters including the Hardy–Weinberg equilibrium probability (HWP), heterozygosity, F_{st}, PIC, power of discrimination (PD), matching probability, power of paternity exclusion, typical paternity index, and subpopulation inbreeding coefficient (F_{is}) of each SNP among Han Chinese were calculated using Genepop 4.0 and Powerstats Software V12 (Promega) (21) as we reported previously (22). Chi-square tests were also carried out to compare the difference in allele distributions between our data with a large sample size ($n \leq 144$ in Han population) and the HapMap records with a

small sample size ($n \leq 60$ in CEU, HCB, JPT, and YRI). Linkage disequilibrium analysis for the six selected SNPs was performed with the online calculating tool *SNPSTATS* (http://bioinfo.iconcologia.net/snpstats/start.htm).

Results

A Whole Genome Screen for Highly Informative SNPs

Based on the HapMap bulk data, r21a, and the NCBI SNP database, Build125, we completed a thorough scan of 3,905,819 nonredundant genotyped SNPs across the whole genome. We found 3047 SNPs that were shared by all four populations (CEU, HCB, JPT, and YRI) when MAF ≥ 0.44, 1439 SNPs when MAF ≥ 0.45, 507 when MAF ≥ 0.46, 175 when MAF ≥ 0.47, 30 when MAF ≥ 0.48, 2 when MAF ≥ 0.49, and 0 when MAF $= 0.50$. Detailed information about such shared SNPs and the total number of genomic SNPs are listed for each chromosome in Table 2.

Identification numbers (rs#) for these 1439 shared SNPs with a MAF ≥ 0.45 are provided in the Electronic Supplementary Materials (publicly available at http://leg-med.blogspot.com/ or http://legal-med.blog.sohu.com/). For example, rs4722616 has two alleles, G and A, in all four populations; the allele frequency for G/A in CEU is 0.517/0.483, in HCB, it is 0.467/0.533, in JPT, it is 0.500/0.500, and in YRI, it is 0.483/0.517. The frequency distribution schemes recorded in the HapMap database (http://www.hapmap.org) are given in Fig. 1. We found that Chromosome 2 had the most abundant MAF ≥ 0.45 SNPs; 117 of them were shared by the four populations. In contrast, Chromosome Y had no such SNPs. The average heterozygosity of the 1439 identified SNPs was 0.500 and the average F_{st} was -0.0061799. Detailed heterozygosity and F_{st} data for each of the 1439 SNPs are also publicly available at http://leg-med.blogspot.com/ or http://legal-med.blog.sohu.com/.

FIG. 1—*The frequency distribution scheme record for rs4722616 in the HapMap/NCBI single nucleotide polymorphism (SNP) databases (selected from http://www.hapmap.org or http://www.ncbi.nlm.nih.gov/snp). This is one of the 1439 SNPs we have identified from the whole genome SNPs with Minor Allele Frequency ≥ 0.45 among the four populations unanimously: CEU, CHB, JPT, and YRI. The G/A allele frequency ratio of rs4722616 in CEU is 0.517/0.483, CHB 0.467/0.533, JPT 0.500/0.500 and YRI 0.483/0.517, respectively. CEU, U.S. residents with northern and western European ancestry at the Centre d'Etude du polymorphisme humain; CHB, Han Chinese in Beijing; JPT, Japanese from Tokyo; YRI, Yorubans of Ibadan, Nigeria.*

Compared to other reports, our results showed that three of these 1439 SNPs (rs447818, rs985492, and rs1019029) were also included in the 40 SNPs recommended by Pakstis et al. (10). None of the 52 SNPs recommended by Sanchez et al. (11), the 24 SNPs recommended by Lee et al. (13), or the 70 SNPs recommended by Vallone et al. (14) were found in our set of 1439 SNPs.

The self-edited VBA programs used for this screening are also listed in the Electronic Supplementary Materials at http://leg-med.blogspot.com/ or http://legal-med.blog.sohu.com/.

Population Genetics Data of Six Tested SNPs in a Large Sample of Han Chinese ($n > 140$)

The pyrograms were clear and self-explanatory. A single high peak appeared if the sample was a homozygote, while two moderately sized peaks appeared if the sample was a heterozygote. Several sequencing pyrograms are shown in Fig. 2.

The utility of the combined set of these six SNPs was evaluated to estimate their potential for human identification and kinship analysis. The allele and genotype frequencies for each SNP are shown in Table 3. All tested SNPs showed relatively even allelic distributions ranging from 0.399/0.601 to 0.493/0.507 among Han Chinese. Fisher's exact test showed that those markers were all in Hardy–Weinberg equilibrium ($p > 0.05$) (data shown in Table 3). F_{is} values, which were used to estimate the variation within populations, were calculated and listed in Table 4.

Duplicate results were compared for 20 samples and no discrepancies were found between duplicates in any of the 20 cases. The TaiI digestion results also indicated that pyrosequencing-confirmed AA homozygotes for rs6427658 were not digested and the original 125 bp PCR bands were uncleaved, while pyrosequencing-confirmed GG homozygotes for rs6427658 were completely digested, resulting in two bands, at 45 and 80 bp, and the absence of the intact 125 bp band. Pyrosequencing-confirmed heterozygotes for rs6427658 were found to produce three bands: 125, 80, and 45 bp. No difference was found between the pyrosequencing results and the TaiI digestion results in any sample.

Linkage disequilibrium analysis showed that the maximal value of D' or r statistics among these six SNPs was 0.204, suggesting that no linkage disequilibrium existed when these SNPs were used in a combined panel.

Statistical Analyses Between Our Data and the HapMap Records

The chi-square tests of allele distribution comparative analysis between our data and the HapMap records showed no significant
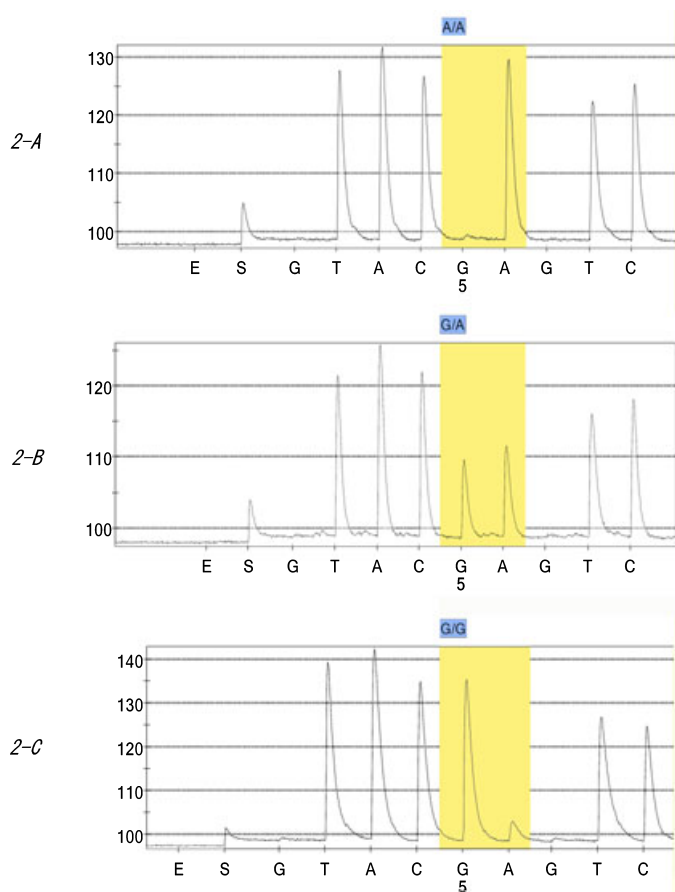
TABLE 2—*List of single nucleotide polymorphisms (SNPs) shared by the four populations with different minor allele frequency (MAF) cutoffs and on different chromosomes (based on the HapMap bulk data public release r21a).*

	MAF ≥ 0.44	≥ 0.45	≥ 0.46	≥ 0.47	≥ 0.48	≥ 0.49	Total SNPs Screened
Chr1	185	60	18	3	1	0	302,414
Chr2	224	117	38	11	1	0	320,297
Chr3	188	77	19	9	1	0	250,156
Chr4	211	90	35	14	2	0	238,977
Chr5	205	76	16	3	0	0	243,347
Chr6	235	111	42	15	1	0	266,915
Chr7	176	76	22	10	4	0	208,948
Chr8	168	86	44	16	2	2	212,265
Chr9	153	91	23	11	1	0	179,317
Chr10	158	87	28	6	1	0	207,152
Chr11	104	38	10	3	2	0	200,670
Chr12	180	105	60	19	4	0	189,348
Chr13	114	19	8	3	2	0	154,880
Chr14	95	52	20	11	0	0	121,310
Chr15	128	69	28	10	1	0	105,022
Chr16	95	42	13	1	0	0	106,996
Chr17	79	49	12	5	0	0	87,411
Chr18	63	38	16	5	2	0	117,678
Chr19	64	43	11	5	0	0	55,195
Chr20	94	48	16	3	1	0	117,982
Chr21	56	30	15	7	3	0	50,154
Chr22	28	13	3	2	0	0	55,484
ChrX	44	22	10	3	1	0	113,837
ChrY	0	0	0	0	0	0	67
Sum	3047	1439	507	175	30	2	3,905,819*

*Average number of nonredundant genomic SNPs in the HapMap bulk data public release r21a.

FIG. 2—*The pyrosequencing profiles of three samples for rs6427658 genotyping. The dispensing order was 5′-ESG TAC GAG TC, which was the same order of the sequence to be analyzed except the E (enzyme) S (substrate) (see Table 2). The underlined "A" was used to detect the allele "A" and the underlined "G" was used to detect the allele "G." For an AA homozygote, there will be no "G"-containing strands in their PCR products, thus a single signal of A-base appears after GA dispensing, as shown by Fig. 2A (Sample A94). On the contrary, there will be no "A"-containing strands in the PCR products of GG homozygote and a single signal of G-base appears after GA dispensing as shown by Fig. 2C (Sample K562). The pyrogram of an AG heterozygote was shown in Fig. 2B (Sample A93), containing both A and G peaks with almost one half of the heights of their neighboring peaks. The genotypes, which were shown in the upper part of the profiles, were automatically generated by the PSQ96MA 2.1 software after the detection.*

difference between our data and the CEU, HCB, JPT, and YRI data at all loci ($p > 0.05$), with the exception of rs4607417 between our data and the JPT data ($0.01 < p < 0.05$), and rs4722616 between our data and the CEU data ($0.01 < p < 0.05$).

The genotype distribution comparative analysis showed the same results as above, finding significant difference at rs4607417 between the Han Chinese and the JPT, and at rs4722616 between the Han Chinese and the CEU.

Discussion

One of the main reasons that STRs have been utilized successfully and globally is that the CODIS STRs are highly informative across various populations worldwide (2,3). Forensic SNP typing has advantages over STRs, if any, in certain situations, such as in cases involving highly degraded biological samples with substantially fragmented DNA. However, it remains unclear whether a core SNP panel can be as universally informative as the CODIS STRs.

Ideal SNPs for individual identification should have two primary characteristics: high heterozygosity (i.e., c. 50% heterozygosity for a biallelic system) and the lowest F_{st} values among various populations. When fitting these criteria as few as 50–100 autosomal loci will be able to reach high levels of PD (23), and fewer reference population datasets will be required for statistical assessments for forensic casework (9–15).

So far, many research groups have endeavored to find such universal SNPs (9–15); the work of Pakstis et al. (10), Sanchez et al. (11), and others have identified valuable SNPs for many populations around the world. Most of their recommendations have good allele frequency distributions among various populations, but some of these SNPs have an extremely low MAF in one or more major populations. For instance, rs560681, recommended by Pakstis et al., has an allele frequency distribution of 0.754/0.246 among Europeans ($n = 57$, http://www.ncbi.nlm.nih.gov/snp, accessed on March 8, 2010) (10) and rs1335873, recommended by Sanchez et al., has an allele frequency distribution of 0.949/0.051 among Sub-Saharan Africans ($n = 59$, http://www.ncbi.nlm.nih.gov/snp, accessed on March 8, 2010) (11). Such uneven frequency distributions will ultimately lead to low heterozygosity, high F_{st} values, low discriminative power among the world major ethnic groups, and reduced utility for worldwide applications.

These uneven distributions may be because of a number of factors, including the smaller sizes of datasets available for early analyses, frequency variations of originally promising SNPs in recently updated samples, or insufficiently broad SNP selection procedures. For these reasons, we have written several computer programs and used them to perform a thorough screening of the c. 4,000,000 whole genome SNPs fitting into the category $0.44 \leq MAF \leq 0.50$ across all four major world populations (CEU, HCB, JPT, and YRI). Our work has shown that, even when restricted to SNPs with a MAF between 0.45 and 0.50, the 1439 qualifying SNPs provide a much larger sample than the maximum required for forensic applications, calculated to be 50–100 for ordinary human

TABLE 3—*Population genetics data for the six single nucleotide polymorphisms in 144 Han Chinese.*

| Rs# | Allele (1,2) | N* | Allele Frequency | | Genotype Frequency | | | HWP | Het | MP | PD | PIC | PPE | TPI | Control DNA | |
			1	2	1, 1	1, 2	2, 2								9947A	K562
rs6427658	A, G	144	0.563	0.437	0.326	0.472	0.201	0.480	0.472	0.37	0.63	0.37	0.164	0.95	A/G	G/G
rs4607417	A, G	143	0.399	0.601	0.161	0.476	0.364	0.872	0.476	0.384	0.616	0.36	0.167	0.95	A/G	A/G
rs4722616	A, G	141	0.592	0.408	0.340	0.504	0.156	0.508	0.504	0.394	0.606	0.37	0.191	1.01	A/G	A/G
rs7907658	T, G	141	0.557	0.443	0.298	0.518	0.184	0.434	0.518	0.391	0.609	0.37	0.203	1.04	G/G	G/G
rs2369522	A, C	144	0.531	0.469	0.271	0.521	0.208	0.457	0.521	0.388	0.612	0.37	0.206	1.04	A/C	C/C
rs749305	A, G	143	0.493	0.507	0.231	0.525	0.245	0.423	0.524	0.388	0.612	0.37	0.21	1.05	A/G	A/G

*N: sample numbers with successful pyrosequencing results; HWP, Hardy–Weinberg equilibrium probability; Het, observed heterozygosity; MP, matching probability; PD, power of discrimination; PIC, polymorphic information content; PPE, power of paternity exclusion; TPI, typical paternity index.

TABLE 4—F_{is} values of our samples and corresponding Han Chinese in Beijing (HCB) records in the HapMap database.

	F_{is} Values	
Rs#	Hans (our samples, $n \le 144$)	HCB (HapMap records, $n = 45$)
rs6427658	0.0440	−0.0602
rs4607417	0.0117	−0.1497
rs4722616	−0.0390	0.0291
rs7907658	−0.0454	−0.1451
rs2369522	−0.0423	0.1598
rs749305	−0.0457	0.1472

A small absolute value for F_{is} usually indicates a small average heterozygote deficiency/excess within a population.

identification and parentage testing by Chakraborty et al. (23). The large pool of SNPs identified here will offer plenty of opportunities to select a panel of SNPs with ideal Hardy–Weinberg equilibrium, appropriate GC content, low linkage, and no relation to disease genes, as well as high heterozygosity, low F_{st}, and high discrimination power across populations; it will also allow for variation in MAF as sample sizes increase.

In our work, we have focused on SNPs with $0.45 \le$ MAF for the genome-wide screening. Statistics and calculated results showed that this criterion has two main benefits: first, it is able to ensure that selected SNPs have the highest possible heterozygosity and the lowest possible F_{st} values across populations; second, it helps us to ensure that the number of selected SNPs does not exceed 3000, which could result in theoretical linkage (19).

Using a MAF ≥ 0.45 presents a simple way to achieve the goal of high heterozygosity and low F_{st} values across populations (7,9,10). For a biallelic SNP, assuming that the allele frequencies in a given population are p and q, the expected heterozygosity (H_{exp}) is calculated as follows:

$$H_{exp} = 1 - \sum \left(p^2 + q^2 \right) \qquad (1)$$

H_{exp} will reach its greatest value when $p = q = 0.50$ and will reach its lowest value when $p = 0.01$ or $q = 0.01$. Hence, our selection criterion ensures that the frequency of one allele is between 0.45 and 0.50 and the frequency of the other allele is between 0.55 and 0.50, keeping H_{exp} values at the highest possible level when $p = q = 0.50$ is not available (Table 3).

Calculations have demonstrated that a MAF ≥ 0.45 across various populations will also select target SNPs with low F_{st} values. According to the well-known work of Weir and Nei (24,25), the F_{st} of the four HapMap populations (CEU, HCB, JPT, and YRI) is calculated as follows:

$$F_{st} = \frac{H_T - H_S}{H_T} = \frac{\left[1 - \sum \left(\bar{p}^2 + \bar{q}^2\right)\right] - \frac{\sum_{i=1}^{4} H_{exp_i} n_i}{\sum_{i=1}^{4} n_i}}{1 - \sum \left(\bar{p}^2 + \bar{q}^2\right)} \qquad (2)$$

$$\bar{p} = \frac{\sum_{i=1}^{4} p_i n_i}{\sum_{i=1}^{4} n_i} \qquad (3)$$

$$\bar{q} = \frac{\sum_{i=1}^{4} q_i n_i}{\sum_{i=1}^{4} n_i} \qquad (4)$$

Here, p and q represent the two allele frequencies, n is the sample size of each population, H_S is the expected heterozygosity

in subpopulations, H_T is the expected heterozygosity for the overall (total) population, and H_{exp} is the expected heterozygosity. These equations show that F_{st} is strongly related to the allele frequency values p and q, and final calculation results showed that our selection of SNPs with a MAF between 0.45 and 0.50 in all four populations reduced F_{st} values to the lowest possible levels, occasionally even achieving small negative values, such as −0.006483 for rs749305 in HapMap records (http://leg-med.blogspot.com/ or http://legal-med.blog.sohu.com/). According to Nei, F_{st} is the correlation between two gametes drawn at random from each subpopulation and should never be negative (26), but it is possible in principle for the F_{st} parameter to have a small negative value, when the probability of identity of genes copies within populations is smaller than the probability between populations. This could perhaps occur under strong balancing selection.

The second reason to use MAF ≥ 0.45 is based on our estimation of the maximal number of SNPs of the whole genome-genotyped SNPs (HapMap r21a, or NCBI SNP database b125) that could be used in a single forensic test. Because there are 3 billion base pairs in total and the minimum distance between two SNPs required to avoid linkage is 1 Mb (19), the maximum number of SNPs that can be used for one single detection is $3 \times 10^9 / 1 \times 10^6$, or 3000. Our statistics show that there are 3047 SNPs shared among the four tested populations with MAF ≥ 0.44; and thus, linkage effects among these SNPs will be theoretically significant. As more and more SNPs are found across the whole genome, the number of SNPs with MAF ≥ 0.44 will only increase and require even more attention to prevent linkage disequilibrium (although such concerns may be less serious when only 50–100 SNPs will be used in a single test).

It is important to note that we have made our SNP selections based on the assumption that frequency records in the public HapMap and NCBI databases are reliable. In fact, these data records are very repeatable, their sample collection procedures are strict and transparent, their sequencing data are generated from five stable genotyping technologies that gained public acceptance, and their genotyping quality is ensured by the inclusion of duplicate or related samples and by periodic quality checks of common sets of SNPs. In our chi-square tests for the allele distribution comparative analysis, we found no significant difference between our data ($n \le 144$) and the HCB data records in HapMap/NCBI ($n = 45$) at the six experimentally tested SNPs.

However, this does not mean that the HapMap/NCBI database is accurate enough for forensic evaluation and that no further investigation is needed. The HapMap/NCBI records are generated from small sample sizes: 30 adult-and-both-parents trios for CEU, 45 unrelated individuals for HCB, 44 unrelated individuals for JPT, and 30 trios for YRI. It is possible that frequency variations will change as sample sizes increase. For example, our data have shown that the G frequency of rs6427658 in our samples is 0.437, while it is 0.533 in HCB samples of HapMap database and that the T frequency of rs4607417 in our samples is 0.399 but 0.467 in HCB samples. We also found differences in F_{is} between our samples and the HapMap HCB samples, which demonstrated that larger sampling was usually more stable. Therefore, larger sampling investigations should be carried out whenever possible to ensure a more robust dataset for complex forensic applications, even when HapMap data are readily available.

Despite these variations, our experimental results have demonstrated that the six experimentally tested SNPs retained high heterozygosity, a low F_{st}, and a high PD among Han Chinese when sample sizes were greatly increased, even though two of the six

MAFs were changed slightly. This implies that most of the 1439 SNPs identified here will maintain MAF between 0.45 and 0.50 in future investigations with larger sample sizes and that they will be a reliable source for future identity testing SNP developments.

Further, we have demonstrated the utility of pyrosequencing as a highly sensitive SNP typing method. Pyrosequencing is a DNA sequencing technology based on the sequencing-by-synthesis principle (16–18). The technique is based on four-enzyme bioluminescence-based real-time monitoring of DNA synthesis using a cascade that produces a detectable light signal upon nucleotide incorporation. This light signal can be quantitatively connected to the number of bases added. Given its internal controls, quantitative properties, and ease of detection, it is clear that this method is a highly sensitive SNP typing platform (27). Although we did not explore the lower limit of the DNA amount required for a successful pyrosequencing assay in our experiments (we used 10–20 ng of DNA, far more than the *c.* 1 ng required for normal STR profiling), data presented by Harrison et al. clearly demonstrated that as little as 8 pg of DNA can be successfully detected and genotyped with this technology (27). This low threshold makes pyrosequencing an even more valuable forensic SNP typing platform.

In summary, the SNPs identified in this paper will be a useful resource for the forensic community, allowing the investigation of a large number of potential forensic SNPs for a final forensic SNPs panel. Meanwhile, we also offer a small but valuable SNP database that will be an essential reference for SNP-based forensic developments in three world major populations as well as in Han Chinese in the future. Investigators worldwide who are interested in these 1439 SNPs or in the six experimentally validated SNPs will be able to launch large-scale investigations to develop unique SNP panels for human identification, if not a universal panel.

References

1. Brinkmann B, Rand S, Bajanowski T. Forensic identification of urine samples. Int J Legal Med 1992;105:59–61.
2. Urquhart A, Kimpton CP, Downes TJ, Gill P. Variation in short tandem repeat sequences: a survey of twelve microsatellite loci for use as forensic identification markers. Int J Legal Med 1994;107:13–20.
3. Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM. Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. J Forensic Sci 1999;44:1277–86.
4. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, et al. Genome-wide detection and characterization of positive selection in human populations. Nature 2007;449:913–8.
5. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–45.
6. Butler JM, Coble MD, Vallone PM. STRs vs. SNPs: thoughts on the future of forensic DNA testing. Forensic Sci Med Pathol 2007;3:200–5.
7. Budowle B, van Daal A. Forensically relevant SNP classes. BioTechniques 2008;44:603–10.
8. Amorim A, Pereira L. Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs. Forensic Sci Int 2005;150(1):17–21.
9. Butler JM, Budowle B, Gill P, Kidd KK, Phillips C, Schneider PM, et al. Report on ISFG SNP Panel Discussion. Forensic Sci Int Genet Suppl 2008;1(1):471–2.
10. Pakstis AJ, Speed WC, Kidd JR, Kidd KK. Candidate SNPs for a universal individual identification panel. Hum Genet 2007;121:305–17.
11. Sanchez JJ, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M, et al. A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 2006;27:1713–24.
12. Børsting C, Sanchez J, Hansen H, Hansen A, Bruun H, Morling N. Performance of the SNPforID 52 SNP-plex assay in paternity testing. Forensic Sci Int Genet 2008;2(4):292–300.
13. Lee HY, Park MJ, Yoo JE, Chung U, Han GR, Shin KJ. Selection of twenty-four highly informative SNP markers for human identification and paternity analysis in Koreans. Forensic Sci Int 2005;148:107–12.
14. Vallone PM, Decker AE, Butler JM. Allele frequencies for 70 autosomal SNP loci with U.S. Caucasian, African-American, and Hispanic samples. Forensic Sci Int 2005;149:279–86.
15. Shi M, Hou Y, Yan J, Bai R, Yu X. 6 Y-SNP typing of China and Korean samples using primer extension and DHPLC. J Forensic Sci 2007;52:235–6.
16. Ahmadian A, Ehn M, Hober S. Pyrosequencing history, biochemistry and future. Clin Chim Acta 2006;363:83–94.
17. Ronaghi M. Pyrosequencing sheds light on DNA sequencing. Genome Res 2001;11:3–11.
18. Iwamoto K, Bundo M, Kato T. Estimating RNA editing efficiency of five editing sites in the serotonin 2C receptor by pyrosequencing. RNA 2005;11:1596–603.
19. Phillips C, Lareu M, Sanchez J, Brion M, Sobrino B, Morling N, et al. Selecting single nucleotide polymorphisms for forensic applications. Int Congr Ser 2004;1261:18–20.
20. Zeng ZS, Yan HT, Zheng XD, Hu GZ, Chen Y, Ding M. High GC amplification: a comparative study of betaine, DMSO, formamide and glycerol as additives. Life Sci J 2006;3:67–71.
21. Tereba A. Tools for analysis of population statistics. Prof DNA 4 1999;2:14–6. http://www.promega.com/geneticidtools (accessed on March 8, 2010).
22. Zeng ZS, Zheng XD, Zhu YL, Wang ZQ, Xiang ZD, Meng XS, et al. Population genetic data of 15 STR loci in Han population of Henan province (central China). Leg Med (Tokyo) 2007;9:30–2.
23. Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B. The utility of STR loci beyond human identification: implications for the development of new DNA typing systems. Electrophoresis 1999;20:1682–96.
24. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution 1984;38:1358–70.
25. Nei M. Molecular evolutionary genetics. New York, NY: Columbia University Press, 1987; 187–92.
26. Nei M. Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 1973;70:3321–3.
27. Harrison C, Musgrave-Brown E, Bender K, Carracedo A, Morling N, Schneider P, et al. A sensitive issue: pyrosequencing as a valuable forensic SNP typing platform. Int Congr Ser 2006;1288:52–4.

Additional information and reprint requests:
Zhaoshu Zeng, Ph.D. or Ziming Dong, M.D., Ph.D.
Center of Forensic Identification
School of Basic Medical Sciences
Zhengzhou University
100 Science Road
Zhengzhou, Henan 450001
China.
E-mail: zs_zeng@yahoo.com or dongzm@zzu.edu.cn

PAPER

CRIMINALISTICS

*Helen W. Kreuzer,[1] Ph.D.; Jon H. Wahl,[1] Ph.D.; Candace N. Metoyer,[1,†] Ph.D.;
Heather A. Colburn,[1] Ph.D.; and Karen L. Wahl,[1] Ph.D.*

Detection of Acetone Processing of Castor Bean Mash for Forensic Investigation of Ricin Preparation Methods*

ABSTRACT: Samples containing the toxic castor bean protein ricin have been recently seized in connection with biocriminal activity. Analytical methods that enable investigators to determine how the samples were prepared and to match seized samples to potential source materials are needed. One commonly described crude ricin preparation method is acetone extraction of crushed castor beans. Here, we describe the use of solid-phase microextraction and headspace analysis to determine whether castor beans were processed by acetone extraction. We prepared acetone-extracted castor bean mash, along with controls of unextracted mash and mash extracted with nonacetone organic solvents. Samples of acetone-extracted mash and unextracted mash were stored in closed containers for up to 109 days at both room temperature and −20°C, and in open containers at room temperature for up to 94 days. Acetone-extracted bean mash could consistently be statistically distinguished from controls, even after storage in open containers for 94 days.

KEYWORDS: forensic science, ricin, preparation method, acetone, solid-phase microextraction, headspace analysis, gas chromatography-mass spectrometry

Ricin is a toxic protein produced by the castor plant *Ricinus communis*. It exerts its lethal effect by depurinating an adenine residue within eukaryotic ribosomal RNA, ablating a binding site for an elongation factor, and thereby inhibiting protein synthesis (1). The lethal dose has been estimated to be 5–10 mg/kg body weight, if the toxin is inhaled or injected (2). The median lethal oral dose in mice is approximately 1000 times greater than the inhaled or injected lethal dose, and the lethal oral dose for humans has been estimated to be 1–20 mg ricin per kilogram body weight (3). Ricin was investigated as an offensive weapon by the U.S. military after World War I, and other governments have reportedly explored its use (4). It is classified as a Schedule 1 controlled substance under the Chemical Weapons Convention and a Category B (biological agents∕toxin) substance under the Biological Toxins Weapons Convention. Ricin is also included as a Category B select agent set forth by the U.S. Department of Health and Human Services (42 CFR Part 73).

Much of the concern over ricin as a potential biological threat agent stems from the widespread availability of the source material, the seeds of the castor plant. Castor is cultivated as an oil-seed crop in tropical and subtropical regions throughout the world, and the plant has naturalized in warm regions around the world. Castor plants, which have a striking appearance, are also common in horticulture. Castor seeds, commonly referred to as castor beans, are an unregulated agricultural product, but as soon as they are broken open with the intent of obtaining ricin, the resulting material is considered a ricin sample. Thus, demonstrating attempted purification of ricin is an important forensic task.

Numerous procedures for purification and partial purification of ricin from castor beans have been published. Approximately 50% of the weight of a castor seed is oil, which finds numerous uses in the chemical and lubricant industries and has been used as a lamp fuel, purgative, and body ointment since ancient times (5). Ricin comprises about 1.5% of the oil-free meal and, as a water-soluble protein, is not extracted when oil is removed by organic solvent extraction (5). A process for producing large volumes of ultra-pure material patented by the U.S. Army (6), as well as research laboratory procedures based on chromatographic separation, can be found in the scientific literature (7). In addition to these highly technical procedures for preparing purified ricin, however, numerous "kitchen" recipes can be found in books characterized as "anarchist literature" and∕or posted on the internet (8–11). These procedures give rise to ricin-containing preparations of various levels of purity, although all are identified as "ricin" preparation methods, are easy to carry out, and do not require sophisticated equipment. The product of any of these procedures is considered a ricin sample. The most elementary of these procedures call for simply peeling and grinding castor beans into a mash. The next simplest procedure calls for peeling and grinding the beans, then removing oil from the resulting mash, usually by

acetone extraction. Procedures for acetone extraction of ground castor beans are posted on the Internet and found in many of the references cited earlier.

In recent years, ricin-containing samples have been seized primarily in connection with biocriminal activity (12). Analytical methods are needed that enable investigators to determine how the samples were prepared, to match seized samples to potential source materials, and to identify samples that may have been prepared by the same method using the same source materials. We are working to develop procedures to (1) determine whether acetone extraction was used to prepare a ricin sample and (2) as a follow-on, use the impurities that carry over from the acetone into an extracted sample as a signature for further characterizing the acetone and/or matching to specific acetone samples. Here, we describe the development of a screening method to distinguish castor bean mash that was extracted with acetone from bean mash that was not, using recipes from the anarchist literature to produce the samples.

The method employs headspace analysis by solid-phase microextraction (SPME) followed by gas chromatography–mass spectrometry (GCMS) for separation and identification of the compounds in the headspace, a well-recognized method for analyzing volatile organic compounds (13–15). SPME–GCMS analysis has been previously used to detect acetone and other volatiles in the headspace of human saliva (16). In saliva analysis, one-dimensional GC analysis was used. We employed two-dimensional GC, not because it was necessary to detect acetone, but because of our second goal of analyzing trace impurities in the sample headspace and linking them to impurities in the acetone. We have characterized the effects of storage of acetone-extracted and unextracted castor bean mash under different conditions and for varying amounts of time on the ability to detect processing with acetone.

Materials and Methods

Safety Considerations

Ricin is a protein and can be inactivated by a number of methods, including exposure to 10% bleach or 0.6 N NaOH for at least 30 min (17). Solutions of ricin are inactivated by treatment at 80°C for 10 min (17). As a first step in our investigation, we determined that heating castor seeds at 100°C for 30 min did not appear to affect the results of acetone analyses (see Results). We therefore used heated beans in all experiments to reduce risk associated with toxin exposure. All preparation of castor mash, whether the seeds had been preheated or not, was conducted in a biosafety cabinet, and solvent extractions were performed in a filtered fume hood. Laboratory personnel wore gloves, lab coats, and safety glasses when preparing castor mash. The mash itself and all materials exposed to castor mash were inactivated by either soaking for a minimum of 1 h in 10% bleach, heat treatment for a minimum of 30 min at 100°C, or autoclaving. The treated material was collected for incineration.

Sample Preparation

Castor beans were purchased from Bouncing Bear Botanicals (Lawrence, KS). Castor mash was prepared from heated and unheated beans according to recipes from the anarchist literature as follows: beans were soaked in 3 N NaOH for 1 h to soften the husk. They were rinsed in water, patted dry, and the husks were removed. The peeled beans were placed in a plastic weigh boat, covered with a sheet of glassine paper, and thoroughly crushed with a pestle. Typically, the resulting mash was separated into two

portions, one to serve as unextracted control mash, and one to be solvent-extracted. Control mash was transferred to headspace vials and stored at the designated temperature (either −20°C or room temperature, as described in the Results section) before any solvent extraction was begun. After removal of the control sample, the portion to be extracted was transferred to a glass flask and covered with approximately 10 volumes of solvent. A stir bar was added. The flask was covered with aluminum foil, sealed with Parafilm® (Pechiney Plastic Packaging, Menasha, WI), and stirred on a magnetic stir plate for 3 days. The extracted mash was recovered by filtering the mixture through two nested coffee filters, then allowing the mash to air-dry in the fume hood. The effect of drying the mash for different lengths of time was investigated, because the procedures for acetone extraction of ricin in the anarchist literature did not specify a single drying time (typically, one was instructed to allow the extracted mash to "dry"). We allowed extracted bean mash to dry in the fume hood from 1 to 7 days, as described in the Results section. The dry mash was transferred to 10-mL headspace vials and stored as described.

Analytical Instrumentation and Procedure

Analysis of the headspace in the sample vials was carried out by SPME analysis on a Leco Pegasus 4D GC × GC-TOFMS system (LECO Corp., St. Joseph, MI) equipped with a Gerstel multipurpose sampler (MPS2) for automated SPME analysis (Gerstel Inc., Baltimore, MD) including a liquid-cooled sample tray and fiber bakeout station. While the analytical capability of the instrument far exceeded that required for this procedure, we plan in the future to analyze impurities in different batches of solvents and resulting extracts as potential signatures for sample matching, and two-dimensional separation will be important in that phase of the project. The analyses described in this report could be carried out on any GCMS instrument equipped for SPME headspace analysis.

Prior to analysis, the headspace vials containing the samples were opened, and 1.0 mL 18 MΩ water and 0.5 g NaCl were added. Samples were analyzed in blocks of six to eight samples with the order determined by a random number generator. An empty vial was analyzed before and after each sample as a check for potential carry-over. The sample vials were placed in a liquid-cooled (4°C) tray prior to the run. To equilibrate and ready samples for analysis, a vial was moved to a heated (40°C) agitator incubator and agitated in cycles of 45 sec on and 15 sec off for 15 min. Prior to insertion into the sample vial, the entire SPME fiber was baked at 275°C for 4 min. Following the 15-min sample preheating step, the SPME fiber was inserted into the vial and 3 mm of the fiber was inserted exposed for 60 sec. For GC injection, the entire fiber was exposed in the heated injector (250°C), which was operated in pulsed splitless mode. After 60 sec desorption, the split vent was opened and the injector was purged for 120 sec before the fiber was withdrawn. The entire SPME fiber was then exposed and further baked at 275°C for 8 min in the fiber bakeout station. Column 1 was a 30 m, 250 μm ID, 0.25 μm df SolGel wax column (SGE, Inc., Austin, TX). The primary oven had the following temperature program: initial temperature of 35°C with a hold time of 10 min, then a 5°C/min ramp to 250°C, with a 5 min hold time. Column 2 was a 2 m, 100 μm ID, 0.40 μm df-1701 column (Quadrex Corporation, Woodbridge, CT). The secondary oven had the following temperature program: initial temperature of 45°C and a 10 min hold time, then a 5°C/min ramp to 260°C with a 5 min hold time. A 5-sec modulation period was utilized. The transfer line to the mass spectrometer was maintained at a temperature of 250°C. Ultra

TABLE 1—*Log-transformed* m/z = 43 *peak volume from the acetone peak in control water/acetone solutions prepared and analyzed in six separate runs over the course of approximately 3 months. Note that there is no significant difference in the values from 0 M acetone, 6.08E-8 M, and 6.08E-6 M acetone.*

	Acetone Concentration (M)				
Solution concentration	0	6.08E-8	6.08E-6	6.08E-4	6.08E-2
Headspace concentration*	0	1.4E-9	3.2E-8	3.1E-6	3.1E-4
Analytical run	Log peak volume				
1	4.13	4.22	4.61	6.29	7.58
2	4.32	4.36	4.54	6.09	7.63
3	4.49	4.42	4.61	6.18	7.46
4	4.67	4.62	4.74	6.30	7.54
5	4.85	4.76	4.82	6.03	7.59
6	4.52	4.55	4.74	6.37	7.91
Mean	4.50	4.49	4.68	6.21	7.62
Standard deviation	0.25	0.19	0.10	0.13	0.16
% Relative standard deviation	5.6%	4.3%	2.2%	2.1%	2.0%

*Headspace concentration was estimated via Henry's law using standard atmospheric pressure and $T = 40°C$ with the assumptions that the acetone–water solutions were ideal solutions and headspace volume = 9.0 mL.

high purity helium carrier gas flow rate was 1 mL/min for the run duration. Mass scan range was 35–400 *m/z*, at 100 spectra/sec.

Results

Instrument Response

To assess instrument response over time, solutions of acetone in 18 MΩ water at the following concentrations were run along with sample sets over a period of approximately 3 months: 6.8×10^{-2}, 6.8×10^{-4}, 6.8×10^{-6}, 6.8×10^{-8}, and 0 M (water blanks). Data are shown in Table 1. For each solution, the acetone signal intensity (reported as signal intensity equal to the volume of the acetone zone utilizing *m/z* 43 from the acetone peak) remained relatively constant, which indicated that the instrument was reporting similar acetone signal intensities when solutions of the same concentration were analyzed. For each concentration, a linear trend line was used to model the response (acetone signal intensity) versus time (the date when the reading was taken). The slopes were not statistically significant at the family-wide level of confidence $\alpha = 0.05$. Therefore, we could conclude that from day to day, the instrument was providing similar responses for comparable inputs, and that no instrument response correction was necessary. Table 1 also shows the estimated headspace concentration of acetone for the different acetone/water solutions, calculated using Henry's law, with the assumption that the solutions were ideal solutions.

Preheating Castor Beans

CDC guidelines include heat treatment as a method to inactivate ricin (www.cdc.gov/od/ohs/biosfty/bmbl5/bmb/5toc.htm). In an effort to minimize hazards associated with working with active ricin, we investigated whether heated castor beans could be substituted for raw beans in the preparations. We heated two batches of castor beans for 30 min at 100°C on two different days approximately 1 month apart and prepared mash from these batches as well as unheated beans on the same day. During the preparation, headspace vials containing 1 mL water were left open in the

biological safety cabinet, then closed and used as control blanks. We also prepared acetone-extracted mash from raw and heated beans. Here, we allowed portions of the extracted mash to sit on filter paper in the hood for 1 and 7 days after filtering to investigate whether the air-dry time would significantly affect the intensity of the acetone signal in the headspace. Chromatographic profiles of unextracted mash from raw and heated beans showed no noticeable differences (profiles not shown). The mean acetone signal intensities of each of the three unextracted bean groups (the two batches of heated beans and the raw beans) were significantly greater than that of the control blank group ($p = 0.002$ between raw and control, $p < 0.001$ between heated bean batch #1 and control, and $p = 0.0016$ between heated bean batch #2 and control). Statistical analysis of the $m/z = 43$ signal intensities in the acetone peaks observed in the headspace over unextracted mash of raw and heated beans showed no difference ($p = 0.442$ for batch #1 heated vs. raw beans, $p = 0.854$ for batch #2 heated vs. raw beans). Likewise, the acetone signal in the headspace of acetone-extracted mashes of raw and heated beans showed no significant difference ($p = 0.265$). Further, there was no significant difference in the headspace acetone signal intensities of the extracted mashes left to dry for 1 or 7 days on filter paper before being sealed into headspace vials ($p = 0.197$). Therefore, we concluded that heating the castor beans would not affect our experimental results, and the remaining experiments were all conducted with preheated beans.

Initial Analyses

Figure 1 shows two-dimensional chromatograms that illustrate the selection of analytical conditions for the headspace volatiles in our samples. In our initial analyses of acetone-extracted and unextracted castor bean mash, acetone was detected in both extracted and unextracted bean mash, and the acetone signal from the extracted mash overloaded the detector (Fig. 1A,B). Analytical conditions were adjusted so that they permitted detection of signal from unextracted mash but did not overload the detector in the case of acetone-extracted mash (Fig. 1C). These conditions were used for all subsequent analyses and are described in Materials and Methods. The statistical significance of differences in acetone signal intensity and the effects of storage on that signal are explored in the following experiments.

The two-dimensional chromatogram in Fig. 1A illustrates the complexity of the volatiles' profiles in the extracted mash. Previous analysis of the headspace above raw beans (although not castor beans) have found over 60 (18) and over 100 (19) different compounds using SPME and one-dimensional GC–MS. Figure 1 additionally illustrates that different analytical approaches are required to characterize the minor components such as impurities from acetone in the headspace versus detection of acetone itself.

Differentiating Acetone-extracted Castor Bean Mash from Control Mash

To compare acetone-extracted castor bean mash to mash that was not extracted with acetone, we prepared unextracted castor bean mash and extracted bean mash with three different solvents: laboratory grade acetone, laboratory grade chloroform, and commercial grade mineral spirits purchased from a hardware store. Each solvent extraction was performed at different time periods to minimize the chance of cross-contamination. We followed acetone extraction procedures described in the anarchist literature for these preparations, because these procedures are readily available to the nontechnical public and are easy to carry out. However, the
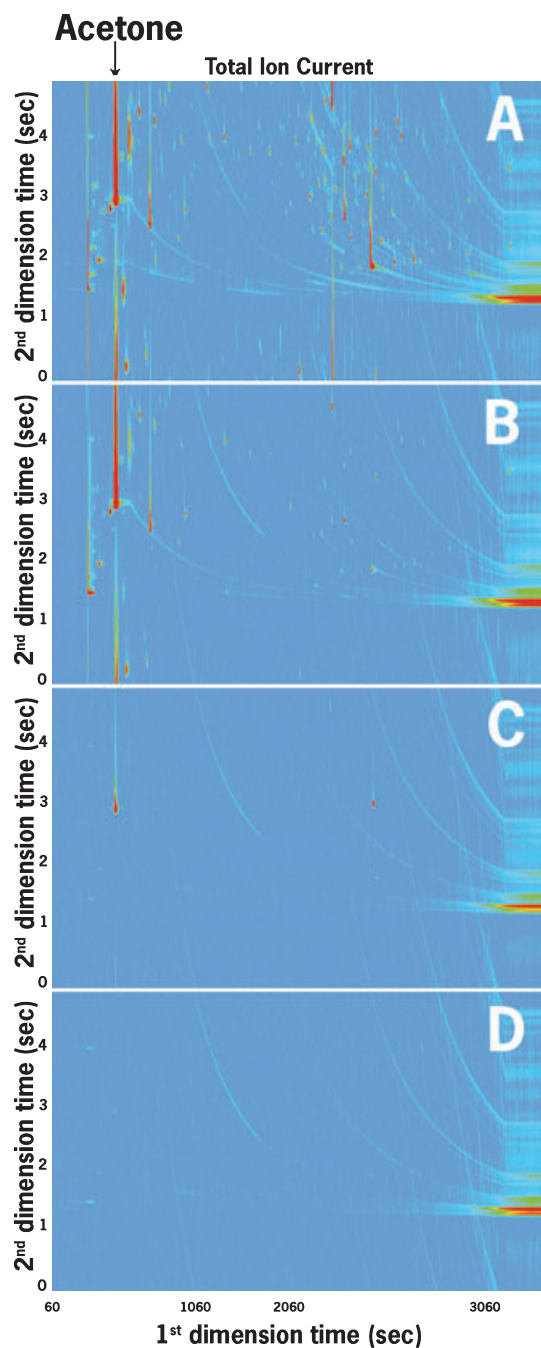
FIG. 1—*Two-dimensional chromatograms of volatiles in headspace of acetone-extracted castor bean mash obtained under different solid-phase microextraction absorption conditions. (A) Full fiber length exposed, 68-min exposure time. (B) Full fiber length, 1-min exposure. (C) 3-mm fiber length exposed, 1-min exposure. (D) Water blank, 3-mm fiber length, 1-min exposure. All signal intensities are normalized to the same scale. Signal intensity (peak volume) increases from aqua to red.*

procedures as published are not as precisely described as those that appear in the scientific literature. In particular, after filtering the acetone-extracted mash, the procedures typically direct the preparer to allow the mash to "dry" without saying how long. The air-dried mash is then identified as "ricin." We therefore assumed that crude ricin prepared by acetone extraction might have been left to dry for varying amounts of time, and so varied the amounts of time we allowed our preparations to air-dry. All of the air-dried samples appeared dry at the time they were placed in vials. The

chloroform-extracted mash was allowed to dry on the filter paper for 4 days before aliquots were placed into tared headspace vials. For the acetone and mineral spirits extractions, bean mash was allowed to dry for up to 7 days on the filter paper, and aliquots of dried mash were placed into tared headspace vials after 1, 4, and 7 days of drying time. Mash samples were stored at −20°C prior to analysis. During acetone and mineral spirits processing, open beakers of water were placed near, but not inside the fume hood to act as background controls. These beakers remained in place throughout processing and drying. At the time the bean mash samples were collected from the filters, 1 mL aliquots of the water were placed in headspace vials and stored along with the mash samples. No water blank was maintained during chloroform extraction because of equipment repairs (not related to this investigation) that occurred in the room during extraction. Duplicate water blanks were set up when water and salt were added to the chloroform-extracted and unextracted mash samples prior to analysis. Figure 2 shows the mean log-transformed acetone peak volumes ($m/z = 43$) from the chloroform-extracted, mineral spirits-extracted, acetone-extracted, and unextracted samples, as well as the accompanying water blanks. The acetone responses in the extracted mash were approximately 100-fold higher than in the unextracted samples, those in the samples extracted with other solvents, or the water blanks.

Statistical analyses showed no significant effect of drying time on the intensity of the acetone signal in either the acetone-extracted mash or the mineral spirits-extracted mash, as was found in the comparison of acetone-extracted mash from heated and unheated beans. The mean signal intensities of the 1-, 4-, and 7-day drying time acetone-extracted mash groups were significantly higher than the intensities of the corresponding mineral spirits-extracted mash groups ($p < 0.001$, $p < 0.001$, and $p = 0.002$, respectively). The mean signal intensity of the acetone-extracted mash group was significantly higher than that of the chloroform-extracted mash group ($p < 0.001$) and the unextracted mash group ($p < 0.001$). The mean intensity of the acetone peak in the headspace of the unextracted mash group was not significantly different from that of the samples extracted with nonacetone solvents ($p = 0.193$ for the mineral spirits-extracted mash group and $p = 0.616$ for the chloroform-extracted mash group).

Effect of Storage in Closed Vials at Room Temperature and −20°C

Acetone is a volatile solvent, and it is highly likely that samples of ricin preparations recovered in the course of an investigation would undergo storage at some point between production and analysis; thus, it was important to investigate the effects of storage on the ability to differentiate acetone-extracted bean mash from unextracted beans. We therefore prepared a batch of unextracted bean mash and placed aliquots of it in tared headspace vials to serve as controls. After the control mash had been stored, we prepared an equal number of castor beans, extracted the resulting mash with acetone, and placed aliquots of the extracted material in tared headspace vials. A beaker of 18 MΩ water was set in the laboratory nearby, but not inside the fume hood when the extraction procedure was begun, and left open until the extracted mash was harvested. At that time, 1 mL aliquots of the water were transferred to headspace vials and stored with the castor mash. In all, 21 samples of unextracted mash, 21 samples of extracted mash, and 21 water blanks were stored in closed headspace vials at room temperature, and an identical set of samples was stored at −20°C for up to 16 weeks. Triplicate sets of samples were pulled at 1, 8, 15, 27, 61, 77, or 109 days for analysis. Figure 3A presents a comparison
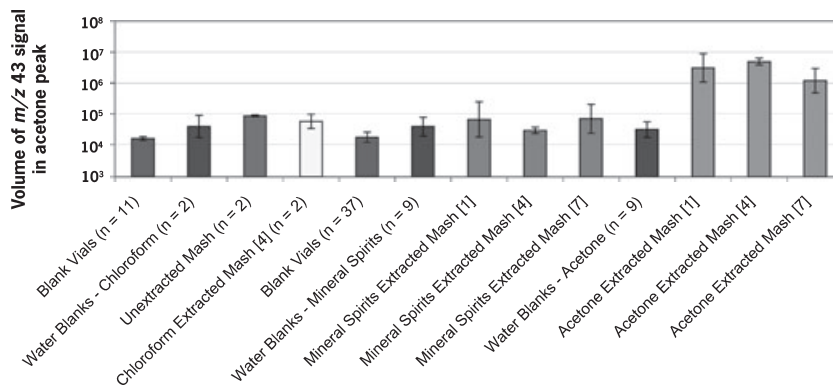
FIG. 2—*Average log-transformed acetone signal intensity for acetone-extracted bean mash and the following controls: unextracted mash, chloroform-extracted mash, and mineral spirits-extracted mash. Samples were prepared in triplicate unless otherwise noted. The number in brackets after the sample identifier indicates the number of days the extracted mash was left to dry prior to being placed in headspace vials. Error bars represent ± one standard error of measurement.*

of the intensity of the $m/z = 43$ signal from acetone in the headspace of samples stored at −20°C for varying lengths of time. Figure 3B presents a similar comparison of samples stored at room temperature. As in Fig. 2, the acetone responses in the extracted mash were approximately 100-fold higher than in the unextracted samples and the water blanks.

A two-factor one-covariate analysis of covariance (ANCOVA) model was used to study both factor effects (*treatment method*: acetone extraction vs. no extraction and *storage temperature*: room temperature vs. −20°C) and regression effects (effects of storage time on signal intensity) (20). Diagnostics were performed to verify that the data met the assumptions of the ANCOVA model. Acetone extraction was shown to have a significant effect on signal size for all the samples examined ($p < 0.001$), with the intensity of the m/z signal in the acetone peak in the headspace of extracted samples significantly larger than the corresponding intensity from unextracted samples. In contrast, the storage temperature did not have a significant effect on the acetone signal intensity ($p = 0.249$). However, an analysis for treatment method × storage temperature

interaction effects demonstrated that extracted and unextracted mashes had different responses to storage temperature ($p = 0.007$). The average signal intensity of acetone-extracted mash was diminished in the samples stored at room temperature compared to that of samples stored at −20°C. In contrast, the average signal intensity of unextracted mash stored at room temperature was greater than that stored at −20°C.

To test whether there was a statistically significant effect of storage time on acetone signal intensity, a point estimate for the regression slope coefficient between the log-transformed acetone signal intensity Y (based on m/z 43 signal intensity in the acetone peak) and (mean-centered) days in storage, x, was computed. The point estimate for the regression slope coefficient between Y and x is $\hat{\gamma} = -0.00026$ (standard error = 0.00049). The test of the null hypothesis H_0: $\gamma = 0$ versus H_a: $\gamma \neq 0$ has $p = 0.600$ and the null hypothesis is not rejected. Thus, for the range of x values that we considered, i.e., for up to 109 days, the number of days in storage had no statistically significant effect on acetone signal intensity at either temperature.
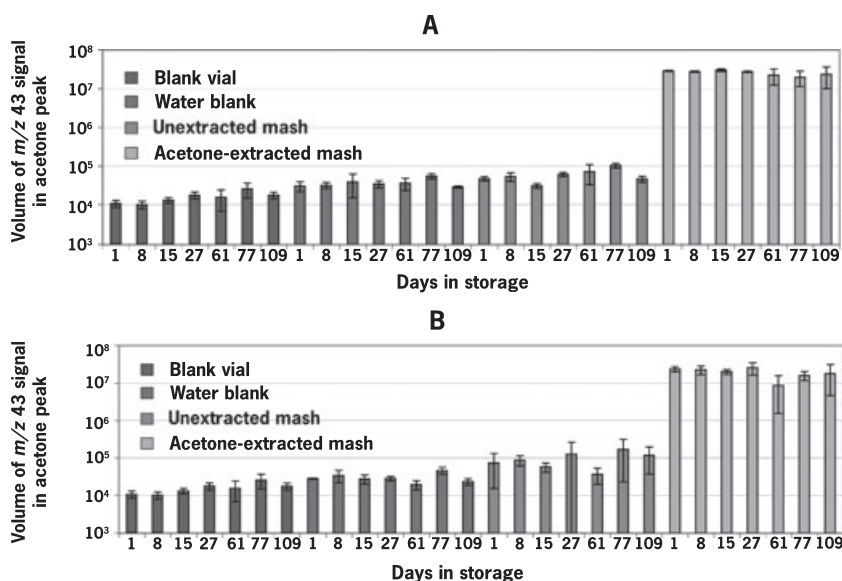
FIG. 3—*Average acetone signal intensity for samples stored 1, 8, 15, 27, 61, 77, or 109 days in closed vials. (A) Samples stored at −20°C. (B) Samples stored at room temperature. Error bars represent ± one standard error of measurement.*
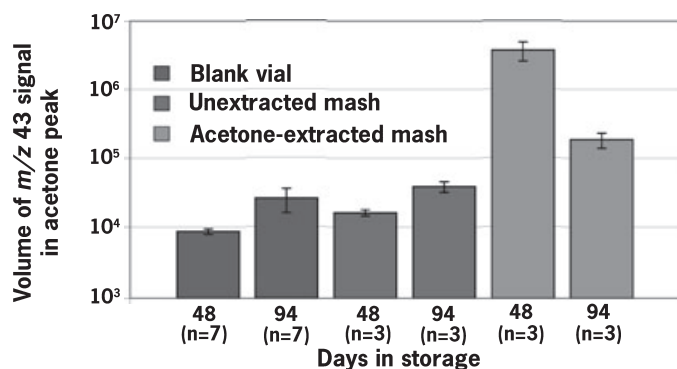
FIG. 4—*Average acetone signal intensity for samples stored unsealed at room temperature for 48 days or 94 days. Error bars represent ± one standard error of measurement.*

TABLE 2—*Mean and variance of* m/z 43 *signal volume in acetone peak in combined classes of samples stored in closed containers.*

Sample Type	Number of Samples	Log-Transformed Sample Mean	Variance
Acetone-extracted mash	50	7.15	0.18
Mineral spirits-extracted mash	9	4.70	0.17
Chloroform-extracted mash	2	4.75	0.054
Unextracted mash	44	4.80	0.068
Water blanks	62	4.51	0.038
Empty vials	179	4.17	0.029

Storage in Open Vials

The data shown in Fig. 3 were obtained from samples stored in closed headspace vials. To examine the effect of storage at room temperature in an open container, six aliquots of the extracted and unextracted mash prepared for the experiment just discussed were transferred to open headspace vials and stored at room temperature in the analysis laboratory, which is in a separate building compared to the preparation laboratory. Three vials of each were analyzed after 48 days, and the remaining three vials were analyzed after 94 days. Figure 4 contains a plot of the average transformed signal intensity at $m/z = 43$ in the acetone peak of the samples' headspace.

To study the differences in the four treatment groups of interest (acetone-extracted [48 days], unextracted [48 days], acetone-extracted [94 days], and unextracted [94 days]), Tukey simultaneous 95% confidence intervals for all pairwise differences in mean log-transformed signal intensity were computed. After both 48 and 94 days, the mean acetone signal intensity of the acetone-extracted mash was significantly higher than that of the corresponding unextracted mash. Additionally, over time, the mean acetone signal intensity of the acetone-extracted mash decreased and the mean acetone signal intensity of the unextracted mash increased.

Therefore, even when left unsealed in a potentially contaminating environment for up to 94 days, the acetone-extracted mash could be distinguished from the unextracted mash. This result is consistent with the finding above that air-drying extracted bean mash for 1 to 7 days on a coffee filter before putting into vials made no significant difference in acetone signal intensity. Because the acetone signal intensity of the acetone-extracted mash tended to decrease and the acetone signal intensity of the unextracted mash tended to increase over time, one would expect that eventually it might become impossible to distinguish extracted from unextracted mash. However, even after 94 days, the mean difference in acetone signal intensity was still statistically significant at the $\alpha = 0.05$ level. It may be of use to conduct additional studies for extended lengths of time to find out if/when the signals from the acetone-extracted mash and unextracted mash become indistinguishable.

Data Summary

None of the variables we investigated in these separate experiments (drying time, storage time, or storage temperature) had a significant impact on the $m/z = 43$ signal intensity in the acetone peak from sample headspace when samples were stored in closed vials. We therefore compared the acetone signal intensity from

broad classes of samples generated over the several-month course of this project and stored in closed vials (Table 2). The sample sizes and sample variances for the six groups of interest varied; thus, Welch's approximate *t*-test was used to study the pair-wise differences in mean response (21).

In all cases, the acetone-extracted bean mash could be distinguished from the mash not processed with acetone. The *p*-values for relevant comparisons follow. The mean acetone signal intensity for the acetone-extracted bean mash was significantly higher than that of the mineral spirits-extracted mash ($p < 0.001$), chloroform-extracted mash ($p = 0.002$; however, this test lacks power because of the small number of chloroform-processed samples), unextracted mash ($p < 0.001$), water blanks ($p < 0.001$), and blank vials ($p < 0.001$). The acetone signal intensity in the unextracted mash was not significantly different from that in the mineral spirits-extracted mash ($p = 0.507$) nor from that in the chloroform-extracted mash ($p = 0.804$; again, this test lacks power because of the small number of chloroform-processed samples).

The signal from unextracted mash was significantly greater than that from the water blanks and the empty vials ($p < 0.001$ in both cases), and the signal from the water blanks was significantly greater than the signal from the empty vials ($p < 0.001$).

All of the unextracted castor bean samples analyzed in these experiments exhibited a small amount of acetone, as did water exposed to laboratory air during the extraction procedures. Many factors could contribute to this signal. For example, acetone is a normal cellular metabolite; therefore, the acetone in the unextracted mash could have originated in the bean. Alternatively, it could have been absorbed from laboratory air, as acetone was routinely used in the laboratory environment, and humans exhale acetone in their breath (16). The analytical techniques we used did not allow us to distinguish between these possibilities. The fact that the signal intensities from unextracted bean mash tended to increase in the samples stored in open vials suggests that castor bean mash can absorb acetone (and possibly other solvents) from the environment. Thus, trace organic solvents in the headspace of castor bean mash samples might serve as indicators of environmental contaminants to which the sample was exposed.

Discussion

We have established conditions for using SPME headspace analysis from castor bean mash or ricin samples to determine whether they were extracted with acetone, and characterized the effects of storage conditions on signal intensity. For safety considerations, the majority of this work was conducted with samples that had been prepared from heated beans so that the ricin in the samples would be inactive; however, preliminary results demonstrated that samples prepared from heated and unheated beans behaved indistinguishably in the assay. In all cases, acetone-extracted castor bean mash could

be distinguished from mash that had not been extracted with acetone, whether it was unextracted or extracted with nonacetone organic solvents, based on the intensity of the $m/z = 43$ signal in the acetone peak. Storage in closed headspace vials for 109 days had no effect on signal intensity in extracted or unextracted bean mash. While storage in open containers in the laboratory resulted in decreasing acetone signals from extracted mash and increasing acetone signals from unextracted mash, the two sample types could be readily distinguished even after 94 days.

Solid-phase microextraction analysis of headspace thus appears to provide a simple and robust screening tool to detect acetone extraction in castor bean mash that was stored in closed containers and is applicable to samples stored in open containers up to 3 months and possibly longer. Our results suggest that ricin samples recovered in the course of a criminal investigation can be stored in tightly closed containers for several months without loss of acetone signal and that acetone-extracted castor bean mash recovered from the field in closed containers is likely to retain the acetone processing signal.

References

1. Endo Y, Tsurugi K. The RNA N-glycosidase activity of ricin A-chain—the characteristics of the enzymatic-activity of ricin A-chain with ribosomes and with ribosomal-RNA. J Biol Chem 1988;263(18):8735–9.
2. Bradberry SM, Dickers KJ, Rice P, Griffiths GD, Vale JA. Ricin poisoning. Toxicol Rev 2003;22(1):65–70.
3. Audi J, Belson M, Patel M, Schier J, Osterloh J. Ricin poisoning: a comprehensive review. JAMA 2005;294(18):2342–51.
4. Kirby R. Ricin toxin: a military history. Army Chemical Review 2004;(PB 3-04-1).
5. Weiss EA. Castor. Oilseed crops. London and New York: Longman, 1983.
6. Craig HL, Alderks OH, Corwin AH, Dieke SH, Karel CL, inventors. Preparation of toxic ricin. United States of America patent 3,060,165; 1962.
7. Simmons BM, Russell JH. A single affinity column step method for the purification of ricin toxin from castor beans (Ricinus communis). Anal Biochem 1985;146(1):206–10.
8. Saxon K. The poor man's James Bond. El Dorado, AR: Desert Publications, 1991.
9. Harber D. Assorted nasties. El Dorado, AR: Desert Publications, 1993.
10. Hutchkinson M. The poisoner's handbook. El Dorado, AR: Desert Publications, 2000.
11. Preisler S. Silent death, 2nd edn. Green Bay, WI: Festering Publications, 1997.
12. Bale JM, Bhattacharjee A, Croddy E, Pilch R. Ricin found in London: an al-Qa'ida connection? Monterey, CA: James Martin Center for Nonproliferation Studies, 2002 [updated 2002 2008; cited], http://cns.miis.edu.
13. Kataoka H. Automated sample preparation using in-tube solid-phase microextraction and its application—a review. Anal Bioanal Chem 2002;373(1–2):31–45.
14. Arthur CL, Pratt K, Motlagh S, Pawliszyn J, Belardi RP. Environmental-analysis of organic-compounds in water using solid-phase micro extraction. J High Resolut Chromatogr 1992;15(11):741–4.
15. Pawliszyn J, Pedersen-Bjergaard S. Analytical microextraction: current status and future trends. J Chromatogr Sci 2006;44(6):291–307.
16. Wang VS, Lu MY. Application of solid-phase microextraction and gas chromatography–mass spectrometry for measuring chemicals in saliva of synthetic leather workers. J Chromatogr B Biomed Sci Appl 2009;877(1–2):24–32.
17. Wannemacher RW, Creasia DA, Hines HB, Thompson WL, Dinterman RE. Toxicity, stability, and inactivation of Ricin (RCA_{60}). Toxicologist 1990;10(1):166.
18. Oomah BD, Liang LSY, Balasubramanian P. Volatile compounds of dry beans (Phaseolus vulgaris L.). Plant Foods Hum Nutr 2007;62(4):177–83.
19. Barra A, Baldovini N, Loiseau AM, Albino L, Lesecq C, Cuvelier LL. Chemical analysis of French beans (Phaseolus vulgaris L.) by headspace solid phase microextraction (HS-SPME) and simultaneous distillation/extraction (SDE). Food Chem 2007;101(3):1279–84.
20. Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. Applied linear statistical models, 4th edn. Boston, MA: McGraw-Hill, 1996.
21. Hogg RV, Tanis EA. Probability and statistical inference, 5th edn. New Jersey: Prentice-Hall, 1997.

Additional information and reprint requests:
Helen Kreuzer, Ph.D.
Chemical and Biological Signature Sciences
Pacific Northwest National Laboratory
999 Battelle Blvd, MSIN P7-50
Richland, WA 99352
E-mail: Helen.Kreuzer@pnl.gov

PAPER

CRIMINALISTICS

Kelly M. Brinsko,[1] *M.S.*

Optical Characterization of Some Modern "Eco-Friendly" Fibers*

ABSTRACT: Fibers that are termed "eco-friendly" or "biodegradable" by manufacturers are increasingly being used in textile products such as apparel and carpeting to appeal to the ever more environmentally aware public. As such, these modern fibers are expected to begin showing more often in forensic casework, and it is important that the forensic examiner recognize them. This study employed polarized light microscopy (PLM) and Fourier transform infrared (FTIR) microspectroscopy to characterize selected fibers of azlon, polylactic acid (PLA), cellulose composites of alginate or chitin, and bamboo (viscose rayon). Fiber cross-sections, refractive indices, melting points, solubilities, and FTIR measurements were conducted. Results indicate that the azlons and PLA fibers are easily distinguishable from other textile fibers by their optical and chemical properties. The cellulose composites show only small differences in comparison with other cellulose-based fibers, while bamboo viscose rayon is indistinguishable from normal viscose rayon.

KEYWORDS: forensic science, criminalistics, fiber identification, azlon, polylactic acid, alginate, chitin, bamboo rayon, polarized light microscopy, Fourier transform infrared spectroscopy

The public's increased awareness of their environmental impact, together with improved manufacturing methods and inexpensive raw materials, have resulted in the large-scale production of certain fibers and fiber types that are marketed to appeal to such consumers. Because these fibers are made from biologically occurring polymers of corn, milk, soybeans, seaweed, bamboo, and others, they are proclaimed by manufacturers as being biodegradable, "eco-friendly," and an alternative to petroleum-based or persistent fibers such as polyester. Currently, the fibers are woven into textiles for clothing, upholstery, and carpeting, and so are expected to begin turning up more often in forensic casework. It is important that forensic scientists are aware of these modern fibers and are able to recognize them when they are encountered. This study characterized several different fibers of azlon, polylactic acid (PLA), cellulose composites of alginate or chitin, and bamboo viscose rayon. Polarized light microscopy (PLM) and Fourier transform infrared (FTIR) microspectroscopy techniques were used to examine fiber cross-sections, refractive indices, melting points, solubilities, and infrared spectra.

Production of Azlon Fibers

Azlon fibers are defined by the Federal Trade Commission (FTC) (1) as a manufactured fiber comprised of any regenerated, naturally occurring protein. The proteins used are typically the by-products of other manufacturing processes and usually result in a reduced cost: casein from milk after the butterfat is removed, zein from corn after starch manufacture, or soybean protein after the beans are pressed and oil is removed. Azlons were first made in

the 1890s, after it was found that coagulated casein from acid-treated milk could be dissolved in NaOH and then extruded through spinnerets into an acid bath to form the fiber. Proteins from corn and peanuts were also used in a similar process. These azlon fibers, however, were rather brittle and too fragile in the presence of water to be of much use as a textile fiber. In 1935, a stronger, more practical azlon was produced via treatment with formaldehyde or aluminum salts, creating cross-linkages between the proteins in the fiber. The fiber underwent large-scale production under the trade names of Aralac, Fibrolane, Lanital, Ardil, and Vicara, among others, and was commonly blended or felted with wool or rabbit hair. With the advent of the more durable polyesters and nylons in the 1960s, the production of most azlon fibers was discontinued (2). But modern azlons are now seeing a reemergence in China, and textiles produced may be 100% azlon or a blend, typically with cotton, silk, nylon, or wool. The fabric is promoted for its softness and is used in undergarments, T-shirts, sweaters, and baby clothing by many brand-name manufacturers.

Production of Polylactic Acid Fibers

PLA fibers were designated by the FTC as a new fiber type in 2002 (3,4). It was first produced in 1932 by placing lactic acid under high vacuum (5); however, the commercial method of PLA production begins with starch (usually corn starch) as the raw material, which undergoes enzymatic hydrolysis for conversion into dextrose. The dextrose is then fermented to produce lactic acid, which can form two stereoisomers: D- or L-lactic acid. Synthesizing lactic acid chemically results in a mixture of approximately 50% D- and 50% L-lactic acid, but natural fermentation results in the majority (99.5%) of L-lactic acid (6). The L-lactic acid is used to form highly crystalline polymers, such as PLA fibers, while D-lactic acid is more amorphous (6,7). Polymers formed from varying proportions of L- and D-lactic acid will therefore have

[1]McCrone Research Institute, 2820 S. Michigan Ave, Chicago, IL 60616

dissimilar physical properties, such as melting point and refractive index. The lactic acid next undergoes a condensation reaction into the cyclic lactide. Heating the lactide in the presence of a tin catalyst (typically stannous octoate) allows for the ring-opening and subsequent polymerization of the polylactic acid (6). PLA is currently being mass-produced by companies in the United States (Natureworks, LLC) (3,4), Japan (Kanebo) (6,7), and China. PLA textile products include apparel such as sportswear, upholstery, and carpet fibers. Carpet fibers in Japan's Toyota Prius are PLA (6). PLA has also been used in medical sutures (8).

Production of Cellulose Composite Fibers

Alginic acid, or alginate, exists in the cell walls of brown seaweed and is analogous to cellulose in land plants, with a similar chemical structure (2). Pure alginate polymers have been used for medical applications, including sutures or drug delivery devices (8), but the alginate textile fiber is a cellulose composite produced using the lyocell process (9). The lyocell process was developed in 1992 by Courtaulds and uses a procedure that is largely free from waste or pollution. In this method, cellulose is dissolved in N-methyl-morpholine-oxide (NMMO), which can later be recycled and reused (10). The alginate powder is added to the cellulose solution before extrusion. Adding alginate to a fiber purportedly incorporates the nutritive elements that are found in seaweed and promotes "well-being" (9,11). Zimmer AG, a textile manufacturer in Frankfurt, Germany, has produced such a fiber, named SeaCell® Pure. Silver or other metal ions may be added to achieve antimicrobial effects (11), as is the case with Zimmer AG's SeaCell® Active fibers.

Similarly, chitin can be added to cellulose solutions as well. Chitin is a polysaccharide of N-acetylglucosamine, similar in structure to cellulose and alginate and is found in many places within the animal and fungi kingdoms, including the exoskeletons of crustaceans and insects (2). The chitin may be added to the cellulose solution in either powder form or as a solution itself (12). The chitin/cellulose solution is then extruded in the viscose process. These composite fibers are also said to inhibit the growth of bacteria (13).

Production of Bamboo Rayon

Viscose rayon that is made from bamboo cellulose, instead of the typical softwoods such as spruce or pine, is publicized as being environmentally friendly because bamboo grows so quickly as a renewable resource. The bamboo is pulped, and cellulose is dissolved in sodium hydroxide and carbon disulfide, which is then extruded through a spinneret into a coagulating bath (2). The resulting "bamboo rayon," as it is often called, is comparable to regular viscose rayon and is used in a variety of consumer products, ranging from apparel to towels and bedding.

Methods

Cross-Sections

Cross-sections were produced using the plate method as outlined by Skirius (14). A thin plastic sheet containing several holes 1 mm in diameter (available from Insulfab Plastics, Inc, Spartanburg, SC) was utilized to hold bundled fibers that had been threaded through the holes. In some cases, black cotton embroidery thread was used alongside the test fibers in the bundle to provide support and cushion, as well as minimize the number of test fibers needed. A Teflon®-coated (DuPont, Wilmington, DE) razor blade held parallel to the sheet was then used to slice off the excess fibers from each side of the plastic sheet. Cutting out the small section and taping it to a glass microscope slide allows the analyst to observe the cross-sections using a transmitted light microscope.

Polarized Light Microscopy

Refractive index measurements were made for the n-parallel ($n_\parallel$) and n-perpendicular ($n_\perp$) of each fiber using Cargille Refractive Index Liquids (Cargille Labs, Cedar Grove, NJ) by the Becke line immersion method. Using a PLM (with the polarizer aligned east–west), short lengths of individual fibers were mounted in successive Cargille liquids until the edges of the fiber in the parallel or perpendicular orientation (relative to the polarizer) "disappeared" when illuminated with sodium D light (589.3 nm). Multiple fibers were measured, so in some cases a narrow range of refractive index values was noted because of small variations in the fibers themselves. Birefringence was calculated as the difference between $n_\parallel$ and $n_\perp$, and sign of elongation was verified with a first-order red compensator.

Melting point was determined microscopically using a calibrated Mettler FP-5 hotstage and FP-52 attachment (Mettler-Toledo, Columbus, OH), using a temperature range from 25°C to 300°C. Short lengths of fibers were mounted in silicone oil (Dow Corning 710 fluid; Dow Corning, Midland, MI) and inserted into the hotstage at approximately 30°C. The preparation was heated at a rate of 3°C per minute and observed between crossed polars with the first-order red compensator inserted for contrast. The temperature at which the fibers became isotropic was defined as the melting point. Three preparations were made for each fiber being tested, and melting point temperatures were recorded as a range of the highest and lowest values.

Solubility was determined using seven solvents: formic acid, chloroform, glacial acetic acid, acetone, dimethylformamide (DMF), concentrated sulfuric acid, and concentrated nitric acid. The fibers were placed on a clean glass microscope slide and coverslipped dry. Then, while observing the dry fibers through the microscope, a drop of solvent was allowed to run underneath the coverslip and immerse the fibers. The fibers were observed for 1 min and described as "soluble," "insoluble," "partially soluble," or "swells." Additionally, any emission of gas bubbles or gelling was noted.

Fourier Transform Infrared (FTIR) Microspectroscopy

Individual fibers were prepped for FTIR by flattening with a microroller onto a potassium bromide salt plate. FTIR spectra were obtained using a Mattson Galaxy 5020 equipped with a Quantum Infrared Microscope and WinFIRST™ Software (Mattson Instruments, Inc, Madison, WI). Spectra were manipulated and searched against existing libraries using the OMNIC E.S.P. software program (Nicolet Instrument Corp., Madison, WI).

Results

Azlons

Two azlon fiber subtypes were examined: Silk Latte™ (Southwest Trading Company, Tempe, AZ), a fiber formed from milk protein (casein), and Soy Silk® (Southwest Trading Company), formed from soybean protein. Soy Silk® is available in two colors, "white" or "natural" (a light tan color by reflected light), both of which were characterized in this study. These three fibers are available from Southwest Trading Company, a company which markets fibers primarily to knitters. The cross-sections of all three fibers are
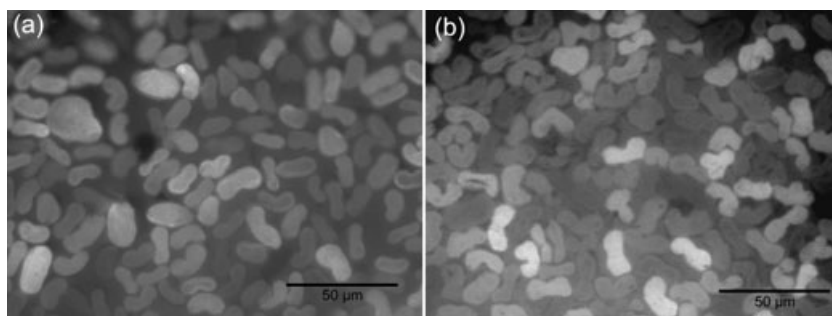
FIG. 1—*Fiber cross sections of Silk Latte™ (a) and Soy Silk® (b).*
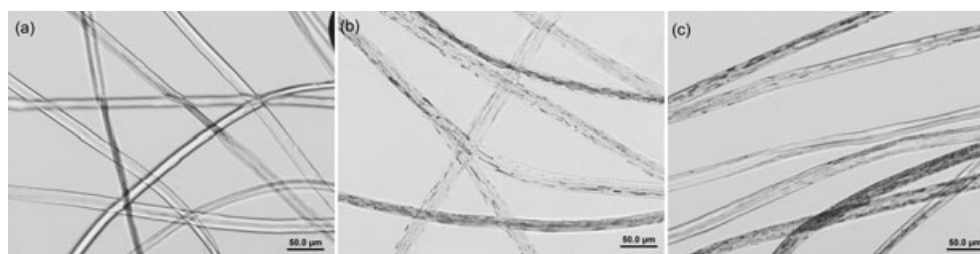

FIG. 2—*Transmitted light photomicrographs of Silk Latte™ (a), Soy Silk® "natural" color (b), and Soy Silk® "white" color (c). All fibers are mounted in Cargille Refractive Index Liquid n = 1.662 and viewed using plane polarized light.*

bean- or peanut-shaped (Fig. 1a,b). It should be noted here that earlier-generation azlon fibers are listed in references as having circular cross-sections (2). By PLM, the Silk Latte™ fibers are largely transparent, although the presence of small air bubble inclusions in some fibers results in a translucent or pseudo-opaque behavior; these fibers appear to be a light brown color by transmitted light because of light scattering (Fig. 2a). The colors of both Soy Silk® fibers appear identical to each other and similar to Silk Latte™ by PLM, but with more pronounced and elongated air bubble inclusions than the Silk Latte™ (Fig. 2b,c). No pigment or delustering pigment is present in any of the fibers. All fibers appear flattened or ribbon-like, which is consistent with observed cross-sections, having diameters of 10–25 μm. The fibers show parallel extinction between crossed polars. Most fibers exhibit complete extinction, although some demonstrate incomplete, undulose extinction behavior.

Refractive index ranges and birefringence for Silk Latte™ and Soy Silk® are listed in Table 1. Also included in Table 1 is the n_{iso}[1], which serves to give a good approximation of the refractive index of the unaligned polymer. Silk Latte™ has an $n_{\parallel}$ refractive index range of 1.536–1.544 and $n_{\perp}$ = 1.520. Soy Silk® is slightly higher, with $n_{\parallel}$ = 1.544–1.547 and $n_{\perp}$ = 1.520–1.524. Correspondingly, the birefringence range of Soy Silk® is slightly higher (B = 0.023–0.024) than that of Silk Latte™ (B = 0.016–0.024). As is illustrated in Table 1, previous data for the earlier generation of azlons has a much lower birefringence; ranging from isotropic (0.000) to 0.006 (15,16). Nylon 6,6, a polyamide, and silk, a natural protein fiber, both have much higher refractive indices and birefringence values. This is likely because of the fact that the monomer units for both nylon and silk are small and relatively linear, which allows for closer packing in the fiber. The azlons, made of bulkier proteins with structures that cannot be packed as closely, exhibit lower birefringence as a result. Rayon is the only common man-made fiber that has refractive indices and birefringence similar to the new azlons.

The melting range for Silk Latte™, 235–245°C, is lower than Soy Silk®, 250–260°C (Table 1). Both fiber subtypes shrink markedly as they are heated, but retain their general shape at the melting point. Vicara™, an earlier-generation azlon made from corn protein, is listed by Gaudette as melting between 265–275°C, while nylon 6,6 melts in nearly the same range as Soy Silk® (15). Thus, melting point is not a discriminating method for identifying an azlon fiber, although it may help when comparing two azlons or excluding rayon.

Solubility tests did allow for differentiation between Silk Latte™ and Soy Silk® (Table 2). When concentrated sulfuric acid (H_2SO_4) is added to the fibers, the Soy Silk® evolves gas bubbles and is only partially soluble, while Silk Latte™ is soluble and gels. This suggests that a solubility test may be one way to subclassify azlons.

TABLE 1—*Refractive indices, birefringence, and melting points for azlons and related fibers.*

Fiber	$n_{\parallel}$	$n_{\perp}$	Birefringence	n_{iso}	Melting Point, °C
Silk Latte™	1.536–1.544	1.520	0.016–0.024	1.53	235–245
Soy Silk®	1.544–1.547	1.520–1.524	0.021–0.027	1.53	250–260
Azlon* (15,16)	1.532–1.545	1.526–1.545	0.000–0.006	1.53–1.54	265–275[†]
Nylon 6,6 (15)	1.58	1.52	0.06	1.54	250–264
Silk (16)	1.591	1.538	0.053	1.56	DNM
Viscose Rayon	1.541–1.549	1.520–1.521	0.020–0.028	1.53	DNM

n_{iso}, Isotropic refractive index of the unaligned polymer, $[(2n_{\perp} + n_{\parallel})/3]$.
DNM, does not melt.
*Azlon fibers produced in the mid-20th century.
[†]Melting point determined for Vicara, a zein (corn protein) fiber.

[1]$n_{iso} = [(2n_{\perp} + n_{\parallel})/3]$

TABLE 2—*Solubilities.*

Fiber	Formic Acid	Chloroform	AcOH	Acetone	DMF	Conc. H$_2$SO$_4$	Conc. HNO$_3$
Silk Latte™	Swells	Insol.	Insol.	Insol.	Insol.	Sol., Gels	Sol., Gels
Soy Silk®	Swells	Insol.	Insol.	Insol.	Insol.	Part. Sol. (gas)	Part. Sol.
Ingeo™	Insol.	Sol.	Insol.	Insol.	Insol.	Sol.	Sol., Gels
China PLA	Swells	Sol.	Swells	Part. Sol.*	Swells	Sol.	Sol., Gels
SeaCell®	Swells	Insol.	Insol.	Insol.	Insol.	Sol.	Part. Sol.
Lyocell	Swells	Insol.	Insol.	Insol.	Insol.	Sol.	Part. Sol.
Chitin composite	Sol.	Insol.	Part. Sol.	Part. Sol.	Part. Sol.	Sol.	Part. Sol.
Viscose rayon	Insol.	Insol.	Insol.	Insol.	Insol.	Sol.	Swells
Bamboo rayon	Insol.	Insol.	Insol.	Insol.	Insol.	Sol.	Swells

*"Spinal column" formation; fiber appeared segmented and notched (Fig. 6).
PLA, polylactic acid; DMF, dimethylformamide.

IR spectra for both Silk Latte™ and Soy Silk® are similar (Fig. 3), and a library search did not return any close matches. Silk and nylon show similar spectra but can easily be differentiated from the azlon IR spectra. The azlons exhibit the characteristic amide I and II doublet, the C=O stretch at 1640 cm^{-1}, and the N–H deformations at 1530 cm^{-1} (17).

Polylactic Acid

PLA fibers from two manufacturers were examined in this study. One is produced by Natureworks, LLC under the brand name Ingeo™. The second was produced in China by an unknown manufacturer and will be denoted in this article as "China PLA." The cross-sections of these fibers examined are circular (Fig. 4), though other manufacturers such as FIT, Kanebo, Unifi, and Interface Carpet have made fibers with polygonal or triangular cross-sections (7), and Natureworks, LLC has produced other PLA fibers with trilobal, polygonal (false twist textured), and circular hollow cross-sections (3). The Ingeo™ fiber cross-sections are very uniform, with a narrow size range of 18–20 μm, while the China PLA fibers show a larger size distribution of 15–35 μm. The two fibers also show differences by PLM (Fig. 5a,b). The China PLA fibers, while mostly transparent, also exhibit transverse striations in some areas. These striations are dark and, by transmitted light, the fibers appear dark brown or tan. This pseudo-opacity is possibly because of the poor alignment of the polymer chains in that area of the fiber and may be an artifact of the extrusion process (7). The Ingeo™ fiber displays some transverse banding as well, though they are not nearly as dark or as numerous. Extinction is parallel and complete for both fibers, though the areas with banding exhibit undulose extinction, which is expected in poorly aligned polymers.

Refractive indices are lower for PLA than for any other common textile fiber (Table 3). Both the Ingeo™ and China PLA have comparable values, with the n$_{||}$ = 1.473. The n$_{\perp}$ = 1.445 for Ingeo™ and ranges from 1.445–1.447 for China PLA. Birefringence for both is around 0.028. These two PLA fibers have a slightly
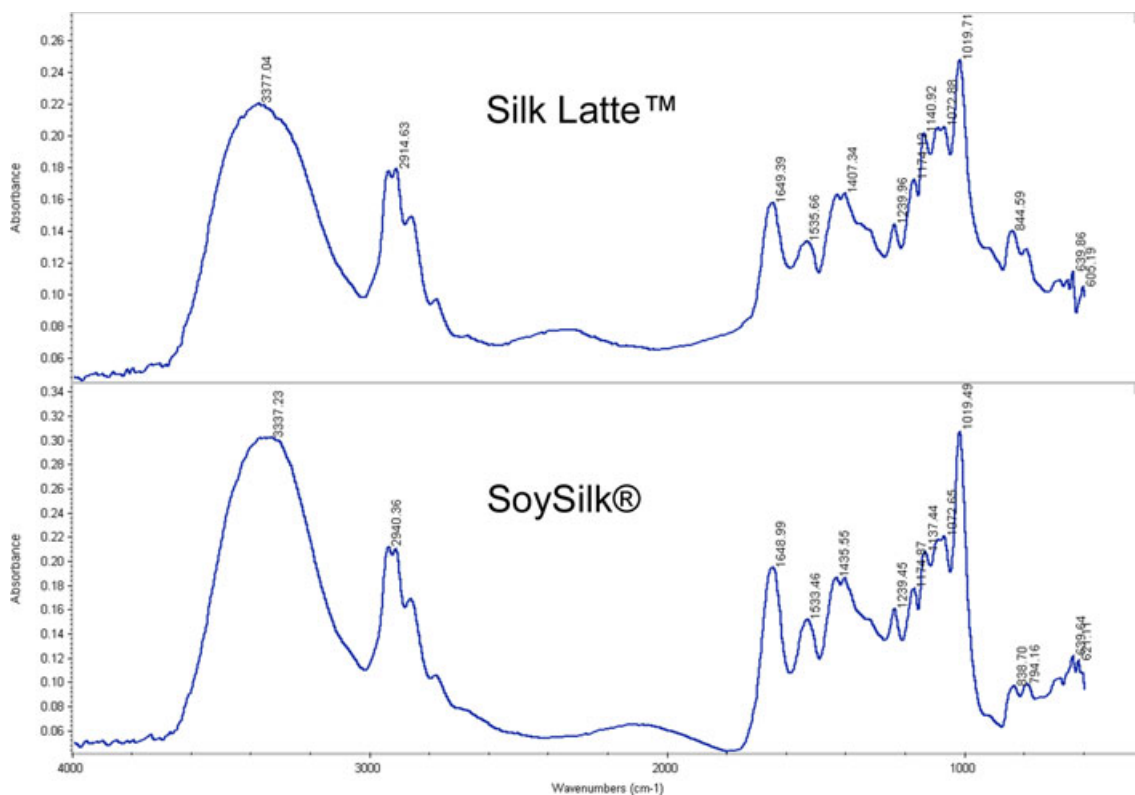
FIG. 3—*Fourier transform infrared spectra for SilkLatte™ (top) and Soy Silk® (bottom).*
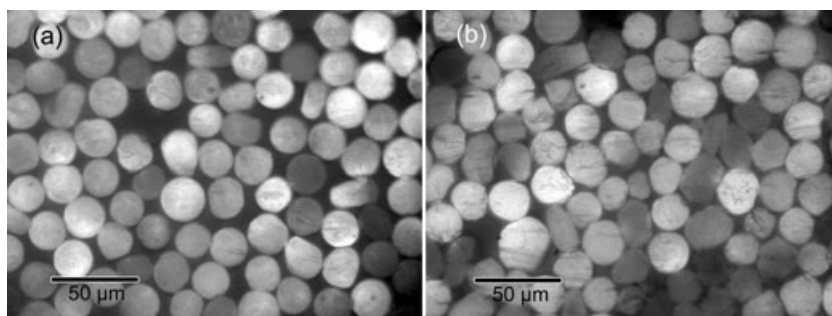
FIG. 4—*Fiber cross-sections of Ingeo™ (a) and China polylactic acid (b).*
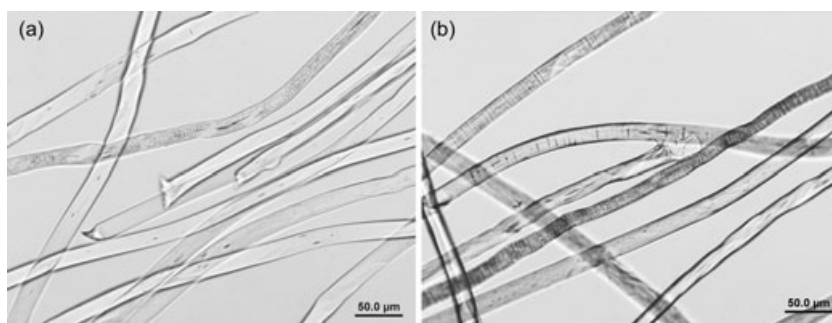
FIG. 5—*Transmitted light photomicrographs of Ingeo™ (a), mounted in Cargille Refractive Index Liquid (n = 1.530), and China polylactic acid (b) mounted in Cargille Refractive Index Liquid n = 1.662. Both preparations are viewed using plane polarized light.*

different n_{iso}, 1.45 for Ingeo™ and 1.46 for China PLA. However, because n_{iso} is only a calculated theoretical value, this difference may not be significant. Refractive index values measured here agree with the reference data (3,4,6,7). The only other common synthetic fiber that has refractive indices near PLA is acetate rayon, but acetate's higher perpendicular refractive index and lower birefringence (by an order of magnitude) (18) easily distinguishes acetate from PLA.

Melting points range from 160–161°C for China PLA, and 174–175°C for Ingeo™ (Table 3). The higher melting point for Ingeo™ may be attributed to more highly aligned polymer chains or a higher percentage of L-lactic acid compared to China PLA. The melting point for the olefin polypropylene ranges from 160–175°C (15), which encompasses both PLA fibers here. Melting point, therefore, is not a distinguishing characteristic of PLA fibers, though it would assist in the direct comparisons of PLA fibers.

Solubility data is listed in Table 2 for the tested PLA fibers. There are marked differences between the Ingeo™ and China PLA fibers: China PLA swells in formic acid, acetic acid, and DMF, while Ingeo™ is completely insoluble and experiences no change.

In acetone, China PLA takes on the appearance reminiscent of a human spinal column (Fig. 6), appearing segmented and notched, while Ingeo™ does not react. Peterson (7) also noted differences in solubility between manufacturers.

Infrared spectra for both Ingeo™ and China PLA were comparable, with nearly identical absorption peaks (Fig. 7). The spectrum for lactic acid is also given for comparison and shows that most peaks are similar, with one major discrepancy: the broad peak between 2600–3600 cm^{-1}, because of the hydrogen bonding taking place between the lactic acid molecules. Library searches do not result in any other close matches with the PLA fibers.

Cellulose Composites: Alginate and Chitin Cellulose Fibers

The alginate/cellulose fibers, SeaCell® Pure and SeaCell® Active (manufactured by Zimmer AG) were grouped for comparison with the chitin/cellulose fiber (of unknown manufacturer) because these are all blended with cellulose as composites. However, the SeaCell® fibers are made in the lyocell process, while the chitin composite fiber undergoes the viscose process, so their optical characteristics are not expected to be identical. The cross-sections (Fig. 8) of both SeaCell® fibers are identical, with a size

TABLE 3—*Refractive indices, birefringence, and melting points for PLAs and fiber with similar properties.*

Fiber	$n_{\parallel}$	$n_{\perp}$	Birefringence	n_{iso}	Melting Point, °C
Ingeo™	1.473	1.445	0.028	1.45	174–175
China PLA	1.473	1.445–1.447	0.026–0.028	1.46	160–161
Acetate (15)	1.474–1.479	1.473–1.477	0.002–0.005	1.47	245–260
Polypropylene (15)	1.520–1.530	1.491–1.496	0.028–0.034	1.50	160–175

n_{iso}, Isotropic refractive index of the unaligned polymer, $[(2n_{\perp} + n_{\parallel})/3]$. PLA, polylactic acid.

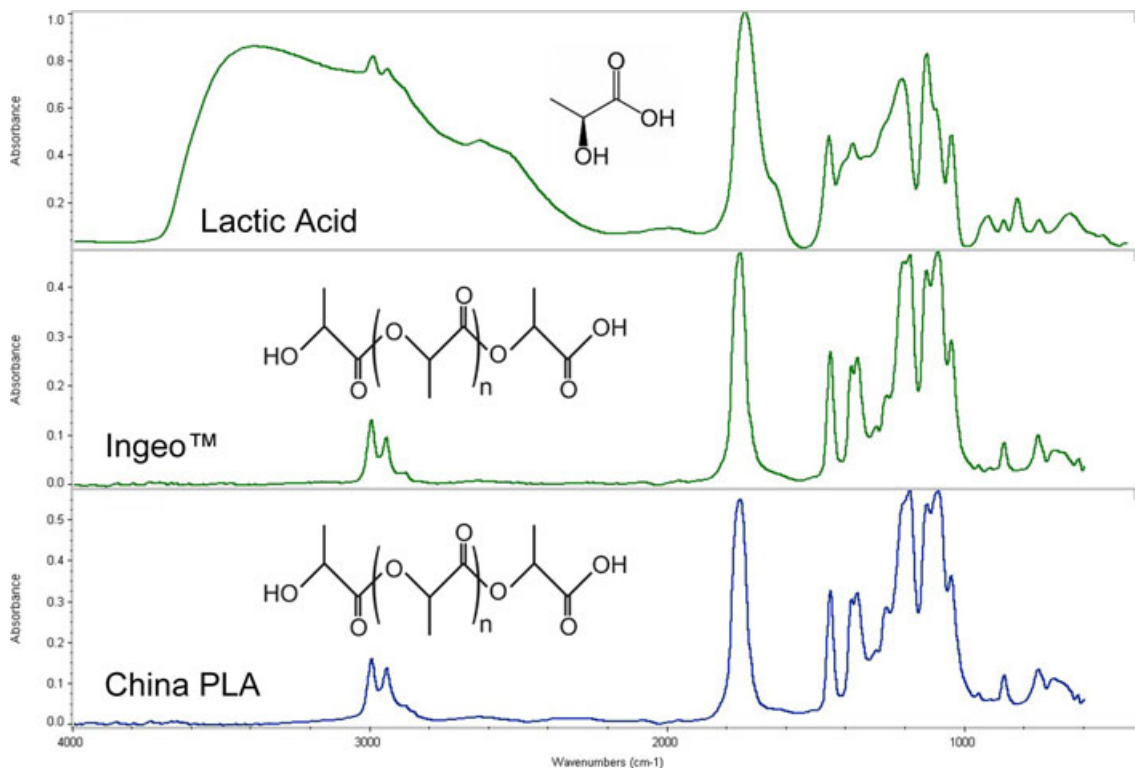FIG. 6—*"Spinal Column" formation of China polylactic acid after addition of acetone.*

FIG. 7—*Fourier transform infrared spectra for lactic acid (top), Ingeo™ (middle), and China polylactic acid (bottom).*

range of 8–13 μm, and typical of a lyocell, rather ovoid with some random edges and corners. The cross-section of the chitin composite looks similar to viscose rayon, with the irregular multi-lobed cross-section, though the chitin composite is more rounded, with fewer lobes than a standard rayon. The fiber diameters range from 10–15 μm. The chitin composite fibers depicted in Fig. 8 are dyed a dark pink hue and blended with cotton fibers to form the yarn (the yarn was approximately 10% chitin composite fiber). Microscopical characterizations were conducted by picking the individual chitin composite fibers out of the cotton yarn for analysis.

By transmitted polarized light, the SeaCell® fibers (Fig. 9) are colorless and smooth with a size range of 8–13 μm. No pigment or delustering pigment was present. Cross-hatchings typically seen in lyocells (10) are evident, as are some areas with small fibrils splitting from the main filament (10). Some parts of the fibers look damaged in plane polarized light, and extinction is undulose in these areas. In some fibers, extinction is complete and parallel. The chitin composite fibers appear colorless or light pink (from the dye) by transmitted light, with slight striations along the length of the fiber corresponding to the fibers' lobes. The chitin fibers are not as highly oriented as most rayons, because between crossed polars their extinction appears to be faintly incomplete and undulose when aligned in the parallel extinction position.

Refractive index values for both SeaCell® fibers are slightly higher than published data for lyocells (Table 4), with the $n_{\parallel}$ for SeaCell® near 1.570, versus 1.564 for most lyocells (18). The $n_{\perp}$ is also higher for SeaCell®, around 1.528 compared to lyocell at 1.522. The birefringence and n_{iso} values of SeaCell® and lyocell are nearly equal. No difference is observed between the SeaCell® Pure and the SeaCell® Active, but the silver present in the Sea-Cell® Active can be detected by SEM/EDS, TXRF, and LA-ICP-MS (19).

The chitin composite fibers were compared to viscose rayon, and several marked differences were found. The $n_{\parallel}$ of the chitin composite is in a range between 1.560 and 1.562, while typical values for viscose rayons are in the range of 1.541–1.549 (18). The $n_{\perp}$ of the chitin composite is also higher, 1.532–1.534, compared to
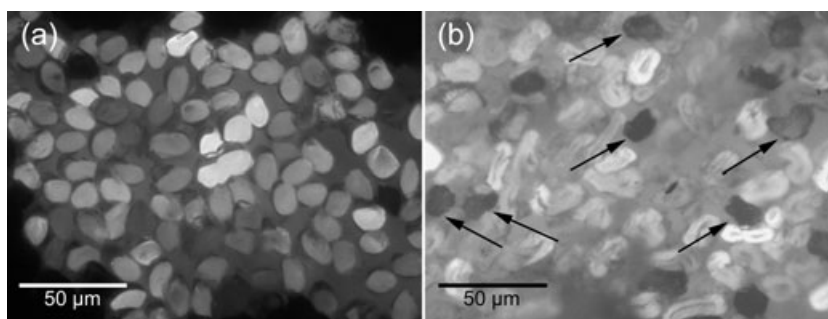
FIG. 8—*Fiber cross-sections of SeaCell® (a) and chitin composite fiber blended with cotton (b). Arrows indicate chitin composite fibers.*

TABLE 4—*Refractive indices, birefringence, and melting points for cellulose composites and cellulose fibers.*

Fiber	$n_∥$	$n_⊥$	Birefringence	n_{iso}	Melting Point, °C
SeaCell®	1.568–1.570	1.526–1.528	0.042	1.54	DNM
Lyocell (18)	1.562–1.564	1.520–1.522	0.044	1.54	DNM
Chitin composite	1.560–1.562	1.532–1.534	0.026–0.030	1.54	DNM
Viscose (15)	1.541–1.549	1.520–1.521	0.020–0.028	1.53	DNM
Bamboo Rayon	1.546	1.518	0.028	1.53	DNM

n_{iso}, Isotropic refractive index of the unaligned polymer, $[(2n_⊥ + n_∥)/3]$. DNM, does not melt.

viscose at 1.520–1.521. However, the birefringence of the chitin composite falls within the normal range for viscose (18).

None of the cellulose composites or cellulose standards melt below 300°C (the upper limit of the hotstage). All fibers discolored to a golden yellow by 250°C, eventually turning a very dark brown by 300°C, typical decomposition of cellulosic fibers.

The solubilities for SeaCell® and lyocell are identical (Table 2), while the chitin composite shows some major differences from viscose rayon. The chitin composite is soluble or partially soluble in formic acid, acetic acid, acetone, and DMF, while viscose rayon is insoluble. The chitin composite is also partially soluble in nitric acid and swells in water (not shown here). Based on solubility results, it seems surprising that this chitin composite is advertised as being ideal as a yarn for socks.

Infrared spectra from SeaCell® indicate that it is indistinguishable from lyocell or viscose (Fig. 10). The chitin composite is similar, but has a broad, shoulder-like peak between 1500 and 1600 cm^{-1} that is not seen in any of the other cellulose traces.

Bamboo Viscose Rayon Fibers

Viscose rayon that is made from bamboo as the raw cellulose material is optically and physically indistinguishable from viscose made from softwoods. The cross-section is multi-lobed like viscose, and by polarized light bamboo rayon has refractive indices and birefringence values comparable with those of viscose (Table 4). Solubility tests do not show any differences between the fibers (Table 2), and IR spectra are identical.

Discussion

Azlon Fibers

This study has shown that the optical properties of at least some of the new azlon fibers are not consistent with previous references (15,16,18) that list azlons as having birefringence values an order

of magnitude lower than those reported here. Refractive indices and birefringence are most consistent with rayon, though microscopical appearance and infrared spectra easily differentiate rayon from azlon. Melting temperatures and solubility tests may also allow an examiner to discriminate between azlons of different protein sources. The results here are promising, but given the small sample size, more studies should be carried out on a greater number of fibers from different manufacturers and protein sources.

Polylactic Acid Fibers

PLA fibers can be identified optically based solely on refractive indices and birefringence, and results here are consistent with data from previously published sources (3,4,6,7). Only acetate and triacetate rayon have refractive indices close to that of PLA, but the low birefringence of acetate and virtually isotropic quality of triacetate allow for a relatively straightforward differentiation. The microscopical appearance, melting points, and solubilities did reveal differences in the two PLA fibers that were analyzed, which suggest that these characteristics may be helpful in the forensic comparison of PLA fibers. In one study, Peterson examined seven different PLA fibers from four manufacturers and reported minimal differences in refractive index and birefringence, but a melting point range of 165.2–176.2°C and differences in solubilities between various manufacturers (7). Peterson's data corroborate the findings of this study and indicate that with such considerable differences between manufacturers, larger studies of the optical characteristics of PLA fibers are warranted.

Cellulose Composite Fibers

Cellulose composite fibers that were investigated here include those made using either the viscose rayon manufacturing process or the lyocell process. SeaCell®, an alginate/cellulose composite, typified the microscopical appearance of lyocell, including the cross-section, cross-hatchings, and presence of fibrillation. Refractive indices are slightly higher for SeaCell® compared to published data for lyocell (18), and birefringence values are comparable. SeaCell® cannot be distinguished from lyocell by melting point, solubility, or infrared spectra. One version of the fiber, SeaCell® Active, does have added silver, which is detectable by other analytical methods (19).

The chitin composite fiber looks similar to viscose rayon in cross-section, but both the parallel and refractive indices are higher than viscose by about 0.02 and 0.01, respectively (18). Birefringence values are comparable, but solubilities are very different. The chitin composite does not perform like a resilient textile fiber would be expected to; it was soluble or partially soluble in all but chloroform. This fiber may need further development before it can
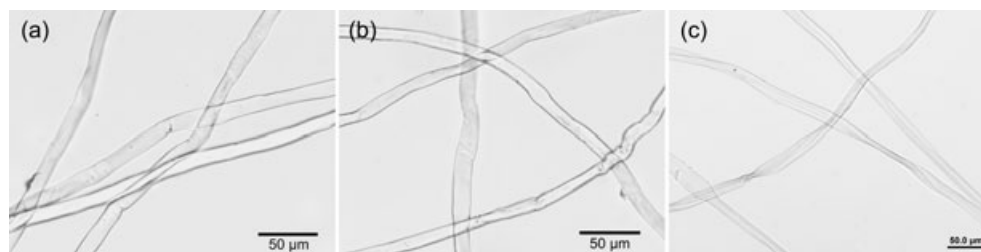

FIG. 9—*Transmitted light photomicrographs of SeaCell® Pure (a), SeaCell® Active (b), and chitin composite fiber (c). SeaCell® fibers are mounted in Cargille Refractive Index Liquid (n = 1.662). Chitin composite fiber is mounted in Cargille Refractive Index Liquid (n = 1.600). All preparations are viewed using plane polarized light.*
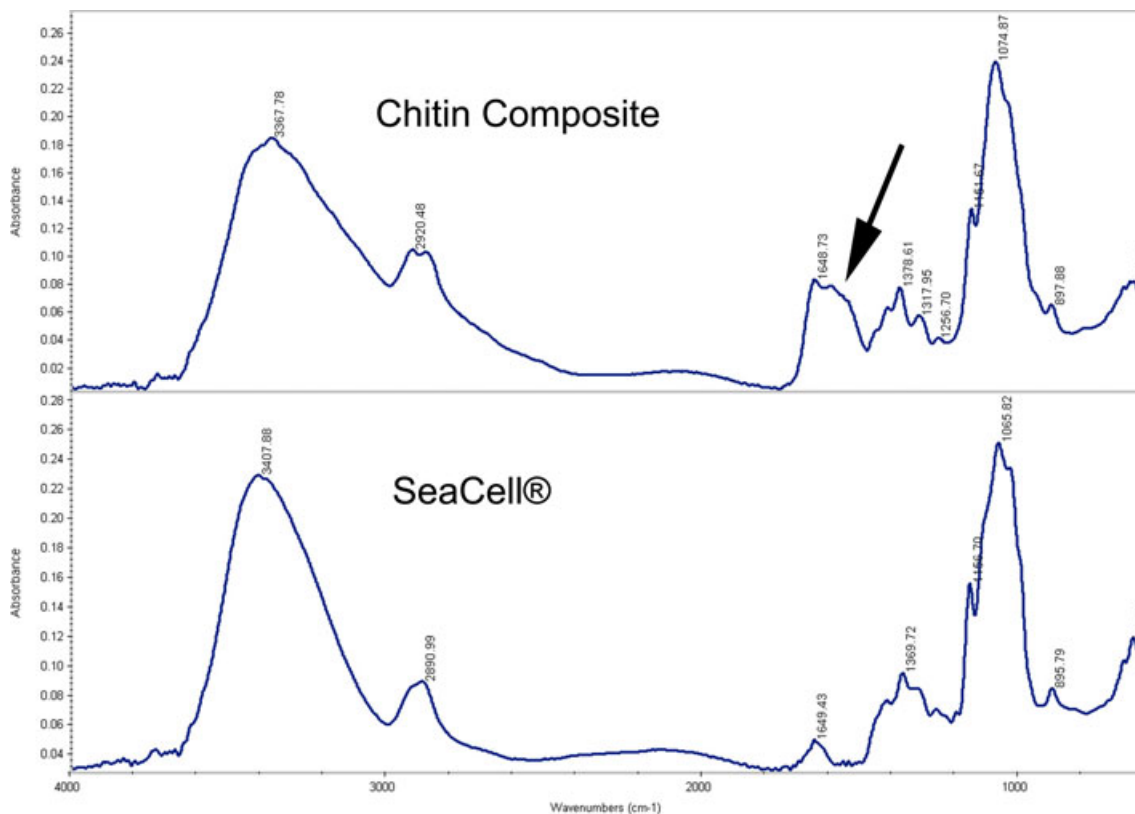
FIG. 10—*Fourier transform infrared spectra for chitin composite fiber (top) and SeaCell® (bottom). Arrow indicates broad, shoulder-like peak between 1500–1600 cm^{-1} that was not seen in any other cellulose traces.*

be adopted into widespread use as a textile. However, because much work is being carried out in the field of cellulose composites, especially in medical applications, it is not unlikely that a forensic examiner will encounter some type of composite during casework. Currently, other published reports of either SeaCell® or chitin composite fibers contain information about the manufacturing process and mechanical properties (for example, tenacity or dye adsorption) but do not discuss optical properties such as refractive index or birefringence. Clearly, these fibers need to be investigated and optically characterized more fully for the forensic fiber examiner.

Bamboo Rayon

Bamboo rayon looks and behaves optically just like ordinary viscose. The two fibers are indistinguishable based on refractive indices, birefringence, melting points, solubilities, and infrared spectra. One company, the China Textile Industry Testing Institute (20), published a method in 2003 that reports a differentiation of bamboo rayon from ordinary viscose based on the increased solubility of bamboo rayon in 55.5% sulfuric acid solution. The method requires 0.5 g of fiber (an impractical amount to the forensic fiber examiner) that is dried and weighed before and after treatment with sulfuric acid. In their report, bamboo rayon, being more soluble in sulfuric acid, will lose over 30% of its weight, while ordinary viscose will lose nearly 20% of its weight. However, the author was unable to replicate these results. If bamboo does possess an increased solubility in sulfuric acid, perhaps this can be exploited by looking at dye penetration or color uptake microscopically. There is currently no way to microscopically differentiate bamboo rayon from ordinary viscose.

References

1. Federal Trade Commission. Rules and regulations under the textile fiber products identification act. 16 CFR Part 303. §303.7(g).
2. Cook JG. Handbook of textile fibers. Vol 2. Man-made fibers, 5th edn. Durham, England: Merrow Publishing Co. Ltd., 1984.
3. Wilson SK, Inafuku RA, Jett JJ. A second glance at polylactic acid (PLA) fibers. Microscope 2007;55:117–25.
4. Velez H. Identification characteristics of PLA fibers: a new generic fiber type. Forensic Sci Commun 2003:5(3), http://www.fbi.gov/hq/lab/fsc/backissu/july2003/velez.htm (accessed June 26, 2008).
5. Gupta B, Revagade N, Hilborn J. Poly(lactic acid) fiber: an overview. Prog Polym Sci 2007;32:455–82.
6. Farrington DW, Lunt J, Davies S, Blackburn RS. Poly(lactic acid) fibers. In: Blackburn RS, editor. Biodegradable and sustainable fibres. Cambridge, England: Woodhead Publishing Limited, 2005;191–220.
7. Peterson LK. Characterization of polylactic acid (PLA) fibers. Microscope 2002;50(1):37–43.
8. Nair LS, Laurencin CT. Biodegradable polymers as biomaterials. Prog Polym Sci 2007;32:762–98.
9. SeaCell GmbH. Technical information: SeaCell pure, http://www.smart-fiber.de/files/infoblaetter/technical-information-pure.pdf (accessed May 2, 2008).
10. Grieve M. New fibre types. In: Robertson J, Grieve M, editors. Forensic examination of fibres, 2nd edn. Boca Raton, FL: CRC Press, 1999;399–419.
11. Zikeli S. Production process of a new cellulosic fiber with antimicrobial properties. Curr Probl Dermatol 2006;33:110–26.

12. Shimizu Y, Nakajima T, Yoshikawa M, Takagishi T. Dyeing chitin/cellulose composite fibers with an acid dye. Text Res J 2002;72(7):563–7.
13. Swicofil AG Textile Services. Technical bulletin: chitosan fibers, http://www.swicofil.com/products/055chitosan.html (accessed December 17, 2008).
14. Skirius SA. An easy crossectioning technique. Microscope 1986;34: 26–7.
15. Gaudette B. The forensic aspects of textile fiber examination. In: Saferstein R, editor. Forensic science handbook. Vol. 2. Englewood Cliffs, NJ: Prentice Hall, 1988;209–72.
16. McCrone Research Institute. Appendix 19: optical properties of textile fibers. Forensic Hair and Fiber Microscopy Course Appendices, 2009. Chicago, IL: McCrone Research Institute, 2009.
17. Kirkbride KP, Tungol MW. Infrared microscopy of fibres. In: Robertson J, Grieve M, editors. Forensic examination of fibres, 2nd edn. Boca Raton, FL: CRC Press, 1999;179–222.
18. Palenik SJ. Microscopical examination of fibres. In: Robertson J, Grieve M, editors. Forensic examination of fibres, 2nd edn. Boca Raton, FL: CRC Press, 1999;153–77.
19. Koch S. Emerging fibers. Global Forensics 2007;1:2–5.
20. China Textile Industry Testing Institute. Bamboo fiber identification rule, http://www.swicofil.com/bambrotexbambooidentification.pdf (accessed January 30, 2009).

Additional information and reprint requests:
Kelly M. Brinsko, M.S.
Research Microscopist/Instructor
McCrone Research Institute
2820 South Michigan Avenue
Chicago, IL 60616
E-mail: kbrinsko@mcri.org

PAPER

CRIMINALISTICS

Oliver Dalby,[1] B.Sc.; David Butler,[1] M.Sc.; and Jason W. Birkett,[1] Ph.D.

Analysis of Gunshot Residue and Associated Materials—A Review

ABSTRACT: A comprehensive review of the scientific literature on gunshot residue (GSR) is presented. Aspects of both inorganic and organic GSR are discussed, from formation and distribution, to sample collection, preparation, and analysis using a variety of techniques. The interpretation of GSR results is also considered including issues surrounding the contamination, distribution, and transfer of GSR. Potential problems with ulterior sources of GSR like particles have been reported in the literature. For example, particles from environmental and occupational sources have been highlighted as exhibiting similar chemical and morphological characteristics to GSR. These findings are put into context with regard to interpreting samples. A move toward a "case by case" approach is argued to be more preferable to a "formal" classification system where possible. The analysis of both inorganic and organic compositions of residue samples as well as morphological considerations is considered to be a more ideal approach to GSR analysis, wherever practicable.

KEYWORDS: forensic science, ballistics, gunshot residue, cartridge discharge residue, firearm discharge residue, scanning electron microscopy, gas chromatography, liquid chromatography, environmental sources

Between 2006/2007 there were 9650 reported incidents involving firearms in England and Wales (excluding air rifles), 566 of which involved serious or fatal injuries. Firearms were discharged in 40% of these reported incidents. Despite a drop in the number of incidents involving firearms in the last 3 years, the general trend appears to show a rise in their numbers (Fig. 1). A variety of firearm types were used in reported incidents and these are summarized in Fig. 2 (1). Despite firearm offenses accounting for only 0.2% of the total crime volume in England and Wales (2006/07) (1), the gravity of their nature means that they must not be overlooked and it must be seen as paramount that forensic evidence involved in such cases is as reliable and accurate as practicable.

Mejia (2) raised a number of questions about the methods currently used for the analysis of gunshot residues (GSR): Which methods of analysis are most effective? How can false negative or positive results be best minimized? Is it actually possible to state with certainty that a person has discharged a firearm?

Ronald Singer, former president of the American Academy of Forensic Sciences thinks not, stating "None of what we do can establish if anybody discharged a firearm" (2).

Aleksandar (3) comments "In my 10 years of very intense practice… I still haven't met the method on the basis of which with absolute reliability, as a court expert working on identification activities of eventual shooters, I could confirm that the suspect 100% fired from the firearm on that occasion."

High profile cases in the U.K. such as the Jill Dando murder trial (4) have brought the evidential value of GSR analysis further into question. Therefore, a review of techniques used for the analysis of GSR and their effectiveness is valuable, relevant and timely.

This review covers the formation and distribution of GSR, the collection, preparation, and analysis of GSR samples, and discusses factors that may affect the interpretation of any given GSR sample.

What is Gunshot Residue?

Gunshot residue, which may also be known as cartridge discharge residue (CDR) or firearms discharge residue are particles produced during the discharge of a firearm. When a cartridge/round is fired in a firearm, combustion products from both the primer and the propellant will be released at the same time (5).

Gunshot residues are composed of unburned and partially burnt propellant powder, particles from the ammunition primer, smoke, grease, lubricants, and metals from the cartridge as well as the weapon itself (6,7). Organic compounds mainly originate from propellant and firearm lubricants, taking the form of unburned and partially burned gunpowder particles, some products of their
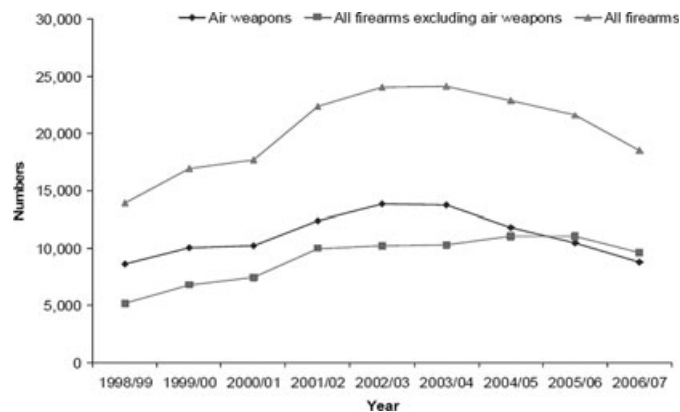
[1]Department of Chemistry and Forensics, School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, U.K.

FIG. 1—*Government statistics for the number of reported incidences involving firearms in England and Wales 2006/2007 (1).*
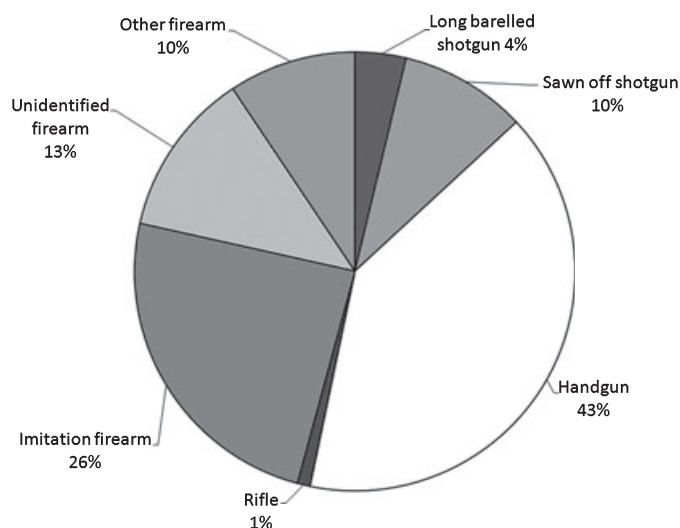
FIG. 2—*Government statistics of the types of firearms used in reported incidents in England and Wales 2006/2007 (1).*

TABLE 1—*Organic compounds that may contribute to gunshot residue (10–14,21,90,93,94,117,127,142).*

Compound	Source of Compound
2,4,6-Trinitrotoluene (TNT)	Propellant powder/primer mix
2,4-Dinitrodiphenylamine (2,4-DPA)	Propellant powder
2,3-Dinitrotoluene (2,3-DNT)	Propellant powder
2,4-Dinitrotoluene (2,4-DNT)	Propellant powder
2,6-Dinitrotoluene (2,6-DNT)	Propellant powder
2-Nitrodiphenylamine (2-NDPA)	Propellant powder
4-Nitrodiphenylamine (4-NDPA)	Propellant powder
Akaridte II (AKII)	Propellant powder
Butyl phthalate	Propellant powder
Butylcentralite (*N,N*-Dibutylcarbanilide)	Propellant powder
Camphor	Propellant powder
Carbanilide	Propellant powder
Carbazole	Propellant powder
Charcoal (major carbon)	Black powder
Cresol	Propellant powder
Dextrin	Primer mix
Diazodinitrophenol	Primer mix
Diazonitrophenol	Primer mix
Dibutyl phthalate	Propellant powder
Diethyl phthalate	Propellant powder
Dimethyl phthalate	Propellant powder
Dimethylsebacate	Propellant powder
Dinitrocresol	Propellant powder
Diphenylamine (DPA)	Propellant powder
Ethyl centralite (*N,N*-Diethylcarbanilide)	Propellant powder
Ethyl phthalate	Propellant powder
Ethylene glycol dinitrate	Propellant powder
Gum Arabic	Primer mix
Gum tragacanth	Primer mix
Karaya gum	Primer mix
Methyl cellulose	Propellant powder
Methyl centralite (*N,N*-Dimethylcarbanilide)	Propellant powder
Methyl phthalate	Propellant powder
Nitrocellulose (NC)	Propellant powder/primer mix
Nitroglycerine (NG)	Propellant powder/primer mix
Nitroguanidine	Propellant powder
Nitrotoluene	Propellant powder
N-nitrosodiphenylamine (N-NDPA)	Propellant powder
Pentaerythritol tetranitrate (PETN)	Propellant powder/primer mix
Picric acid	Propellant powder
RDX (Cyclonite)	Propellant powder
Resorcinol	Propellant powder
Rubber cement	Primer mix
Sodium Alginate	Primer mix
Starch	Propellant powder
Tetracene	Propellant powder/primer mix
Tetryl	Propellant powder/primer mix
Triacetin	Propellant powder

This list is not exhaustive. Some of the substances present may now be obsolete from production but are included as obsolete ammunition may still be in circulation.

transformation, and hydrocarbons. Inorganic residues such as nitrates, nitrites, and metallic particles originate from the primer and propellant as well as the cartridge case, the projectile jacket or its core and from the weapon barrel itself (8). These combustion materials from the primer, propellant, and other sources escape from weapon openings and vaporized materials solidify into particulates. These particulates are usually of sizes ranging from 0.5 μm to 10 μm (diameter) (5), although sizes of up to 100 μm have been reported (9).

Organic Compounds Present in GSR

Organic compounds are predominantly found in the propellant powder and the primer mixture, but can also originate from every part of the ammunition used (10). Table 1 contains a list of organic compounds that may contribute to the composition of organic GSR (OGSR).

The original propellant type used in firearms was black powder. The first recorded European recipe for black powder was produced by Roger Bacon in 1250 (5). Black powder is typically composed of 75% potassium nitrate (saltpeter), 15% sulfur, and 10% charcoal (fuel) (10,11). Today black powder is hardly ever used as a firearm propellant; however, it may still be encountered, usually linked to persons involved in military re-enactments.

Smokeless powders have long since replaced black powder as the main propellant used in firearms. Single base powders are based around nitrocellulose (NC) as an explosive, while double base powders contain both NC and nitroglycerine (NG), with NG increasing the energy potential of a powder. In triple base powders, a portion of NC and NG is replaced by nitroguanidine (10).

All smokeless powders, in addition to explosive ingredients, contain a number of additives, including stabilizers, plasticizers, flash inhibitors, coolants, moderants, surface lubricants, and anti-wear additives. A particular propellant powder will contain one or more of these additives depending on its intended use (10,12–18).

Inorganic Compounds Present in GSR

Inorganic components predominantly derive from primer mixtures. The particular type of primer used in a cartridge will generate materials of interest because of variations in formulation

compositions from manufacturer to manufacturer. The history of primer formulation development can be found elsewhere (5). The first example of a "modern" primer formulation was produced in 1921, containing lead styphnate, barium nitrate, and antimony trisulfate (5).

The most recent major development in primer manufacture has been the introduction of numerous types of lead (Pb) and heavy metal-free primers. These primer types have been produced in response to increased concerns over health problems relating to airborne pollution and exposure to high levels of heavy metals, such as those found in "modern" primer mixtures (19). There is currently a large number of manufacturers' producing completely lead-free ammunition (20), while others produce cartridges with

lead-free primers but not lead-free bullets. The compounds used in lead-free primers differ from manufacturer to manufacturer (19).

Further inorganic constituents may originate from the cartridge case, primer cup, bullet (including jacket), and the barrel of the firearm (21). Table 2 contains a list of inorganic compounds that may contribute to the composition of inorganic GSR.

The Formation of Inorganic GSR Particles

Gunshot residue particles form during the discharge of a firearm. As the firing pin strikes the primer cap, the primer mixture is ignited, creating an environment of rapid temperature and pressure

TABLE 2—*Inorganic compounds that may contribute to gunshot residue (5,11,12,21,60,116,117,127,143).*

Compound	Source of Compound
Aluminum	Primer/case
Aluminum sulfide	Primer mix
Antimony	Case/bullet
Antimony sulfide	Primer mix
Antimony sulfite	Primer mix
Antimony trisulfide	Primer mix
Arsenic	Case
Barium nitrate	Primer mix/propellant powder
Barium peroxide	Primer mix
Bismuth	Case
Boron	Primer mix
Brass	Case
Bronze	Bullet
Calcium carbonate	Propellant powder
Calcium silicide	Primer mix
Chromium	Bullet
Copper	Bullet jacket/primer cup/case
Copper thiocyanate	Primer mix
Cupro-nickel	Bullet jacket
Gold	Primer mix
Ground glass	Primer mix
Iron	Rust inside barrel, bullet
Lead	Bullet
Lead azide	Primer mix
Lead dioxide	Primer mix
Lead nitrate	Primer mix
Lead peroxide	Primer mix
Lead stifnate (styphnate)	Primer mix
Lead thiocyanate	Primer mix
Magnesium	Primer mix
Mercury	Primer mix
Mercury fulminate	Primer mix
Nickel	Case
Nitrate	Black powder
Phosphorus	Case
Potassium chlorate	Primer mix
Potassium nitrate	Propellant powder/primer mix
Prussian blue	Primer mix
Red brass	Bullet jacket
Silicon	Primer mix
Sodium nitrate	Primer mix
Sodium sulphate	Propellant powder
Steel	Bullet core/case
Strontium nitrate	Primer mix
Sulphur	Primer mix/black powder
Tin	Primer mix
Titanium	Primer mix/Lead free primer mix
Tungsten	Bullet
Yellow brass	Bullet jacket/case
Zinc	Primer cup
Zinc peroxide	Primer mix
Zirconium	Primer mix

This list is not exhaustive. Some of the substances present may now be obsolete from production but are included as obsolete ammunition may still be in circulation.

increases within the cartridge. This increase in temperature melts the primer mixture and within a few milliseconds the vaporization points of lead (Pb), barium (Ba), and antimony (Sb) are exceeded (Pb 1620/C, Ba 1140/C, Sb 1380/C). The effects of supersaturation cause vaporized particles to condense back onto the liquefied primer surface as droplets. There has been evidence to suggest that inorganic GSR particles of materials originating solely from the primer (primer GSR) are formed even before the propellant is ignited (22).

As the primer mix ignites the propellant powder, a second rapid increase in pressure and temperature occurs and the bullet is expelled from the firearm barrel. During this process, the particles involved are subjected to extreme temperature and pressure followed by rapid cooling. Particles form as liquid droplets, which subsequently solidify (22).

Wolten and Nesbitt (23) suggested that GSR particles formed from inorganic substances can be divided into two categories, "primer particles" and "bullet particles." Primer particles were shown to contain oxides, sulfides and salts in which the anion contained oxygen (oxysalts), such as barium meta-antimonate and basic lead sulfate (lanarkite). It was reported that primer ingredients are initially compounds and therefore cannot be expected to be reduced to elements in the oxidizing environment of the primer ignition. Elemental particles should therefore originate from bullet materials (23).

Basu (22) further divided inorganic GSR particles originating from the primer (with a lead styphnate, barium nitrate and antimony trisulfate mixture) into three categories (Table 3), each formed differently, depending on how they interact during firearm discharge. Categories II and III are larger sized particles that are thought to travel slower through the propellant ignition front and are therefore subjected to a second, greater increase in pressure and temperature. This causes these particles to pass through various metastable states because of boiling, fragmenting, or etching until a stable shape is again reached (22).

Burnett (24) suggested an alternative reason for the formation of irregular inorganic GSR particles (Basu's category II type), reporting the observation of particles still being molten upon impact onto target materials. Particles were observed to be splattered or flattened on target impact causing sometimes drastic modification of spherical form. This occurred especially at muzzle to target distances between 20 and 30 cm (although observed between 10–70 cm). Larger particles (>2 μm diameter) that came into contact with a target within 40 cm of the muzzle were often molten on impact (9 mm semi-automatic pistol). At distances >50 cm, it was shown that particles may shatter or adhere to the target surface (24).

GSR Collection Techniques

The areas from which GSR may be collected are wide ranging. Skin, vehicles (seats and seat backs, doors, windows, dashboards, headliners, interiors, and exteriors), the surroundings of an incident, doors, windows, body parts, clothing, and any surfaces in the immediate vicinity of a firearm discharge may all be sample targets (9). There are numerous techniques that can be used for GSR sample collection and selecting the most appropriate one is important in ensuring maximum collection efficiency.

Tape Lifts

Tape lifts are the most commonly used procedure for the collection of inorganic residues from skin surfaces (7). It has also been applied to sample collection from hair (25) and other mediums (26,27).

TABLE 3—*Inorganic primer residue particle types proposed by Basu (22).*

Particle Category	Mechanisms of Formation
I	Most commonly found type on the hands of a person who has previously discharged a firearm. Small spheroid particles that may have small nodules/bulges upon their surfaces. Nodules are usually of single elemental composition, which may originate from the primer or bullet and may be nonspecific (not Pb, Ba, or Sb). Believed nodules are captured by the main particle mass when a degree of cooling has occurred and the particle is solidifying. The main body of these particles contains a uniform and concurrent mixture of Pb, Ba, and Sb. Sizes range from 2 μm to 10 μm in diameter. Generally solid with no space in the core. Particles likely homogeneous due to thermal agitation and Brownian motion. Particles are formed prior to the ignition of the propellant mix and because of their smaller size travel through the propellant ignition detonation front much faster than larger particles leaving unaffected. Represented 68% of the total gunshot residue on a firer's hands
II	Twenty five percent of all particles found. Inhomogeneous and have a discontinuous distribution of Pb, Ba, and Sb. Heterogeneity produced by a final irregular distribution of lead, barium, and antimony which may reflect the way in which the particle grows. Particles often have air central cavities, which suggest they have been subjected to some form of disturbance during formation, which may explain elemental distribution
III	The least commonly found. They consist of a Pb crust, which surround a homogeneous core of Ba and Sb. It is believed that the Ba and Sb core is in the process of solidifying when it captures Pb vapours from burnt residues and the etched bullet (caused by barrel rifling as the bullet travels down the barrel). The Pb coating may also take on a peeled orange appearance in certain conditions

Wrobel (27) compared the efficiency of a number of different adhesives for the collection of inorganic GSR particles. Fifteen assorted adhesives were investigated, with a variety of criteria being used to assess the suitability of each adhesive. Eight double-sided tapes, three adhesive tabs, two adhesive liquids, a glue stick, and a carbon conductive cement were all tested. Out of all the adhesives tested Sellotape 404 double-sided tape was the medium chosen as the best performer.

Shaffer and Yi (26) compared tape/sticky lifts to swabs (isopropanol as solvent) for the collection of inorganic GSR for scanning electron microscope (SEM) analysis. Results showed that tape lifting was much more effective than swabbing. The average lifting efficiency for tape was 389 particles in total, with 126 being classified as "unique" (the classification of particles is discussed in a later section). The average for swabbing was 60 particles in total with only three particles classified as "unique."

DeGaetano (28) compared three techniques developed for the sampling and analysis of inorganic GSR by SEM with energy dispersive X-ray detection (EDX); tape lift (3 M brand adhesive), glue lift (29), and a centrifugal concentration technique (30). The number of particles of inorganic GSR found on the collection surface in 1 h was used as an indicator of the lift efficiency of each medium. Of the three techniques, tape lifts were reported to be the most effective, being cheap, having good collection efficiency, and performing well in the SEM. A shelf life of at least 6 weeks was reported and samples were shown to be able to withstand being in the high temperatures that may be expected in a crime scene vehicle, for at least 12 h without being effected. However, a number of problems were reported with tape lifts including the large surface area to be searched (dependent on tape size), the requirement to carbon coat samples prior to analysis (something which can now be avoided by using carbon-coated adhesives), and the collection of debris that may mask particles. The sample concentration technique tested, which theoretically should reduce the search area (by repeated centrifugal concentration through high density liquid), was shown to produce highly variable results and be less efficient than tape lifting.

Zeichner (31) also found concentration techniques problematic, concluding that the build-up of debris on filter surfaces was to such a degree that direct observation of a tape or glue lift was preferable.

Wallace and Keely (32) reported a successful concentration technique. Samples were suspended in nonpolar solvent as particles and subjected to a two-stage filtration process firstly removing extraneous materials and secondly retaining small particles including any inorganic GSR. Particles from sizes 0.5 to 2 μm were captured on the second filter. Recovery levels as high as 90% for Pb and 91% for barium were shown. However, as GSR particles may be as large as 100 μm in diameter (9), potentially helpful particles may be lost using this method. This must be seen as a drawback to this particular concentration technique.

Zeichner (33) reported a technique for extracting OGSR from tape stubs following SEM-EDX analysis. Extraction with an aqueous (0.1% w/v sodium azide)/ethanol mix (80/20) at 80°C for 15 min, followed by further extraction with methylene chloride and concentration by evaporation was shown to be the optimal procedure for gas chromatography with thermal energy analysis (GC–TEA) and ion mobility spectrometry (IMS). However, tests on single base powder showed variable results, with recovery levels for NG and DNT ranging from 30% to 90%.

Plasma ashing of tape samples has been reported to be successful for reducing organic material (skin cells, etc.) that may be present on the surface of a tape lift, making GSR particles easier to analyze (34). However, Varetto (35) argued that a combination of contamination by the electron beam of the SEM and oxygen plasma ashing was required to destroy the cells of the skin epidermis, leaving only thin filaments. Plasma ashing alone was reported to be ineffective.

Zeichner (36) reported that in fact skin debris posed little threat to impeding the analysis of inorganic GSR particles using SEM. Particles completely covered in skin cells that could not be detected by the secondary electron image were detected without problem by the backscatter electron image and EDX detector. Although morphological information could not be gained, Zeichner argued that the American Society for the Testing of Materials (ASTM) considered morphology a secondary criterion for GSR identification. This must though be seen as problematic, as particles from ulterior environmental sources have been shown to have elemental compositions that may be easily mistaken for GSR (37–39); in such situations it is the morphological information that is of key importance in correctly determining a particle's origin.

Collecting samples from clothing using tape lifts may also create problems with fibers and other debris. This detritus is likely to be nonconductive and may hold charge during SEM analysis. Carbon/gold coating of the sample may therefore be required which involves extra time and expense (40).

The collection of inorganic GSR particles from hair has been reported as important because of the longevity of particle retention when compared to hands (41). Tape lifts have been reported to be unsuitable for GSR collection from hair (41). However, Zeichner (25) reported that tape lifting was an acceptable method for use on hair (both curly and straight) and no significant difference between tape stubs and more complicated hair comb swab and solvent damp cloth were observed. Maximum collection efficiency was reported at 200–300 dabs (60–120 dabs for hands).

Vacuum Lifts

Vacuuming is used primarily for the collection of GSR from clothing. Speers et al. (42) reported the successful application of a vacuum collection technique for both organic and inorganic residues from items of clothing. Residues were collected from clothes following only one firearm discharge. More organic then inorganic residues were seen. Solid phase extraction (SPE) was used to clean up and concentrate samples in order to maximize the high performance liquid chromatography-pendent mercury drop electrode (HPLC–PMDE) and gas chromatography–mass spectrometry (GC–MS) systems being used. Recovery levels from the vacuum filters used were in the region of 57–78% depending on the compound of interest. Debris and garment fibers present on the filter were reported as a possible cause of reduced extraction efficiency.

Zeichner et al. (43) investigated the vacuum collection of OGSR with two different types of vacuum filter, fiber glass and Teflon. Four solvents (acetone, methylene chloride, ethyl acetate, and chloroform) were also tested for their ability to extract residue collected on the filters. No significant differences between the solvents in their extraction efficiency of the propellant components were found. Collection levels were highly variable, with between 30–100% yields for the same solvent. Methylene chloride was chosen because of the advantage of not dissolving/extracting NC (which is not volatile enough to be analyzed by GC-MS [44]). Teflon filters were shown to have greater collection efficiency. The use of tape lifts on clothing for the collection of inorganic GSR, followed by vacuum lifting to collect organic residues was shown to be an effective technique (43).

Andrasko and Pettersson (45) reported the use of a double filtration vacuum system. A prefilter with a pore size of 20 μm allowed the separation of residue particles from debris and fibers. Residues were collected on the second filter (0.8 μm), and concentrated onto a tape stub for SEM analysis. As with previously discussed concentration methods, the range of filters used in this study could potentially lead to the loss of inorganic GSR particles larger than 20 μm.

Mastruko (40) reported that vacuum lifts from clothing collect particles from the surface of the material but also the depth of the clothing. This was determined to be problematic as it increased the difficulty of interpreting sample analysis, as particles from other shoots (for example a hunter or sports shooter) may be present. In this respect, tape lifting is advantageous as it only lifts particles settled on the surface of a material. However, Andrasko and Petterson (45) reported that tape lifting was not very suitable for the collection of GSR from clothing as the loss of tape stickiness restricted the area that could be sampled and also fibers and other unwanted particles were transferred to the tape making analysis using SEM more difficult.

Swabbing

Twibell et al. (46) assessed the efficiency of eight solvents for the collection of NG from hands followed by GC analysis. Solvent efficiency was determined based on the amount of NG removed from the hands, the amount of interfering material removed from the hands, and the stability of NG within the solvent. Aqueous solvents showed the best recoveries, when thin layer chromatography was used for partial purification. However, NG was degraded by micro-organisms that grew in the solutions. Ethanol was determined to have performed best, with the most complete, stable, and consistent recovery.

Organic solvents are commonly used for the collection of explosive residues as such residues are readily dissolved in them.

However, this causes the problem that many other compounds are also dissolved causing interference issues. Thompson (47) reported that this matrix of interfering compounds may be removed by using water as an extraction agent and adding an additional step, SPE. Water extraction followed by SPE was reported to be an effective process for treating organic explosive residues on cotton swabs for subsequent analysis by liquid chromatography (LC) or GC–MS and fast GC-TEA. When compared to a solvent extraction (acetone) and direct injection method, the water/SPE was shown to be just as effective at recovering organic explosives. Water/SPE also gave much greater selectivity in most cases.

Lloyd and King (48) reported a method by which swabs used to collect explosives and firearm residues were extracted and cleaned up by SPE in the containers issued for the return of samples to the laboratory. Recovery over the range 63–75% was reported. Inorganic GSR particulates also remained on the swab after organic compounds had been extracted and could be subsequently recovered for characterization by SEM by sonication in an organic solvent followed by membrane filtration of the extract. This method was reported to reduce the possible problems of sample loss because of transfer and minimize the chances of cross contamination.

Reardon and MacCrehan (49) compared supercritical fluid extraction (SFE) and ultrasonic solvent extractions (USE) in order to determine whether a reliable quantitative extraction technique for smokeless powders could be achieved. SFE was shown to be unsuitable for quantitative extraction of double base powders, although it has been shown as successful with regard to single base powders. Even after optimization of the extraction process, yields of below 90% with smokeless powder standards were recorded. NG was also shown to readily react with stabilizers under the conditions of SFE. 2-butanol:methanol (1:3) was reported as the most efficient solvent for USE. The most desirable extraction time was determined to be 15 min (handgun powders). For ball type rifle powders, 75 min extractions were recommended. No significant differences between temperature of 0°C and 50°C were reported and therefore 25°C was selected.

Glue Lifts

Glue lifts have been applied to the collection of GSR from hands (29,50). Basu and Ferriss (29) reported that glue lifting was a very usable technique for the collection of GSR from the surface of hands. When compared to tape lifts, glue lifts required less dabs on the skin surface and collected less debris because of the glue being less tacky than tape lifts. This increased the speed of SEM analysis. The glue also contained no elements of high atomic numbers that could potentially interfere with the GSR particle analysis with a SEM.

In contradiction, DeGaetano et al. (28) reported glue lifts to be an ineffective lifting medium. However, a different type of glue lifting planchet was used to that of Basu and Ferriss (29), which could potentially have led to the different results.

Nasal Collection

Schwartz et al. (51) reported the development of a technique for the collection of airborne inorganic GSR particles from human nasal mucus. Samples were collected on a 5 × 5 piece of substrate by normal nose blowing. SEM-EDX was used for sample analysis. Inorganic GSR particles were recorded at times >48-h post firing. This collection method must be seen as promising in terms of determining whether a person has been in the vicinity of a firearm

discharge. However, it may not be of value in trying to determine whether a person actually discharged a firearm, as airborne GSR particles have been shown to take relatively long periods of time to settle (52).

Collection of GSR from Hair

Smoke plumes exiting the breach of a weapon during discharge frequently extend posterior to the face and head (53), with potentially useful GSR being deposited in the hair. A number of methods for the collection of GSR from hair have been reported including a swab and comb method and tape lifting (25,54).

MacCrehan et al. (53) used a fine toothed comb to collect residue samples from head hair. Nearly intact grains of incompletely burned propellant powder residues approaching mm diameters were recovered from handgun firings. Particles smaller than the gaps in the teeth of the comb were also recovered with this method. Difficulties with curly hair were reported because of the fine teeth of the comb. Twenty positive results from 23 tests were reported for human hair wig tests. Tests on shooters show NG positive results for all three different shooters tested. NG and ethyl centralite (EC) were found on the collections from rifles and revolvers. When compounds from unfired powders, exemplar residues (taken from the inside of a cartridge) and residues collected from the hair were compared, it was reported that there was an amount of variation between the unburned powder and hair combed residue but that the exemplar and the combed residue were in good agreement. It was concluded that although EC was detected in some of the residue samples there was not an effective enough extraction for it to be reliably detected using capillary electrophoresis (CE). It was also concluded that a more effective protocol for hair residue collection and analysis would be required to enable reliable detection of stabilizers such as EC that are present in OGSR.

GSR Analysis

Color/Spot Testing

Color/spot tests are most commonly used for the estimation of firing distances (9). However, they can also be used as a rapid test for the presence of GSR and the determination of bullet holes/entrance wounds (55). Such tests have been in use since 1933, when the dermal nitrate or paraffin test was introduced (7,10). Table 4 contains a list of some of the most common color/spot tests used for GSR detection.

The main problem that arises from using spot/color tests is their presumptive nature. However, despite this they are still used in case work in some countries (e.g., Brazil) for determining the presence of GSR (19).

Inorganic GSR Analysis

Neutron Activation Analysis—Neutron activation analysis (NAA) has been used as a bulk analysis method for various elements, which can be found in inorganic GSR. Ba and Sb are the two main elements identified using this method (56–59). However, Cu and Au (60), Ag, Ni, and Co (61) have also been analyzed. Capannesi and Sedda (61) used NAA to examine the trace elements present in lead core, jacketed bullet fragments. With this method, as many as 13 trace elements could be analyzed. As the elements present in bullets can become part of inorganic GSR, this NAA method could theoretically be used for GSR analysis. NAA has also been used for the determination of firing distances (62) and the determination of GSR on the hands of shooters (57,59,60,62,63).

The levels of compounds in samples are usually compared to hand blanks in order to determine if elevated levels of the compounds of interest are present. Ruch et al. (58) reported average levels on hands of nonfirearm firers (130 samples) of Ba and Sb to be 0.05–0.10 μg and 0.01–0.03 μg, respectively. A smaller study (14 samples) by Pillay et al. (60) reported Ba and Sb levels at 0.061 μg and 0.036 μg and Cu and Au levels as 1.085 μg and 0.020 μg. Kilty (63) determined the average Sb and Ba as 0.01 μg and 0.32 μg respectively, although daily fluctuations in concentration levels, especially with Ba were reported.

A number of problems have been reported with NAA; the technique cannot be applied to Pb analysis, samples must be irradiated, which requires a nuclear reactor as a neutron source, trained personnel are required to carry out the analysis procedure (7), and it is also an expensive and time consuming technique (9).

TABLE 4—*Color/spot tests that have been applied to GSR (7,9,10,55,113,114,127,143–148).*

Test Name	Compounds Detected	Additional Notes
Dermal nitrate/ paraffin test	Nitro groups	False positive results may be found by reaction with compounds present in tobacco, leguminous plants, fertilizers, pharmaceuticals, fingernail polish, urine, and as a result of striking a match, tires, and outer garments, solid rocket fuels, pesticides and dyes, pharmaceuticals, veterinary medicines, and the storage preservation of apples. Chlorates, dichromates, iodates, bromates, permanganates, and higher metal oxides may all cause reactions that may lead to false positives. Aleksandar tested 250 persons that had not handled firearms using the dermal-nitrate test. 117 of them produced positive results (46.8%)
Walker test/Griess test	Nitrites	Tests specific for nitrites but not for GSR
Modified Griess test	Nitrites	Test specific for nitrites but not for GSR
Harrison and Gillroy test	Pb, Ba, and Sb	Reported as much more successful than the dermal nitrate test as it produced much fewer numbers false positive results. Test specific for Pb, Ba, and Sb but not GSR
Alizarin red S (ARS)	Ca and other metal ions, including Pb and Ba present in primer residues	None specific and will stain many heavy metal ions including Fe, Ba, Sr, Be, Cd, La, Pb, and U
Sodium rhodizonate test	Pb	Test specific for Pb but not GSR
Marshal and Tewari test	Nitrites	Test specific for nitrites but not GSR
Lunge	NC	Test specific for NC but not GSR
Zincon reagent	Zn and Ti	Applied to Pb free ammunition. Will also react with copper. Test not specific to GSR

GSR, gunshot residue.

Atomic Absorption Spectrometry

Conventional atomic absorption spectrometry (AAS) has been reported sensitive enough for the detection of Pb in GSR samples, but inadequate for Ba and Sb. However, the introduction of electro thermal atomizers (carbon rod, tantalum, and graphite tube furnace) made flameless AAS suitable for the analysis of Ba and Sb in GSR samples (7). Samples are most commonly collected using cotton-tipped swabs and 5% nitric acid (64,65).

Flameless AAS has been reported as a successful technique for the analysis of inorganic GSR as it is both readily available and cost effective (66). It has an advantage over NAA, having excellent sensitivity for Ba and Sb and can be used to detect other elements of interest including Pb (67).

AAS has been applied to the determination of shooting distances, based on concentration patterns of Pb around bullet holes, (68) and the detection of GSR on collection swabs taken from hands by the determination of antimony and barium concentrations (67).

Ravreby (69) reported the use of flame and flameless AAS for the analysis of residues collected from bullet holes. Elements originating from the bullet, case, primer and firearms were analyzed (Cu, Zn, Pb, Sb, Ni, Fe, Ba, K, Sr [from the paint on the bullet tips of tracer rounds] and tin). The results provided a means for identifying the type of ammunition used.

Reed et al. (70) applied AAS to the analysis of GSR on the hands of 112 suicide cases. With threshold levels set at 0.2 μg for Sb and 0.3 μg for Ba positive results were found in 38% of cases. The most important factor in obtaining good results was reported to be the condition of the hands being sampled. Dry clean hand sampled at the scene or protected by paper bags during transport provided the most GSR consistent results. Time delays, weapon characteristic, and body location were all secondary considerations.

Cooper et al. (64) used flameless AAS for the analysis of Sb and Ba levels on the hands of persons in close proximity to firearm discharges, using porcine skin as a substitute. Threshold levels were set at 0.05 μg/mL for Sb and 0.50 μg/mL for Ba. In most cases, levels of Sb and Ba were only slightly elevated at a distance of 1 foot from discharge.

Koons (65) discussed a number of problems associated with AAS. Incomplete extraction of Sb from collection swabs was shown to be an issue; even with an optimized method only 60–70% was extracted, compared to a nearly complete extraction of Pb and Ba. Further problems arising from the extraction process included variable absorbance-time profiles for Sb and the enhancement of Ba absorbance caused by various matrix constituents. Heavily soiled swabs that potentially prevented effective extractions were also reported as problematic. Ashing methods have been proposed to overcome this problem; however, Koons also reported good recoveries from soiled swabs using a 10% nitric acid sample digestion method (65).

The worth of AAS in term of GSR analysis was further brought into question by Aleksandar (3), who criticized the method on the basis of the large number of false negative results it has been shown to produce (about 40%).

Inductively Coupled Plasma

Inductively coupled plasma (ICP) is a bulk analysis technique that is usually used to analyze trace amounts of Pb, Ba, and Sb in primer residues (9).

Koons (71) reported the use of ICP with atomic emissions spectroscopy (AES) for the analysis of Ba levels in swabs. The method was determined more successful than the previously used AAS because of the lack of interference from common swab constituents, a wide linear dynamic range, and good precision and accuracy (limits of detection for barium, defined as three times the baseline noise level were determined to be 0.002 μg/mL for AAS and 0.0008 μg/mL for ICP-AES. Relative standard deviations of multiple measurements of the same solution having a barium concentration of 0.05 μg/mL were reported as 5% to 10% for AAS and less than 1% with ICP-AES).

Koons (72) also reported the use of ICP-MS for the analysis of residues originating from primers. This method was chosen because of superior detection limits in comparison with ICP-AES and graphite furnace-AAS (GFAAS) (0.052, 0.020, and 0.14 ng/mL for Sb, Ba, and Pb, respectively) and faster analysis times than GFAAS. The use of MS allowed the detection of several isotopes for each of the elements of interest (Pb, Ba, Sb). ICP-MS was reported as potentially useful in several areas; the determination of levels of additional elements which may be associated with the handling of a firearm or ammunition component, or elements which may be present in specific ammunitions, such as strontium in some nontoxic primers, cobalt in Nyclad™ bullets, or copper, nickel, or zinc in jacketed bullets. ICP-MS also provided the possibility of some sourcing of primer derived Pb by isotope distribution (72).

Steffen et al. (73) reported the use of ICP-MS in conjunction with SEM-EDX. Following SEM analysis, GSR particles on the adhesive stub were dissolved using 2% v/v HNO_3 in nanopure water. Pb isotope ratios were determined for eight lead-based primer ammunitions. Ratio comparisons were shown to be promising and some of the lead-based primers could be distinguished from one another. However, such a method was not recommended for everyday case work, being both destructive and time consuming.

Zeichner (74) investigated lead isotope ratios in inorganic GSR using multiple-collector ICP-MS (MC-ICP-MS). Lead isotope levels were reported as potentially useful in specific types of scenarios, for example in a shoot out situation where several firearms and ammunition are discharged. It was reported to be possible to link a bullet hole to the firearm that discharged the bullet. However, problems were reported with "lead memory." Even after thorough cleaning of a firearm, some amounts of Pb from previous firings remained and contributed to the residues of following discharges. This memory was reported to lower the levels of association between residues collected from the firearm barrel and residues from the fired ammunition (bullet and case).

Sarkis (75) employed sector field, high-resolution-ICP-MS (SF-HR-ICP-MS) to determine the levels of Pb, Ba, and Sb in residue samples, allowing identification of these elements at concentrations as low as 1 ng/mL. Ternary graphs were used to better visualize the results and allow direct comparisons between the relative percentages of the three compounds of interest in different samples. Such graphs were reported to provide strong evidence concerning the origins of sample components on hands (GSR or environmental).

Scanning Electron Microscopy

Scanning electron microscope equipped with an X-ray detector (wavelength dispersion, or energy dispersion [EDX]) is the most commonly used method of analyzing inorganic GSR. Information on the principles of SEM and the development of this technique for inorganic GSR analysis are presented elsewhere and will therefore not be included here (7).

The advantage that SEM and X-ray detection has over bulk analysis techniques is the ability to analyze individual particles of

inorganic GSR both morphologically and chemically. Bulk elemental techniques are problematic as the total sample levels of specific compounds (usually Pb, Ba, and Sb) must be above predetermined thresholds for a positive result to be concluded. Therefore, because of the small levels of GSR that are often encountered in case work, many cases involving such analysis techniques are determined inconclusive. A further problem with bulk analysis lies with the fact that the elements being analyzed are not exclusively found in GSR and therefore elevated levels of such compounds on the hand of a suspect must not be taken as conclusive evidence that they have discharged a firearm (76).

Aleksandar (3) comments that the ability of SEM-EDX to morphologically and chemically analyze samples makes it possible to determine that particles could have only originated from the discharge of a firearm and no other way. Despite this SEM-EDX cannot determine whether a person discharged a weapon on any given occasion.

Brozek-Mucha (8) used SEM-EDX to examine the differences between inorganic airborne and cartridge case residues. Both similarities and differences were observed in the two residue types. It was reported that the composition of a residue is influenced by two main factors:

• The direction of movement of the expanding products of burning propellant at the stage of internal ballistics.

• The kind of materials that were applied to construct the gun and ammunition.

Lebiedzik and Johnson (77) reported the use of specific morphological and elemental indicators within inorganic GSR particles to differentiate between the firearm used to discharge a round, the case and bullet material of the ammunition and ultimately differentiate one type of ammunition from another. Twenty-one elements, within 60 types of ammunition were assessed by individual particle analysis using SEM, with automated image analysis and X-ray micro analysis. In ideal conditions, it was reported possible to differentiate between residues produced by firearms with barrels made of stainless steel, carbon steel, and titanium. Ammunition caliber could be estimated and information on the materials used for the bullet and case of a cartridge could be determined.

Brozek-Mucha and Jankowicz (78) compared the residues collected from the hands of six different shooters using six different lead-based ammunition types. Purely visual inspections of the results allowed differentiation between three ammunition types. Statistical comparison approaches to the results were also evaluated. The Wilcoxon rank sum method was shown to be unsuitable for such an application. T-Kendall rank correlation coefficients and the R Spearman methods proved suitable for the evaluation of mutual relationships among frequencies of occurrence of certain chemical classes of primer residue. One ammunition type could be distinguished from the others and two other ammunition types could be differentiated from one another. It was concluded that further research would be required in order to determine the most suitable statistical analysis method for such applications.

Steffen et al. (73) reported the use of chemometric classification of inorganic GSR. Residues from eight lead-based and seven lead-free primer ammunitions were analyzed. Regularized discriminant analysis presented the ability to enter unknown samples into a distribution model consisting of various classes. Differentiation of lead-based primers was achieved and successful classification rates varied from 51–86%, with a misclassification risk of 9%. Problems in classification occurred because of the very similar qualitative compositions of lead-based primers, which were inorganically consistent. Therefore, only the semi-qualitative information gained from the EDX could be used for classification purposes. Lead-free primers proved much easier to classify as both qualitative and quantitative differences could be used for classification. The risk of misclassification was determined to be as low as 2.4%, with particles being correctly assigned 90–100% of the time.

Miyauchi et al. (79) used SEM-EDX to analyze smokeless powders from 20 different ammunition types. Cur, Si, K, S, Al, Ca, Fe, Cl, Ba, and Zn were all found. It was suggested that propellant powders may contribute to inorganic residues to a greater degree than was previously thought.

Collins et al. (80) reported the analysis of glass fragments in inorganic GSR in .22 caliber ammunitions. Particles containing glass fragments were found on the hands of shooters, with sizes varying from 2–100 μm. Particles were described as having morphologies indicative of the fusion of Pb and Ba compounds with glass of varying viscosities. Such particles were reported to provide a further area of interest which may allow discrimination between inorganic GSR and environment sources of GSR like particles. Only break linings were determined to be a potential source of similar particles.

One of the biggest drawbacks to SEM analysis was the tedious and time-consuming task of manual particle searching within samples. The introduction of automated GSR search programs has eliminated such a problem, reducing the total search time for a sample, freeing the SEM operator to do other tasks and reducing the number of potential human error sources (81).

Germani (82) investigated the effects of changing SEM-EDX variables on search results, in an attempt to compile a set of standard settings for the analysis of inorganic GSR particles. It was reported that no one set of standard operating procedures could be determined, as the required search parameters between different samples are likely to be variable. For example, a sample collected from a suspected shooter a number of hours after firearm discharge is likely to contain smaller sized particles than a sample collected immediately and a search program would need to be tailored accordingly.

Zeichner and Levin (54) reported the success rates in case work of inorganic GSR analysis using SEM-EDX on samples from hands, hair, and clothing over a period of 6 years. The overall success rate for finding GSR was reported at about 10%, which was determined to increase the probative value of positive results, as it diminished the danger of accidental sample contamination leading to positives. In 39% of positive cases, only one particle of inorganic GSR was discovered.

The analysis of the inorganic GSR can provide a great deal of information. However, in some cases where inorganic residues are not present in a sample, or are only present in relatively low levels, OGSR components may be used. Even when inorganic residues are present at large enough levels to be useful, the analysis of OGSR may provide complementary and additional information that may strengthen the probative value of a sample (10) and potentially provide additional means of differentiating between GSR and environmentally sourced residues.

Organic Analysis Methods

The analysis of OGSR up until the year 1997 has been reviewed in detail previously (10). Therefore work up until 1997 will be briefly covered, while more attention will be directed toward work published after this point. Techniques covered previously where no additional noteworthy articles were found during the compilation of this review are not included.

Gas Chromatography

Gas chromatography coupled with a number of different detectors has been applied to OGSR analysis, including flame ionization, TEA, electron capture, and MS. TEA has been most commonly used for OGSR analysis.

Andrasko et al. (83) reported the use of GC–TEA and GC–MS for the detection of various constituents and degradation products from smokeless powders in the barrels of firearms after test shootings. GC–TEA was employed for the analysis of samples and GC–MS was used to identify some of the compounds of interest.

Zeichner (43) assessed the value of GC–TEA and GC–MS alongside IMS for the analysis of OGSR. GC–TEA showed a good level of sensitivity for the compounds of interest. Limits of detection were: 0.2 ng for NG, 0.05 ng for 2,4 DNT and 0.05 ng for 2,6 DNT. The considerably lower sensitivity of GC–TEA for NG compared to DNT was reported as being because of thermal decomposition of the NG in the GC column, creating a nonlinearity of the NG peak height as a function of concentration, in particular when approaching the limit of detection. An increase in column length (15–30 m) yielded two peaks for NG, a second smaller peak being determined as 1,2-GDN, a thermal decomposition product of NG. This decomposition was reported as a drawback to sensitivity. However, the presence of two peaks increased the probability of identifying NG using GC–TEA. Two GC–MS systems were used to analyze standard mixtures, details of which can be found elsewhere (43). Neither of these systems was optimized for explosive analysis; however, limits of detection for the compounds of interest were reported at several nanograms. Not one case work example which was found positive for NG by GC–TEA could be confirmed by GC–MS. It was therefore reported that GC–MS was not sensitive enough for the majority of real life cases. IMS was shown as a good complementary technique for GC–TEA for the detection OGSR.

Kirkbride et al. (84) reported the use of GC–MS for the detection of explosive compounds, many of which may be found in firearm ammunition. As reported by Zeichner (43) the limits of detection for many of the compounds were several nanograms (NG 10 ng, TNT 10 ng, RDX 10 ng, Pentaerythritol tetranitrate [PETN] 0.5 ng).

A potentially useful method for the combined analysis of inorganic and organic GSR by SEM-EDX, and GC–TEA and IMS has been reported (33). Samples were collected onto aluminum stubs and inorganic analysis carried out by SEM-EDX. Samples were then extracted from the stub with a water/ethanol mixture (80/20) at 80°C with sonication for 15 min, followed by further extraction with methylene chloride and concentrated by evaporation for analysis of OGSR by GC–TEA and IMS (33).

Gas chromatography techniques suffer from the inability to analyze NC, as it is not sufficiently volatile. Its introduction (as a major component) to a GC column may also accelerate the deterioration of the column's performance (43). Nitrate esters which are frequently found in GSR are also incompatible with the usual condition of GC systems because of their thermal instability. Such compounds will decompose on improperly prepared columns. PETN has been reported to suffer particularly from this (10). It has also been reported that GC is unsuitable for the analysis of stabilizers such as N-nitrosodiphenyleamine because the temperatures involved in such systems may cause denitrosation to diphenylamine (DPA) (10,16).

High Performance Liquid Chromatography

HPLC has been applied to OGSR analysis with various types of detector. NG, 2,4-DNT, and DPA have been analyzed with electrochemical detection (85). Lloyd (86) applied amperometric detection at a mercury drop electrode, and coulometric detection at porous carbon electrodes to the analysis of NG and DPA and size exclusion chromatography/PMDE for NC.

Speers et al. (42) reported the use of HPLC–PMDE and GC–MS in the analysis of organic propellant powder residues combined with SEM-EDX for the analysis of inorganic residues. HPLC–PMDE was used to detect NG and 2,4-DNT (GC–MS for DPA, EC,MC). The analysis of residues for organic components was shown to yield more positive results than inorganic analysis. The combined analysis of both kinds of residue was reported to produce the most useful/powerful results. HPLC–PMDE and GC–TEA have also been used in conjunction with one another (87).

Laza et al. (14) reported the use of LC–MS/MS for the quantitative analysis of common propellant powder stabilizers in OGSR. Residues were collected from the hands of shooters with cotton swabs (ethanol, water 75/25% v/v). An SPE technique was used for the concentration and purification of the samples. DPA, N-nitrosodiphenylamine (N-NDPA), 4-nitrodiphenylemaine (4-NDPA), 2-nitrodiphenylemaine (2-NDPA), Akardite II (AKII), Methylcentralite (MC), and EC were all resolved from standard mixtures using multiple reaction monitoring, allowing the presence of precursor ions in the samples to be determined and therefore the identification of the corresponding target compounds. The limits of detection were reported to be 0.29, 0.27, 0.34, and 0.21 nmol/L for DPA, N-NO-DPA, 4-NO2- DPA, and 2-NO2-DPA, respectively. EC as well as MC had a limit of detection of 0.07 nmol/L and a limit of detection of 1.3 nmol/L was determined for AK II. Converted into the equivalent amount of target compounds injected onto the column, the limits of detection were of 5, 6, 20, 27, 32, 34, and 115 µg injected onto the column for EC, MC, 2-NDPA, N-NDPA, 4-NDPA, DPA, and AK II, respectively. The method was determined to be very sensitive for the centralites (EC and MC) and poor AKII results were linked to problems with recovery levels during SPE. The analysis of samples collected from the hands of shooters confirmed that the method was suitable for routine analysis of OGSR. However, it was reported that further investigation into the longevity of OGSR components on the hands of shooters was required before the true value of such a technique could be truly determined.

Xu et al. (88) developed a combination of three HPLC–atomic pressure ionization (API)–MS systems for the analysis of 21 nitro-aromatics, nitramines, and nitric esters, some of which may be found in firearm ammunition. Limits of detection for most compounds were reported to be between 0.012 and 1.2 ng. The method was also determined to be superior to a previously developed HPLC-thermospray ionization–MS, allowing a greater screening range and an increase in selectivity by a factor of ten. The HPLC–API–MS method was reported as specific, stable, and reproducible.

Mathis and McCord (89) reported the application of a reverse-phase LC–electospray ionization–MS for the comparison of organic additives in several smokeless powders. The method was determined useful in the analysis of compositional variations in smokeless powders.

Cascio et al. (90) compared HPLC and micellar electrokinetic capillary chromatography (MEKC) for their ability to analyze OGSR. Results indicated that both reverse phase HPLC and MEKC with UV detectors were capable of resolving standard mixtures of organic components of smokeless powders. Statistical analysis using the Spearman's rank correlation test showed the separate patterns from the two systems were highly correlated. Because of the broader range of analytes detected, better suitability for diode array detection, and lower operation costs MEKC, diode array UV detection looked particularly interesting as a screening technique.

Capillary Electrophoresis

CE is an important analytical technique, which can provide rapid, high-resolution separations of complex mixtures. Although electrically neutral compounds such as those found in OGSR cannot be separated by conventional CE, micellar electrokinetic capillary electrophoresis (MECE) permits the separation of such substances (10).

Reardon et al. (91) reported the use of CE for the analysis of NG, DPA and N-NDPA and EC in seven reloading propellant powders (one single base and six double base). Both bulk samples and single particles were analyzed. Compositional variations between particles were shown to be significant in some cases. Because of potential blending in finished propellant powders, it was reported that individual particles may not be sufficient in representing the sample bulk. The ratio of propellant/total amount of stabilizer (p/s) was shown to be a more robust way of linking residues to powders. In five of six of the double base powders, the p/s ratio of ten particles taken from the bulk was in agreement with the p/s of the bulk. Forty-nine of 60 particles analyzed gave reliable comparisons to bulk samples. It was also reported that the combination of quantitative and qualitative information with details of particle size, shape, and color could help associate unknown powders or OGSR with a known sample.

MacCrehan et al. (92) also used the propellant to stabilizer ratio in order to associate handgun fired OGSR with unfired propellant powders. Of seven powders analyzed, four could be easily differentiated. However, when visual examinations of particle morphologies were combined with the results of the P/S ratios all seven of the powders could be reliably differentiated. It was reported that a much larger sample of ammunitions would however need to be evaluated before such observations could be considered as generally acceptable.

Northorp (93,94) assessed the application of MECE to case work. SEM and MECE were used together to provide information on both inorganic and organic substances present within propellant powders and residue samples. Samples were collected on adhesive stubs and analyzed using SEM before being extracted with methanol for MECE analysis. The limits of detection for 13 characteristic organic gunpowder components (2,3-DNT, 2,4-DNT, 2,6-DNT, 3,4-DNT, 2-NDPA, 4-NDPA, Dibutylphthalate, Diethylphthalate, DPA, EC, MC, NG, N-NDPA) ranged from 0.9–3.8 pg for standard solutions. One hundred commercial propellant powders were analyzed in order to create a reference library. It was concluded that the detection of characteristic organic gunpowder components was a strong indication of OGSR with little or no likelihood that they resulted from environmental exposure. MECE found detectable residues on all ammunition types examined except for .22 caliber. Levels of OGSR were also shown to vary between firings carried out under the same conditions, leading to the conclusion that residue deposition mechanisms and collection efficiency significantly affected the outcome of OGSR analysis. In case work both inorganic and organic compounds were identified on some samples collected from the hands of suspect shooters and the clothing of victims. MECE was concluded to be a potentially valuable tool in the examination of OGSR.

CE has also been applied to the analysis of inorganic GSR. Hopper and McCord (95) reported the use of capillary zone electrophoresis for the analysis of inorganic ion profiles from smokeless powders. Seven smokeless powders were analyzed in both unburned and burnt states. Results demonstrated that ionic profiles could be used for characterizing smokeless powders.

Romolo et al. (96) reported the worth of CE for the detection of nitrate and nitrite ions in GSR. The used method proved simpler, cheaper, and faster than the traditional approaches to GSR analysis based on AAS, AES, ICP-MS, or SEM. Hair and skin samples from a victim shot in the head were also successfully analyzed for the presence of nitrite and nitrate. This preliminary study indicated that CE offered a sound potential for forensic GSR investigation as a screening technique prior to more expensive and time-consuming analytical methods. It was also reported that CE could be useful in the detection of residues produced by lead-free ammunition, where the evidential value of analytical results obtained with SEM can be poor (96).

Morales and Vazquez (6) developed a CE method for the simultaneous detection of both inorganic and organic substances in GSR, with 11 organic and 10 inorganic compounds being analyzed. The method did suffer from poor detection limits for some compounds, including Ba and Sb. A preconcentration method did however allow OGSR to be detected. It was concluded that two separate runs for inorganic and organic residues might be a better option, or alternatively the analysis of inorganic compounds with CE and organic compound with another technique, for example GC.

Other Techniques That Have Been Applied to GSR Analysis

Tong et al. (97) reported the use of a tandem MS/MS method for the quantitative analysis of DPA and its nitrated derivatives in smokeless gunpowders. Detection limits for DPA, N-NDPA, and 4-NDPA were 1.0, 0.5, 2.5 ng/mL, respectively. The method was determined highly selective and sensitive.

Meng and Caddy (98) reported the use of a fluorescence method for the detection of EC in GSR. The limit of detection for EC in spiked swabs and standard solutions was 5 ng and 1 ng, respectively. Three of six test firings were determined to contain EC. One of these swabs was collected 180 min following firearm discharge.

Time-of-Flight Secondary Ion Mass Spectrometry (TOF–SIMS) has been applied to both organic and inorganic GSR analysis. Cumbaros et al. (99) reported TOF–SIMS analysis as a useful method of detecting inorganic residues. The technique was determined to have a number of advantages over SEM–EDX including lower detection limits, the elimination of signal overlaps, the ability to analyze both organic and inorganic compounds, near surface depth profiling allowing elemental distribution to be monitored as a function of depth and elemental and compound mapping from raw data. However, it was reported that because of a lack of high-resolution imaging capabilities, such as those provided by SEM back scatter electron detection, SEM-EDX should still be considered the method of choice for routine detection and counting of potential GSR particles.

Mahoney et al. (100) applied TOF–SIMS to the analysis of both organic (such as EC, dibutyl phthalate, and NC) and inorganic substances (such as potassium nitrate, potassium perchlorate, and sulfur) in smokeless and black powders (and black powder substitutes). However, because of the high-vacuum conditions inside the instrument the technique was reported to be unsuitable for more volatile components such as NG.

X-ray microfluorescence has been reported as an applicable method for the analysis of inorganic GSR (101–103). However, despite several advantages over SEM-EDX being reported (larger potential scanning area [20 × 20 cm] and direct visualization of target materials e.g., fabrics), the inability to analyze particles smaller than 10 μm led to the conclusion by Flynn (101) that X-ray microfluorescence could not replace SEM-EDX as the method of choice for inorganic residue analysis. Berendes (103) reported that although X-ray microfluorescence could not replace

other commonly used techniques in GSR investigation it does offer a good supplement, especially with nontoxic primer residues.

Niewohner and Wenz (104) applied focused ion beam systems to GSR investigation. It was reported that scanning ion microscopy allowed particles to be cross sectioned, revealing interior morphologies that could be used for the identification of ammunition manufactures. Some particles were shown to be solid throughout, in which case such cross sectional information was not useful. However, where morphological details were present, different ammunition types produced characteristics that could be used to differentiate between them. However, the study only investigated four types of ammunition, and although each of these ammunitions did produce unique interior morphologies an investigation of a much larger number of ammunition types would be required to confirm the true worth of such analysis.

Pun and Gallusser (105) reported the worth of macroscopic observations of the morphological characteristics of propellant powders, investigating 181 cartridges of different calibers. The color and shape of the powder grains were shown to be an effective way of linking partially burnt powders from discharged rounds back to unburned samples. On these observations alone, a number of potential ammunitions matches for any residue sample could be determined. The correct source ammunition was always within those selected as possibilities for any given sample. However, a quantity of particles was required for accurate analysis of both color and shape. In real situations, factors such as weapon type, spatial distribution of GSR and problems with contamination arising from previously discharged cartridges were all reported as potential hindrances to the successful application of such observations.

Environmental Sources of GSR-Like Particles and Their Influence on GSR Classification and Analysis

Environmental Sources of Inorganic GSR-Like Particles

The possibility that there could be other sources of GSR-like particles is very important. If it were found that any other process or activity could produce particles with indistinguishable morphological and/or compositional characteristics to those of GSR then the weight of such particles as forensic evidence would be greatly reduced.

The analysis of GSR with bulk analysis techniques such as NAA or AAS does not take into account the morphology of individual particles being analyzed and therefore the possibility of false positive results is much greater. The three elements most commonly analyzed using these techniques are Pb, Ba, and Sb, which may be picked up from a number of environmental sources. Pb particles can be found in the emissions of leaded petrol, plumbing materials, battery plates, solder, glass, and paints. Sb is found in several alloys, often with Pb, and its oxides are used as a fire retardant in cotton and polyester blend fibers. Ba is found in paint, car grease, and barium sulfate from paper (7).

The application of SEM and X-ray detectors allows the morphology and elemental composition of individual particles to be determined. Wolten et al. (21) produced a list of particle types that were considered unique (characteristic) and consistent to inorganic GSR particles based on particle composition (Table 5), size, and morphology. In the majority of cases, 70–100% of particles analyzed were spheroid, with smooth or fuzzy surfaces, scaly or covered with smaller spheres. The remaining (rarely above 30% depending on ammunition) particles had irregular morphologies. The vast majority of particles had diameters less than 5 μm.

TABLE 5—*Classification of inorganic gunshot residue particles (21).*

Unique	Characteristic
Pb,Sb,Ba	Pb,Sb
Ba,Ca,Si with trace S	Pb,Ba
Sb,Ba	Pb
	Ba if S is absent or only trace levels

The aforementioned particles may also contain the following Si, Ca, Al, Cu, Fe, S, P (rare), Zn (only if Cu>Zn), Ni (rare and only with Cu and Zn), K Cl and Sn (in obsolete ammunition).

A number of studies have since been published that have investigated whether or not inorganic GSR particles of such compositions are indeed unique to firearm discharge.

Wolten et al. (106) carried out SEM-EDX analysis on samples taken from the hands of people working in areas that were considered to be possible sources of GSR-like particles. Areas with the most potential for producing GSR-like particles were described as critical occupations. These were made up of industrial and commercial operations involving metals or compounds of Pb, Ba, and Sb. Sectors that involved melting and/or vaporization of such elements were of particular interest. An overview of their results is presented in Table 6. The possibility of particles being created in some of the areas investigated that were of similar morphological and compositional structures to inorganic GSR was observed. It was reported that such particles would be classified as indicative of GSR using the "formal" classification system presented in the Aerospace report (21). Such particles were seen in residue samples collected from stud guns. It was concluded that individual particles should be considered within the context of all other particles within the sample in which they are found if the potential for misinterpretation was to be minimized. The occupation of the person from which a sample is taken was also considered as potentially relevant to the interpretation of samples (106). None of the samples analyzed contained particles of Pb, Ba and Sb or Ba and Sb, which were still considered to be unique to GSR (21). The results of this study showed the possibility that particles from other sources could be confused with GSR. This must be seen as an important finding, especially if a sample being analyzed has only a small number of particles present.

Wallace and McQuillan (107) carried out a more detailed study of particles produced by stud guns, analyzing (SEM-EDX) the primer and residue compositions of six types of stud gun ammunition. All of the particles found were between 1–12 μm in diameter and both spherical and irregular morphologies were observed. Morphologies were consistent with vaporizing or melting of the compounds involved. No crystalline structures were seen in any of the collected particles (which would indicate the particles would have been created in an environment not consistent with those involved in a firearm discharge [21]). It was concluded that the physical characteristics of the samples analyzed were indistinguishable from GSR. The results from Wallace and McQuillan (107), together with their past experience led to a conclusion that the original classification system for inorganic GSR particles (Table 5) should be modified to take into account particles generated from cartridge operated tools (Table 7). This classification system was stated to be only applicable to brass cased, Pb, Ba and Sb primed and Pb, Ba primed ammunition types. A simplified version of this system was adopted in the ASTM standard guide for inorganic GSR analysis by SEM/EDX (108,109).

Wallace and McQuillan (107) also suggested that the term GSR should be reworked in relation to the classification system, suggesting the term "primer residues" should be adopted as this would

TABLE 6—*Results of investigations into environmental sources of inorganic GSR-like particles (106).*

Vocation/Activity	Findings
Stud guns	Two particles consistent with GSR particles were found in samples from one brand of stud gun cartridge (Remington); these were small spheroid particles containing Ba, Pb, Zn, and Cu. It was only the abundance of larger non-GSR-like particles that led to sample being determined as not GSR. Thirteen particles consistent with GSR were found in samples collected from Omark brand cartridges. A typical composition of these particles was Cu and Zn (3:1) with Pb, Fe, and Ba (major) and Si, K, Cu, Ca, and Zn (minor). Again it was only the abundance of non-GSR-like particles that meant samples were not determined to be GSR
Cap guns	No GSR-like particles were found in these samples. Sb was seen but there was no Pb or Ba present. Both crystalline (Sb, S [major] Cl and P [minor]) and spheroid (Cl, Ca and P and Ca, P, Zn and K [major] Sb, Cl and S [minor]) morphologies seen
Blanks	No GSR particles found with two blank types Winchester 0.22 with black powder and Winchester Western 0.38 with smokeless powder
Lead smelting	Pb and Sb particles were found in samples, 50% were spheroid and 50% irregular (determined not consistent with GSR). None of the samples were determined to be GSR
Car brake mechanics	Two particles consistent with GSR (Pb and barium with other element associated with GSR). The first contained Fe, Cu, S, Si, Ba, and Pb. The second, Pb, Fe, Cu, Si, Ba, and Cl. The elemental proportions were however determined not to be that of GSR and the particles were dismissed
Pb acid battery assemblers	Particles consistent with GSR were found. These included the following 1. Pb and Sb. 2. Pb, Sb, Fe, Si, Zn. 3. Pb, Ca, Ba, Fe, and Zn. 4. Pb, Fe, and Ba. 5. Pb, Cu, Si, and Zn. 6. Pb, Ca, Ba, Fe, and Zn. However by taking into account all particles present, samples were discounted as GSR. The size distribution of particles was also deemed inconsistent with GSR samples
Car mechanics, exhaust fitters, and environmental lead.	Nothing relating to GSR particles was found

GSR, gunshot residue.

incorporate all percussion-activated charges based on Pb, Ba, Sb compounds, therefore including cartridge-based tools.

In addition to this proposed change, Wallace and McQuillan (107) also suggested a number of other considerations that should be taken into account when analyzing inorganic GSR. Spent cartridges, they argued may be used for chemical composition comparisons but not for morphological checks, as particles found inside cartridge cases will be different from those ejected from the firearm. The analysis of inorganic GSR by comparing samples collected from hands or other sources to cartridge cases/ammunition is a move away from a "formal" approach in which samples are interpreted following the rules of a formal general interpretation system, to a "case by case" or "specific" approach (an approach concluded as the most appropriate for GSR analysis wherever possible by Romolo and Margot [7]). Further more, individual particles within a sample should be considered in relation to all the others that are present. The fewer the number of particles that are present

TABLE 7—*Inorganic gunshot residue classification system proposed by Wallace and McQuillan (107).*

Unique	Indicative*
Pb, Sb, Ba	Ba, Ca, Si[†]
Sb, Ba	Pb, Sb
	Pb, Ba
	Sb (with or without S)
	Ba[†]
	Pb
	Pb, Sb, and Ba absent[‡,§]

*Indicative particles listed in order of decreasing significance.
[†]S absent or acceptable as trace only when Ba present at a major level.
[‡]Any of the above may also include some or all of the following: Al, Ca, S, Si at major, minor or trace levels; Cl, Cu, Fe, K, Zn (only if Cu also present and Zn:Cu < 1) at minor or trace level; Mg, Na, P at trace levels only.
[§]Particles containing no Pb, Sb, or Ba may be considered indicative if they are composed entirely of the elements in [‡] above and are accompanied by other types of indicative particles.

in a sample the more difficult it is to discriminate between tool and GSRs (107).

Zeichner and Levin (110) agreed with Wallace and McQuillan (107) that GSR-like particles could be found in residues from cartridge-operated tools.

Garofano et al. (111) followed on and expanded upon the work of Wolten et al. (106), investigating the same vocations/activities that were flagged as potential sources of GSR-like particles. One hundred and seventy-five samples from the hands of persons undertaking these activities of interest were analyzed using SEM-EDX. Their results corroborated those of Wolten et al. (106) and confirmed that particles of Pb, Ba, and Sb were unique to inorganic GSR (although they were found in samples collected from stud gun operators). However, Garofano et al. (111) found particles of Sb and Ba in samples relating to car repair and maintenance activities. Even taking into account the morphology of these particles there were some that were hard to distinguish from genuine inorganic GSR Sb, Ba particles. It was therefore suggested that Ba, Sb particles (with iron present) no longer be classed as unique to inorganic GSR (21,106,107) and instead be seen as characteristic (unless found with Pb, Ba and Sb particles in which case they would still be seen as unique). Ba and Sb particles in the absence of Fe should be seen as borderline particles. Garofano et al. (111) also concluded that both the elemental and the morphological compositions of the particles being analyzed are important if a correct interpretation of a particle's source is to be achieved.

Mosher et al. (37) investigated the possibility of fireworks and pyrotechnic devices producing particles that resembled inorganic GSR. Residues from pyrotechnic devices only available to professional display organizers as well as those available to the general public were analyzed with SEM-EDX. Results showed that it was possible to find firework residue particles that could be mistaken for inorganic GSR. Samples collected from the hands of the professional firework display organizers contained some Sb and Ba particles that were morphologically (at least externally) similar to inorganic GSR particles (spheroid, noncrystalline, and between 0.5–5 μm in diameter). Such particles could be classified as unique to

GSR using the original guidelines (21). Other Sb and Ba particles, with irregular morphologies and much larger in size (up to 40 μm) were also present. These particles, although not considered typical of GSR could still be classified as such under the original formal classification system (21).

The analysis of the fireworks available to the general public showed that all of the unburned powders contained both Pb and Ba, with two also containing Sb. Residue samples indicated that the fireworks without antimony could produce Pb and Ba particles (seven were found), which had been reported in inorganic GSR produced from some .22 caliber rim fire primer units (10). Residues from one of the firework types that also contained antimony had particles that consisted of Pb, Ba, and Sb, an elemental composition classified as unique to GSR (21). Many of these particles were irregular and shaped between 10 and 35 μm in size and containing other elements. These other elements and the morphology of the particles would, it was determined, allow a trained GSR examiner to differentiate such particles as coming from some form of pyrotechnic device. However, a few particles contained Mg, which had been reported as being present in some inorganic GSR particles (11). Therefore, a potential for such particles to be wrongly identified as GSR has been shown to exist. There was also two Pb, Ba, and Sb particles found that did not contain Mg, which had the potential to be mistaken for GSR.

Torre et al. (38) carried out an investigation into car brake linings and their potential for producing inorganic GSR-like particles. It was argued that Sb, Ba particles should not be considered unique (as suggested by Wallace and McQuillan [107]) as they can be found readily in urban areas rich in metallic particulates such as those exposed to road traffic. Brake linings appeared to be an obvious source of GSR-like particles, as many contain lead sulfide, antimony sulfide, and barium sulfate in different combinations and while breaking disks can reach temperatures in excess of 600°C, reaching up to 1500°C in friction spots on the surface, similar temperature to those that occur during a firearm discharge (38). The results of tests on 40 different types of brake discs showed that it was possible to obtain Pb, Ba, and Sb particles from the wear of brake linings. Such particles were subdivided by Torre et al. into two groups, "clean" and "unclean." "Unclean" particles contained the three elements of interest, but with other elements or concentrations of elements that were inconsistent with inorganic GSR. "Clean" particles were described as those collected from car front brakes and new brake linings composed of Ba, Sb or Pb, Sb, Ba, which were of a similar size to inorganic GSR with elements and elemental concentrations that were consistent with inorganic GSR. Such particles as these were found mainly in new break linings, while those collected from car front breaks often had iron present within them at major levels. Results were in agreement with other recent publications in that Sb, Ba particles could not be considered "unique" to GSR (111); further more, the observation of "clean" Pb, Ba, Sb particles also suggested that such particles could not be considered "unique" to GSR. The only two compositions that were not found were: "clean" Ba, Sb particles with no sulfur, or sulfur at trace levels and "clean" Pb, Sb, Ba particles with Ba and/or Sb levels higher than lead. These particle types could still in theory be considered "unique"; however, it was suggested that the "unique" classification be dropped altogether in favor of the "more prudent" "consistent" categorization. It was reported that in order to discriminate between primer discharge residue and environmental particles, analysis of particle morphology must be re-evaluated (38,111). Particles identical to primer discharge residue in elemental composition may be found, but never of a typical shape. The only reliable particle would be of the "ideal" morphology, spherical or globular, with a surface either perfectly smooth, pitted with craters, or coated with rounded and smooth nodules, but never, even in part dusty or rough (38). Torre et al. (38) agreed with Wallace and McQuillan (107) suggesting that the term GSR should be dropped and replaced by either "primer discharge residue" or "residues of the detonation of a mixture of Pb, Ba, and Sb compounds." Finally, it was concluded that if ammunition that was recovered from a suspect/scene is not available for comparison with residue samples then any conclusions should be drawn with extreme caution (38).

Cardinetti et al. (39) used X-ray mapping techniques to analyze a number of nonfirearm sources of possible inorganic GSR-like particles. Break lining samples from 42 different makes and models of car were taken. Hand (palm and back) samples from car electricians, motor mechanics, and pyrotechnic and firework operators were also collected. A number of different genuine firearm cartridge residues were also analyzed to re-evaluate compositions and morphologies that may be encountered. Results confirmed that it was possible to find particles containing Pb, Sb and Ba and Sb, Ba in nonfirearm-related samples. Two particles containing Pb, Sb, Ba were found in samples collected from car brake linings, along with 214 Sb, Ba particles. Each of the samples collected from the hands of workers contained Sb, Ba particles. The analysis of residue from within cartridge cases showed that it was also possible to find irregular particles in GSR samples, which could in theory lead to false negative results for actual inorganic GSR. It was reported that X-ray mapping techniques, which can analyze the internal elemental structures of particles, could reduce the probability of mistakes in inorganic GSR analysis occurring, because true GSR particles and those from other environmental sources have very different internal element distributions. Inorganic GSR particles showed homogenous elemental distribution, whereas environmental particles, which had been created under much lower temperatures and pressures, showed a layered structure of individual elemental plaques (39).

The findings of Cardinetti et al. (39) are interesting; however, there are two areas which may be problematic. Firstly, tested residues were taken from inside cartridge cases, which although valid in some respects (chemical comparisons) may not accurately represent the particles that may leave a firearm during a discharge (107). Secondly, it is stated that by looking at the internal structures of particles, it is possible to differentiate GSR from environmental particles by analysis of internal elemental distributions. However, work carried out by Basu (22) showed that a number of different internal structures may be found among GSR particles, including ones in which elements are layered. Therefore, it may not be possible to accurately determine the source of a particle using this method.

Environmental Sources of Organic Compounds Found Within Organic GSR

If the analysis of the organic compounds present in GSR is to improve the evidential value of GSR analysis, it must be determined to what extent these compounds of interest could come from other sources.

Lloyd (112) listed a variety of environmental sources of the stabilizer compound DPA; the surface of apples, tires, and outer garments were all shown to be possible sources. Solid rocket fuels, pesticides, dyes, pharmaceuticals, and veterinary medicines are also formulations known to contain DPA (113).

Direct swabbing of the skins of apples, pears, bananas, and citrus fruit have been reported to produce chromatogram peaks at the same retention time (HPLC) as DPA (114). Grapefruit, oranges, and pears have also been reported to produce peaks at the same

retention time as EC. Items of clothing, shoe polish, and many rubber products all produced interfering peaks (114). Leggett and Lott therefore commented that HPLC analysis of organic compounds is not an infallible method of determining the presence of OGSR in any given sample (114).

*Inorganic GSR from Ammunitions with Lead Free/
Nontoxic Primers*

Another aspect of residue analysis that must be taken into account is ammunition types that do not have primers that contain Pb, Ba, and Sb compounds. A number of studies into GSR produced by lead-free/nontoxic ammunitions have been published and their findings are summarized in Table 8.

Gunaratnam (12) analyzed the inorganic residue particles formed by Sintox lead-free ammunition (cal .375 Magnum) with SEM-EDX. Characteristic particles were reported to contain Zn and Ti, with spheroid morphologies. Morphology was determined as the only way of truly discriminating such particles from environmental sources (irregular particles could not be used). However, a much smaller number of spheroid particles were found in residue samples than in those collected from cartridges with Pb, Sb, and Ba based primers. Ti and Zn are found in paint pigments but never as major elements and never with inorganic GSR characteristic spherical morphologies. However, scenarios involving burning paint and the particles formed have not been fully investigated (12).

Harris (115) analyzed GSR from CCI Blazer lead-free ammunition using SEM-EDX. The only significant metallic element found in the residue particles was Sr. Particles were mostly 0.5–10 μm (some up to 35 μm) and spherical, as would be expected from Pb,Ba,Sb primers. Traces of Ba were seen in some residues, even though not stated by the manufacturer as being included; however, natural deposits of Sr are often found in association with Ba. It was reported that Sr particles with spherical morphologies could result from exposure to flares and fireworks, limiting the evidential value of such particles.

Oommen and Pierce (116) analyzed a number of lead-free ammunition types with SEM-EDX, reporting different elemental compositions for each (Table 8). Similarities were observed between Federal Ballisticlean™ and sparkler pyrolysis residues, although it was determined that a trained analyst should not have problems differentiating them. K, Ca, Mg, and Na salt, present in some of the residues are soluble in water and therefore would be expected to be dissolved by human sweat. These elements would not be expected to persist on shooters' hands for as long as those elements that are not water soluble.

Martiny et al. (19) characterized the inorganic GSR elements present in Magtech Cleanrange® ammunition (SEM-EDX). Two varieties of 9 mm Luger, .40S&W, .380 Auto and .38 special types were analyzed. None of the residue particles analyzed could be considered unique to GSR using the ASTM formal classification system (109). Cleanrange® produced predominantly irregular-shaped particles, which would also be considered atypical using formal classification. The elements present in residues for Cleanrange® may also be found in association with automobile components, lubricants, and combustibles. For this reason, it was reported that elemental compositions of primer resides from Cleanrange® products cannot be used to confirm firearm discharge. The use of alternative analysis techniques, for example organic residue analysis, was reported to be a potential means of improving the evidential value of residues from such ammunitions.

Lead-free and nontoxic ammunitions are problematic when it comes to determining whether residues originated from firearm discharge or environment sources using a "formal approach" (7). Therefore, a "case by case" assessment procedure would be more appropriate for such cases as these (7,111,117). Additional research into possible environmental sources of particles like those produced by lead-free and nontoxic ammunitions is also required before the evidential value of such particles can be truly determined (116).

The Effects of Changing Ammunition Types on the Composition of GSR

The effects of changing ammunition types in any given firearm have been shown to produce inorganic particles of exceptional compositions. Such particles may differ greatly from the "classic" criteria for the identification of inorganic GSR (using a "formal" identification approach) (118). Even the thorough cleaning of a firearm has been shown not to remove all traces of previous residues and lead to a kind of memory effect within the firearm barrel (74,115). Therefore, a "case by case" basis for analysis should be adopted where possible if the potential misinterpretation of the evidence is to be minimized.

TABLE 8—*Elemental compositions of residues from various lead-free/nontoxic ammunitions (12,19,116,117).*

Ammunition	Primer Composition	Residue Particles
Sintox (0.375 Mag)	Diazodinitrophenol (DDNP), tetracene, zinc peroxide, titanium metal powder, nitrocellulose	Zn–Ti Zn–Pb Ti–Zn–Pb–Ba
CCI Blazer® Lead free (0.38 SPL+P)	Tetracene, DDNP, smokeless powder, strontium nitrate	Sr Sr–Ba
Winchester Winclean™ (9 mm and .45ACP)	Copper and zinc (primer cup), DDNP, potassium nitrate, boron, nitrocellulose (MSDS contents)	Cu–Zn K–Al–Si–Na Al–Na,Ca or Mg
Remington/UMC LeadLess™ (9 mm and 0.45ACP)	Copper and zinc (primer cup), DDNP, barium, tetracene (MSDS contents)	Cu–Zn Al–Si–K (trace Na) Al–Si–K (Na or Ca)
Federal Ballisticlean™ (9 mm and 0.45ACP)	Copper and zinc (primer cup), tetracene, barium nitrate, aluminum, nitrocellulose, nitroglycerine	Cu–Sn Al,Si,Ba,K (Na)
Speer lawman Cleanfire™(9 mm and 0.45ACP)	Copper, Zinc, nickel, DDNP, tetracene, strontium nitrate, nitrocellulose, nitroglycerine	Cu Cu-Zn Sr–Al/Si or Cl
Cleanrange 1st generation (9 mm Luger)	DDNP, tetracene, nitrocellulose, strontium nitrate, gum tragacanth (patent claim)	Sr (plume) Sr–Na–K–Fe (hand residue)
Cleanrange 2nd generation (9 mm Luger, 0.40S&W, 0.380 AUTO and 0.38SPL)	DDNP, tetracene, nitrocellulose, potassium nitrate, aluminum powder, ground glass, gum tragacanth (patent claim)	Al,Si,Ca (plume) Al–K–Si–Ca-Fe (S in 0.38 SPL) (hand residue)

Lebiedzik and Johnson (77) reported that in cases where a variety of ammunition types are fired by one firearm, descriptive indicators (21 elements) may be difficult to interpret because of the combined influence of ammunitions (77).

Torre (20) collected samples of lead-free and nontoxic GSR by swabbing the insides of spent cartridges; these were compared to samples collected from hands and targets. Variable pressure SEM-EDX was used for analysis. Results showed that previous shots of lead-based ammunition within a given firearm could influence the composition of residues produced from subsequently discharged lead-free rounds. Because of the poor evidential value of the results obtained from solely using this technique on lead-free and nontoxic GSR, it was recommended that further analysis techniques should be employed to gain as detailed a picture of composition of any given GSR sample as possible.

MacCrehan et al. (119) reported that compositional analysis of residues from firearms in which ammunition types had been changed showed only trace amounts of organic compounds (from propellant powder) from previous firings in the first shot and none in subsequent shots.

Distribution and Transfer of GSR Following a Firearm Discharge

Once a firearm has been discharged, it is essential to understand the distribution of GSR in relation to the shooter and their surroundings, not only in terms of sample collection but also minimization of contamination/secondary transfer within samples which could potentially lead to an inaccurate conclusion as to whether a subject has been involved in handling or discharging firearms. An understanding of these processes may also aid in the accurate interpretation of analysis results (7).

Schwoeble and Exline (9) investigated the evolution of residue plumes from a number of different weapon types (pistols, revolvers, rifles, and shotguns), using high-speed photography. A great amount of variation in plume evolution patterns between firearms was described (Table 9). Blow back or drift of residue plumes toward the chest, shoulder, face, and hair was observed. Cartridge ejection was shown to be a factor in the dispersion of GSR (contradictory to Wolten et al. [21]). Ejection plumes were then shown to spread out in all directions and therefore be subject to the influences of any air turbulence in the vicinity of discharge (9).

Schwoeble and Exline (9) also investigated inorganic GSR particle fall out. Rates were shown to vary widely, based upon the physical characteristics (size, shape, and particle density) of particles.

Carreras et al. (120) also used fast photography to examine the ejection of residues from the muzzle end of a 9 mm parabellum, a 9 mm short gun, a .38 revolver, and a 7.62 carbine rifle. Residues leaving a firearms muzzle were divided into three categories: Primary—residues leaving the barrel prior to the projectile, secondary—residues leaving the barrel at the same time as the projectile, and tertiary—residues leaving the barrel after the projectile. This process will affect the amount of residue available for collection dependent on muzzle to target distance. At close ranges, the gunpowder combustion residues reach the target and cause soiling. The characteristics and the amount of GSR on a target are dependent on the distance of the shot. Contact wound shots may yield little or no external GSR as the residues are forced into the bullet hole (120).

Basu et al. (50) investigated the evolution of inorganic GSR from the rear of firearms, described as "trigger blast." It was demonstrated that deposits of residue on the hands of shooters came almost exclusively from GSR emanating from the rear of the firearm (as suggested by Wolten et al. [21]) and not the muzzle (overlap of muzzle and trigger residues was observed in unclean firearms). This was shown to occur in both open and closed breach weapons. Residues on hands were deposited by blasting because of firearm discharge and not from the deposition of airborne particles. These results were in agreement with Schlesinger (121) who reported that wind velocity has no effect on the deposition of residue particles on the hands of a shooter. Therefore, for a nonfiring individual to have residue present on their hands, they would have to be in very close proximity to a weapon discharge. It was also hypothesized that deposits of GSR that have been found on forward facing portions of shooters' bodies could also be deposited by the firearm discharge blast and not the fallout of airborne particulates (50).

Fojtasek et al. (122) investigated the distribution of GSR particles surrounding a pistol shooter. Experiments carried out in a closed environment showed that "unique" (Pb, Sn, Ba, Si, Ca and Pb, Sn, Sb, Ba [123]) inorganic GSR particles could be found at distances up to 10 m from the shooter. The maximum quantity of residue particles were found at 45° to the right in front of the shooter (several 1000 particles were found in this area); however, several hundred particles were also found in front of the shooter and to their right. In an open environment, distribution patterns were the same, but particle concentrations were 10 times lower in magnitude. This demonstrates a significant influence of climatic conditions on the amount of residue that might be available for collection. There were no particles found after a distance of 6 m (122). There was no indication of any observable differences in the distribution of particles with different sizes as a function of distance. The majority of the particles were less than 3 μm in diameter.

Gerard et al. (124) reported that GSR can travel much further than suggested by Fojtasek et al. (122), detecting inorganic GSR at distances up to 18 m. These particles were reported to have been carried such distances in association with the projectile. It was also demonstrated that inorganic GSR could spread outward laterally from the firearm to at least a distance of 3 m.

Work carried out by Bergman et al. (125) demonstrated that inorganic GSR particles can be consistently found on the bottom of discharged bullets, including those which have been severely deformed on impact or undergone other severe conditions. Many of these particles were shown to have a strong adhesion to the bullet, even after 20 min in an ultrasonic bath, only "loose" particles were removed. The high temperatures and pressures in the firearm on discharge appeared to cause particles to fuse to the base of the bullet. Such particles may prove very useful in determining which weapon a specific bullet was discharged from during an incident

TABLE 9—*Characteristic residue distribution plume patterns by firearm type (9).*

Firearm	Residue Distribution
Smaller caliber semi-automatic weapons (high/forward cartridge ejection)	Sometimes plumes concentrations to the front of the fingers, but in most cases the plumes tend to follow in the direction of the ejected cartridge
Larger caliber revolvers	Widespread plumes
Larger caliber semi-automatic pistols	Smaller more compact plumes (compared to larger revolvers)
Shotguns and rifles	Consistent area of plume concentration in the crook of the support arm

involving multiple shooters, especially if said bullet is greatly deformed, rendering visual comparisons to test firings difficult/impossible (125).

Fojtasek and Kmjec (52) carried out an investigation of inorganic GSR deposition rates. A 9-mm and a 7.65-mm pistol and a .38 special revolver were tested. Results showed that each weapon type had a different time period of maximum deposition. Time periods were dependent on the firearm and the caliber of the ammunition used. With a pistol discharge, the possibility of airborne contamination was reported to exist within ca. 8 min after discharge. With a revolver, the time could be as long as 10 min after discharge. The relationship between particle size and deposition time was also investigated. As would be expected (126) larger particles (6–10 μm) were deposited first, followed by those with smaller diameters (1–5 μm). Different firearms and ammunitions produced varying levels of "unique" particles in different size ranges (the 9 mm produced most "unique" particles in the 1–5 μm range, while the 7.65 were in the range of 6–10 μm).

Andrasko and Pettersson (45) investigated the possibilities of nonfiring persons, present in locations where a firearm discharge had occurred being contaminated with inorganic GSR. A summary of the results from the experiments can be seen in Table 10. Contamination was shown to be possible when a person walked through the particle cloud of a firearm discharge and not just as a consequence of staying close to a shooter. This conclusion is particularly important when considering the evidential value of any given GSR sample.

The papers mentioned earlier present evidence to suggest that there is a real possibility that contamination of persons in or entering a room/location where a firearm has been discharged could occur.

Shooter Activities Post Firearm Discharge and Effects on GSR Loss

It is of importance to understand the potential longevity of GSR particles on hands, other skin, clothing, and other materials from which samples may be collected. Not only is this information useful when determining whether a sample should be taken (in cases where the subject is suspected of discharging a firearm days, weeks or months prior to apprehension), but also in interpreting the results of any given sample analysis.

Gunshot residue deposits on a person are continuously lost as a result of normal activities and as a consequence of this it is very difficult to generalize a time period during which GSR may be retained (10).

Inorganic GSR retention on the hands of shooters has been shown to vary greatly following normal activities. Maximum recovery times ranging from 1–48 h have been reported for particles on shooters' hands (40,59,63,123,127–133). Rapid decreases in particle numbers have been shown to occur within 1–3 h post discharge (63,128,131,133). The rapid loss of inorganic GSR particles from hands is a great disadvantage in terms of collection, but on the other

hand increases its value if found, as this indicates a very short time frame between firearm discharge and sample collection (40).

Specific everyday actions have been shown to radically affect the amount of inorganic GSR on shooters' hands. Kilty (63) investigated the effects of various activities on inorganic GSR levels on hands using bulk analysis, antimony, and barium. Washing hands with soap and water then drying on a paper towel effectively removed all traces of inorganic GSR. Rinsing hands for 3 sec under low-pressure water removed substantial amounts of residue, as did wiping hands on clothing (63), bringing levels close or equal to those levels found on the hands of nonshooters (133). Rubbing the hands together transferred residue from one hand to the other and placing hands in pockets removed residues (63); however, residues could be subsequently detected within the pocket (133).

Lloyd (86) reported that NG could be detected up to 7 h post firing on a shooter's hands, face, and throat, whereas it could be detected on clothes worn without restrictions for up to 5 h only. Douse (134) reported contradictory results showing a 0.5-h detection limit for NG on skin. Northrop (94) reported that recoverable organic GSRs may not persist on skin for more than 1 h.

The longevity of GSR on clothing has been reported to be generally much greater than on skin. Particles have been found days or even weeks after firearm discharge (40). Jane et al. (135) reported that clothing removed and stored post firing as opposed to clothing that was continually worn retained organic residues (NG, NC, and DPA) for a much longer period of time (residue found on the stored cloths the following day, the worn garments retained residue for up to 6 h). Residues deposited on a cotton sheet placed 1 m in front of a revolver that discharged five rounds remained detectable for up to 2 months, if the sheet was undisturbed. These experiments suggest that the loss of GSR is caused by physical activity rather than compound degradation (135). Further evidence for this theory was provided by Douse (134), who reported that inorganic GSR particles may be found on the hands of suicides for up to 48 h and potentially longer, a much greater time period than would normally be expected on a shooter's hands. Inorganic GSR particles have even been recovered from the garments of a badly decomposed man who was discovered after being outside for 2 months, further strengthening the case for physical activity being the major cause of particle loss from a subject.

Zeicher and Levin (54) reported the detection of inorganic GSR in samples collected from hair up to 24 h post discharge, if the hair was not washed. However, this time was seen to be greatly reduced in case work, where inorganic GSR was reported to persist on hands and hair for about the same period of time (hands 2.7 h and hair 3.3 h). However, other factors such as shootings occurring outside and wind affecting the deposition rate of GSR in hair were suggested as potential reasons for this (54).

Machine washing or brushing of clothing has been reported to decrease considerably the amount and the density of inorganic GSR remaining on clothing (136).

TABLE 10—*Contamination experiments and particles found (45).*

Experiment	Particles Found
Walking into a contaminated area. Subject walked in a shooting room in which no shooting had occurred that day. The floor was cleaned every second day	Two particles observed in lower part of the coat
Clothing hung in a shooting gallery: coat hung 2 m behind shooter who fired over 300 rounds (0.22)	Hundreds of gunshot residue and bullet particles observed
Clothing present in the same room as a shooter. Two persons (A 1 m to right and B 4 m behind) in a room with a shooter. Four 357 magnum shots fired. Subjects then walked across area in front of the shooter while leaving the room	A: 6 particles found B: 8 particles found

Studies into Contamination Relating to the Arrest and Transfer of Suspects

Gialamas et al. (137) investigated the levels of GSR that were present on the hands of firearm carrying U.S. police officers at the end of their shifts. Of the 43 officers that were tested, only 3 of them had any "unique" particles on their hands, about 7% of the total. None of the officers were found to have more than one unique particle present on their hands. There were no GSR particles found on 25 of the officers in the study (about 58%). Gialamas et al. (137) commented that although the potential for secondary transfer contamination was present, the low empirical numbers of GSR particles found on these nonshooting officers suggest the potential for this occurrence was relatively low. However, it must be taken into account that GSR particles are lost from the surface of the hands relatively rapidly (depending on activities undertaken); therefore, taking samples at the end of shift may not be the most representative method of assessing transfer potential.

Pettersson (138) showed that around 25% of samples collected from a selection of Swedish police vehicles contained six or more inorganic GSR particles. Samples collected from crime scene investigators also showed positive results in 25% of cases. One sample had as many as 16 inorganic GSR particles (none had handled a firearm for at least 12 h previous to being sampled).

Berk et al. (139) also looked at the possibility of secondary transfer within the police. Two hundred and one samples were collected from law enforcement vehicles and detention facilities in order to assess the probability of inorganic GSR present in these locations being transferred to those persons in detention. Both "unique" and "consistent" inorganic GSR particles were recorded. Of the 201 samples, 173 had no unique particles, while the other 23 contained 56 (two tactical vehicles, 34 table type surfaces, 20 restraining bars). A second collection of samples suggested that there was no accumulation of particles over time. A blank study of persons reputed to have had no primary contact with firearms also presented one case of a "unique" particle being found. Berk et al. (139) stated that although the possibility of unique GSR particles being transferred did exist the low number of unique particles observed suggested the potential for secondary transfer was relatively low.

Berk et al. (139), Pettersson (138), and Gialamas et al. (137) have shown that the potential for secondary GSR transfer does exist within the police. Berk et al. (139) and Gialamas et al. (137) both also commented that the potential for this occurring was relatively low. However, the possibility of even one particle of GSR being transferred to a person who is to be tested for GSR must be seen as important, as such a small quantity of GSR has been used in court as evidence of a person discharging a firearm (4). A survey in 1990 by DeGaetano et al. (140) showed that 41% of laboratories (two surveys of 200 US labs, 1st 51.0% response, 2nd 71.5% response) considered one "unique" particle to be conclusive evidence of GSR. Technically, the presence of one "unique" particle (SEM-EDX) does confirm the presence of GSR; however, it is the opinion of the expert as to the particle's significance that is of paramount importance (140). Another survey of 50 labs in 1996 by Singer et al. (141) showed only one lab considered finding one particle of GSR enough to indicate a person had discharged or been in the vicinity of the discharge of a firearm (one indicated two particles were sufficient and all other stated that their criteria were under review or that interpretation was dependent on the particle type found). Even though it would seem that there has been a general move away from one particle being seen as much of a significant indicator, the existence of a secondary transfer potential of one particle must still be seen as relevant to the interpretation of analysis results. However, the level of significance such a particle holds as a piece of evidence is in the hands of the expert witness and it is an expert's interpretation of any given sample that is of primary importance.

Summary

A review of various aspects of GSR has highlighted a number of areas worthy of consideration. The use of a "formal" classification system has been shown to be problematic. The introduction of numerous lead-free and nontoxic ammunitions as well as ammunitions that do not produce "characteristic" or "indicative" residues, combined with the potential for the misinterpretation of particles from environmental sources means that such an approach must be applied with great care. Particles produced by firearms through which numerous different ammunition types have been discharged may also produce particles that do not fit into a "formal" system.

The adoption of a "case by case" approach to GSR analysis must be seen as preferable. This is in agreement with Romolo et al. (7). The comparison of samples collected from a victim, suspect, or crime scene to firearms, bullets, or cartridge cases has been shown to be an effective approach within a "case by case" framework. The application of statistical evaluation and the construction and utilization of databases have also been shown to be promising approaches to the evaluation of results.

The analysis of both inorganic and organic residues has been shown as a promising method of gaining as much information about any given sample as possible. A combination of these techniques with microscopic or even macroscopic analysis of particle/grain morphologies would be even more favorable. Therefore, this must be seen as the most ideal approach to sample analysis. Having as much information as possible must be seen to increase the probability of the accurate interpretation of results and also increase the evidential value of any given sample.

With regard to the analysis of trace elements/compounds within both organic and inorganic residues, the development of increasingly sensitive and selective analytical techniques has increased the analyst's ability to gain a truer picture of the composition of such residues. With a large number of potential substances that may be present in ammunition, further research into the applicability of such analytical techniques to the analysis of GSR is required.

The interpretation of the results of any sample analysis by an expert witness is incredibly important. A further study into the guidelines which are applied to the interpretation of GSR by independent laboratories/experts would be incredibly valuable in terms of assessing the levels of consistency within the field.

References

1. Povey De, Coleman K, Kaiza P, Hoare J, Jansson J. Homicides, firearm offences and intimate violence 2006/07 (supplementary volume 2 to crime in England and Wales), 3rd edn. London: Home Office Statistical Bulletin, 2008.
2. Mejia R. Why we cannot rely on firearm forensics. New Sci 2005 (2527):6.
3. Aleksandar I. Is there a way to precisely identify that the suspect fired from the firearm? Forensic Sci Int 2003;136(Suppl. 1):158–9.
4. O'Neill S Gunshot particle that helped to convict Jill Dando's murderer 'should be discounted'. Times Online 2007, November 6 2007.
5. Warlow TA. Firearms, the laws and forensic ballistics. United Kingdom: Routledge, 1996.
6. Morales EB, Vazquez ALR. Simultaneous determination of inorganic and organic gunshot residues by capillary electrophoresis. J Chromatogr A 2004;1061(2):225–33.

7. Romolo FS, Margot P. Identification of gunshot residue: a critical review. Forensic Sci Int 2001;119(2):195–211.

8. Brozek-Mucha Z. Comparison of cartridge case and airborne GSR—a study of the elemental composition and morphology by means of SEM-EDX. X-Ray Spectrometry 2007;36(6):398–407.

9. Schwoeble AJ, Exline DL. Current methods in forensic gunshot residue analysis. Boca Raton, FL: CRC Press, 2000.

10. Meng HH, Caddy B. Gunshot residue analysis—a review. J Forensic Sci 1997;42(4):553–70.

11. Wallace JS. Chemical aspects of firearm ammunition. AFTE Journal 1990;22(4):364–89.

12. Gunaratnam L, Himberg K. The identification of gunshot residue particles from lead-free Sintox ammunition. J Forensic Sci 1994;39(2):532–6.

13. Maloney RS, Thornton JI. Color tests for diphenylamine stabilizer and related compounds in smokeless gunpowder. J Forensic Sci 1982;27(2):318–29.

14. Laza D, Nys B, Kinder JD, Mesmaeker AKD, Moucheron C. Development of a quantitative LC-MS/MS method for the analysis of common propellant powder stabilizers in gunshot residue. J Forensic Sci 2007;52(4):842–50.

15. Davis TL. The chemistry of powder and explosives. New York, NY: Wiley, 1943.

16. Espinoza EO, Thornton JI. Characterization of smokeless gunpowder by means of diphenylamine stabilizer and its nitrated derivatives. Anal Chim Acta 1994;288(1–2):57–69.

17. Druet L, Asselin M. A review of stability test methods for gun and mortar propellants, I: the chemistry of propellant ageing. J Eng Mater 1988;6:27–43.

18. Curtis NJ. Isomer distribution of nitro derivatives of diphenylamine in gun propellants: nitrosamine chemistry. Propellants, Explosives, Pyrotechnics 1990;15(5):222–30.

19. Martiny A, Campos APC, Sader MS, Pinto MAL. SEM/EDS analysis and characterization of gunshot residues from Brazilian lead-free ammunition. Forensic Sci Int 2008;177(1):e9–e17.

20. Torre C, Mattutino G. Gunshot residue from lead-free ammunition: inorganic vs. organic analytical techniques. Forensic Sci Int 2003;136(Suppl. 1):150–1.

21. Wolten GM, Nesbitt RS, Calloway AR, Loper GL, Jones PF. Final report on particle analysis for gunshot residue detection, report ATR-77(7915). Segundo, CA: The Aerospace Corp, 1977.

22. Basu S. Formation of gunshot residues. J Forensic Sci 1982;27(1):72–91.

23. Wolten GM, Nesbitt RS. On the mechanism of gunshot residue particle formation. J Forensic Sci 1980;25(3):533–45.

24. Burnett B. The form of gunshot residue is modified by target impact. J Forensic Sci 1989;34(4):808–22.

25. Zeichner A, Levin N. Collection efficiency of gunshot residue (GSR) particles from hair and hands using double-side adhesive tape. J Forensic Sci 1993;38(3):571–84.

26. Shaffer DK, Yi K. A comparison of particle transfer efficiencies of two collection methods for the identification of gunshot residue on fabric surfaces using scanning electron microscopy-energy dispersive spectrometry. Scanning 1999;21(2):99–100.

27. Wrobel HA, Millar JJ, Kijek M. Comparison of properties of adhesive tapes, tabs, and liquids for the collection of gunshot residue and other trace materials for SEM analysis. J Forensic Sci 1998;43(1):178–81.

28. Degaetano D, Siegel JA, Klomparens KL. A comparison of 3 techniques developed for sampling and analysis of gunshot residue by scanning electron-microscopy/energy dispersive X-ray analysis (SEM-EDX). J Forensic Sci 1992;37(1):281–300.

29. Basu S, Ferriss S. A refined collection technique for rapid search of gunshot residue particles in the SEM. Scan Electron Microsc 1980;1:375–84.

30. Ward DC. Gunshot residue collection for scanning electron-microscopy. Scan Electron Microsc 1982;3:1031–6.

31. Zeichner A, Foner HA, Dvorachek M, Bergman P, Levin N. Concentration techniques for the detection of gunshot residue by scanning electron microscopy/energy dispersive X-ray analysis. J Forensic Sci 1989;34:312–20.

32. Wallace JS, Keely RH. A method for preparing firearms residue samples for scanning electron microscopy. Scan Electron Microsc 1979;2:179–84.

33. Zeichner A, Eldar B. A novel method for extraction and analysis of gunpowder residues on double-side adhesive coated stubs. J Forensic Sci 2004;49(6):1–13.

34. Sild EH, Pausak S. Forensic applications of SEM/EDX. Scan Electron Microsc 1979;2:185–92.

35. Varetto L. The use of plasma ashing on sample for detection of gunshot residues with scanning electron microscopy and energy-dispersive X-ray analysis (SEM/EDXA). J Forensic Sci 1990;35(4):964–70.

36. Zeichner A. Is there a real danger of concealing gunshot residue (GSR) particles by skin debris using the tape-lift method for sampling GSR from hands? J Forensic Sci 2001;46(6):1447–55.

37. Mosher PV, McVicar MJ, Randall ED, Sild ED. Gunshot residue—similar particles produced by fireworks. Can Soc Forensic Sci J 1998;31(3):157–68.

38. Torre C, Mattutino G, Vasino V, Robino C. Brake linings: a source of non-GSR particles containing lead, barium, and antimony. J Forensic Sci 2002;47(3):494–504.

39. Cardinetti B, Ciampini C, D'Onofrio C, Orlando G, Gravina L, Ferrari F, et al. X-ray mapping technique: a preliminary study in discriminating gunshot residue particles from aggregates of environmental occupational origin. Forensic Sci Int 2004;143(1):1–19.

40. Mastruko V. Detection of GSR particles on clothing of suspects. Forensic Sci Int 2003;136(Suppl. 1):153–4.

41. SEM/MPA Firearms discharge residues. London, UK: Metropolitan Police Forensic Science Laboratory; 1980.

42. Speers SJ, Doolan K, McQuillan J, Wallace JS. Evaluation of improved methods for the recovery and detection of organic and inorganic cartridge discharge residues. J Chromatogr A 1994;674(1–2):319–27.

43. Zeichner A, Eldar B, Glattstein B, Koffman A, Tamiri T, Muller D. Vacuum collection of gunpowder residues from clothing worn by shooting suspects, and their analysis by GC/TEA, IMS, and GC/MS. J Forensic Sci 2003;48(5):961–72.

44. Zitrin S. Post explosion analysis of explosives by mass spectrometric methods. J Eng Mater 1986;4:199–214.

45. Andrasko J, Pettersson S. A simple method for collection of gunshot residues from clothing. Sci Justice 1991;31(3):321–30.

46. Twibell JD, Home JM, Smalldon KW, Higgs DG, Hayes TS. Assessment of solvents for the recovery of nitroglycerine from the hands using cotton swabs. J Forensic Sci 1982;27(4):792–800.

47. Thompson RQ, Fetterolf DDD, Miller ML, Mothershead II RF. Aqueous recovery from cotton swabs of organic explosives residue followed by solid phase extraction. J Forensic Sci 1999;44(4):795–804.

48. Lloyd JBF, King RM. One pot processing of swabs for organic explosives and firearm residue traces. J Forensic Sci 1990;35(4):956–9.

49. Reardon MR, MacCrehan WA. Developing a quantitative extraction technique for determining the organic additives in smokeless handgun powder. J Forensic Sci 2001;46(4):802–7.

50. Basu S, Boone CE, Denio DJ, Miazga RA. Fundamental studies of gunshot residue deposition by glue-lift. J Forensic Sci 1997;42(4):571–81.

51. Schwartz RH, Zona CA. A recovery method for airborne gunshot residue retained in human nasal mucus. J Forensic Sci 1995;40(4):659–61.

52. Fojtasek L, Kmjec T. Time periods of GSR particles deposition after discharge—final results. Forensic Sci Int 2005;153(2–3):132–5.

53. MacCrehan WA, Layman MJ, Secl JD. Hair combing to collect organic gunshot residues (OGSR). Forensic Sci Int 2003;135(2):167–73.

54. Zeichner A, Levin N. Casework experience of GSR detection in Israel, on samples from hands, hair, and clothing using an autosearch SEM/EDX system. J Forensic Sci 1995;40(6):1082–5.

55. Tugcu H, Yorulmaz C, Karslioglu Y, Uner HB, Koc S, Ozdemir C, et al. Image analysis as an adjunct to sodium rhodizonate test in the evaluation of gunshot residues—an experimental study. Am J Forensic Med Pathol 2006;27(4):296–9.

56. Rudzitis E, Kopina M, Wahlgren M. Optimization of firearm residue detection by neutron activation analysis. J Forensic Sci 1973;18(2):93–100.

57. Rudzitis E, Wahlgren M. Firearm residue detection by instrumental neutron activation analysis. J Forensic Sci 1975;20(1):119–24.

58. Ruch RR, Buchanan VP, Guinn VP, Bellanca SC, Pinker RH. Neutron activation analysis in scientific crime detection. J Forensic Sci 1964;9:119–32.

59. Krishnan SS. Detection of gunshot residue on the hands by neutron activation and atomic absorption analysis. J Forensic Sci 1974;19(4):789–97.

60. Pillay KKS, Jester WA, Fox HA. New method for the collection and analysis of gunshot residues as forensic evidence. J Forensic Sci 1974;19(4):768–83.

61. Capannesi G, Sedda F. Bullet identification: a case of a fatal hunting accident resolved by comparison of lead shot using instrumental neutron activation analysis. J Forensic Sci 1992;37(2):657–62.

62. Krishnan SS. Firing distance determination by neutron activation analysis. J Forensic Sci 1967;12(4):471–83.

63. Kilty JW. Activity after shooting and its effects on the retention of primer residue. J Forensic Sci 1975;20(2):219–30.

64. Cooper R, Guileyardo JM, Stone IC, Hall V, Fletcher L. Primer residue deposited by handguns. Am J Forensic Med Pathol 1994;15(4):325–7.

65. Koons RD, Havekost DG, Peters CA. Analysis of gunshot primer residue collection swabs using flameless atomic-absorption spectrophotometry; a reexamination of extraction and instrument procedures. J Forensic Sci 1987;32(4):846–65.

66. Can M, Uner HB, Koc S, Tok M, Disbudak M. Determination of hand deposited gunshot residue obtained from shootings carried out with handgun cartridges produced by Turkish machinery and chemistry foundation using flameless atomic absorption spectrophotometer. Forensic Sci Int 2003;136(Suppl. 1):147.

67. Koons RD. Flameless atomic-absorption spectrophotometric determination of antimony and barium in gunshot residue collection swabs—a collaborative study. Crime Lab Dig 1993;20(1):19–23.

68. Krishnan SS. Firing distance determination by atomic absorption spectrophotometry. J Forensic Sci 1974;19(2):351–6.

69. Ravreby M. Analysis of long range bullet entrance holes by atomic absorption spectrophotometry and scanning electron microscopy. J Forensic Sci 1982;27(1):92–112.

70. Reed GE, McGuire PJ, Boehm A. Analysis of gunshot residue test results in 112 suicides. J Forensic Sci 1990;35(1):62–8.

71. Koons RD, Havekost DG, Peters CA. Determination of barium in gunshot reside collection swabs using inductively coupled plasma-atomic emission-spectrometry. J Forensic Sci 1988;33(1):35–41.

72. Koons RD. Analysis of gunshot primer residue collection swabs by inductively coupled plasma-mass spectrometry. J Forensic Sci 1998;43(4):748–54.

73. Steffen S, Otto M, Niewoehner L, Barth M, Brozek-Mucha Z, Blegstraaten J, et al. Chemometric classification of gunshot residues based on energy dispersive X-ray microanalysis and inductively coupled plasma analysis with mass-spectrometric detection. Spectrochim Acta Part B At Spectrosc 2007;62(9):1028–36.

74. Zeichner A, Ehrlich S, Shoshani E, Halicz L. Application of lead isotope analysis in shooting incident investigations. Forensic Sci Int 2006;158(1):52–64.

75. Sarkis JES, Neto ON, Viebig S, Durrant SF. Measurements of gunshot residues by sector field inductively coupled plasma mass spectrometry—further studies with pistols. Forensic Sci Int 2007;172(1):63–6.

76. Wolten GM, Nesbitt RS, Calloway AR, Loper GL, Jones PF. Particle analysis for the detection of gunshot residue. I. Scanning electron microscopy energy dispersive X-ray characterization of hand deposits from firing. J Forensic Sci 1979;24(2):409–22.

77. Lebiedzik J, Johnson DL. Handguns and ammunitions indicators extracted from the GSR analysis. J Forensic Sci 2002;47(3):483–93.

78. Brozek-Mucha Z, Jankowicz A. Evaluation of the possibility of differentiation between various types of ammunition by means of GSR examination with SEM-EDX method. Forensic Sci Int 2001;123(1):39–47.

79. Miyauchi H, Kumihashi M, Shibayama T. The contribution of trace elements from smokeless powder to post firing residues. J Forensic Sci 1998;43(1):90–6.

80. Collins P, Coumbaros J, Horsley G, Lynch B, Kirkbride KP, Skinner W. Glass-containing gunshot residue particles: a new type of highly characteristic particle? J Forensic Sci 2003;48(3):538–53.

81. White RS, Owens AD. Automation of gunshot residue detection and analysis by scanning electron microscopy/energy dispersive- X-ray analysis (SEM EDX). J Forensic Sci 1987;32(6):1595–603.

82. Germani MS. Evaluation of instrumental parameters for automated scanning electron-microscopy/gunshot residue particle analysis. J Forensic Sci 1991;36(2):331–42.

83. Andrasko J, Oskarsson J, Stahling S. Ammunition used in the latest shooting. Forensic Sci Int 2003;136(Suppl. 1):146.

84. Kirkbride KP, Klass G, Pigou PE. Application of solid-phase microextraction to the recovery of organic explosives. J Forensic Sci 1998;43(1):76–81.

85. Bratin K, Kissinger PT, Briner RC, Bruntlett CS. Determination of nitro aromatic, nitramine, and nitrate ester explosive compounds in explosive mixtures and gunshot residue by liquid chromatography and reductive electrochemical detection. Anal Chim Acta 1981;130(2):295–311.

86. Lloyd JBF. Liquid chromatography of firearm propellants traces. J Eng Mater 1986;4:239–71.

87. King RM. The work of the explosives and gunshot residue unit of the forensic science service (UK). In: Yinon J, editor. Advances in the analysis and detection of explosives. Dordrecht: Kluwer Academic Publishers, 1993;91–100.

88. Xu X, van de Craats AM, de Bruyn PCAM. Highly sensitive screening method for nitroaromatic, nitramine and nitric ester explosives by high performance liquid chromatography-atomic pressure ionization-mass spectrometry (HPLC-API-MS) in forensic applications. J Forensic Sci 2004;49(6):1–10.

89. Mathis JA, McCord BR. Gradient reverse-phase liquid chromatographic-electrospray ionization mass spectrometric method for the comparison of smokeless powders. J Chromatogr A 2003;988:107–16.

90. Cascio O, Trettene M, Bortolotti F, Milana G, Tagliaro F. Analysis of organic components of smokeless gunpowders: high-performance liquid chromatography vs. micellar electrokinetic capillary chromatography. Electrophoresis 2004;25(10–11):1543–7.

91. Reardon MR, MacCrehan WA, Rowe WF. Comparing the additive composition of smokeless gunpowder and its handgun-fired residues. J Forensic Sci 2000;45(6):1232–8.

92. MacCrehan WA, Reardon MR, Duewer DL. Associating gunpowder and residues from commercial ammunition using compositional analysis. J Forensic Sci 2002;47(2):260–6.

93. Northrop DM. Gunshot residue analysis by micellar electrokinetic capillary electrophoresis: assessment for application to casework. Part I. J Forensic Sci 2001;46(3):549–59.

94. Northrop DM. Gunshot residue analysis by micellar electrokinetic capillary electrophoresis: assessment for application to casework. Part II. J Forensic Sci 2001;46(3):560–72.

95. Hopper KG, McCord BR. A comparison of smokeless powders and mixtures by capillary electrophoresis. J Chromatogr A 2005;50(2):19.

96. Romolo FS, Trettene M, Bortolotti F, Schutz F, Tagliaro F. Rapid, selective and quantitative determination of nitrite and nitrate ions with capillary electrophoresis: a new screening tool for gunshot residue detection. Forensic Sci Int 2003;136:147.

97. Tong Y, Wei ZP, Yang CD, Yu JY, Zhang XR, Yang SJ, et al. Determination of diphenylamine stabilizer and its nitrated derivatives in smokeless gunpowder using a tandem MS method. Analyst 2001;126(4):480–4.

98. Meng H, Caddy B. Fluorescence detection of ethyl centralite in gunshot residue. J Forensic Sci 1994;39(5):1215–26.

99. Coumbaros J, Kirkbride KP, Klass G, Skinner W. Characterization of 0.22 caliber rimfire gunshot residues by time-of-flight secondary ion mass spectrometry (TOF-SIMS): a preliminary study. Forensic Sci Int 2001;119(1):72–81.

100. Mahoney CM, Gillen G, Fahey AJ. Characterization of gunpowder samples using time-of-flight secondary ion mass spectrometry (TOF-SIMS). Forensic Sci Int 2006;158(1):39–51.

101. Flynn J, Stoilovic M, Lennard C, Prior I, Kobus H. Evaluation of X-ray microfluorescence spectrometry for the elemental analysis of firearm discharge residues. Forensic Sci Int 1998;97:21–36.

102. Brazeau J, Wong RK. Analysis of gunshot residues on human tissues and clothing by X-ray microfluorescence. J Forensic Sci 1997;42(3):424–8.

103. Berendes A, Neimke D, Schumacher R, Barth M. A versatile technique for the investigation of gunshot residue patterns on fabrics and other surfaces: m-XRF. J Forensic Sci 2006;51(5):1085–90.

104. Niewohner L, Wenz HW. Applications of focused ion beam systems in gunshot residue investigation. J Forensic Sci 1999;44(1):105–9.

105. Pun K, Gallusser A. Macroscopic observation of the morphological characteristics of the ammunition gunpowder. Forensic Sci Int 2008;175:179–85.

106. Wolten GM, Nesbitt RS, Calloway AR, Loper GL. Particle analysis for the detection of gunshot residue. II: occupational and environmental particles. J Forensic Sci 1979;24(2):423–30.

107. Wallace JS, McQuillan J. Discharge residues from cartridge-operated industrial tools. Sci Justice 1984;24(5):495–508.

108. American Society for Testing and Materials. ASTM standard E 1588-95: standard guide for gunshot residue analysis by scanning electron microscopy/energy dispersive spectroscopy. Annual book of ASTM standards. Vol. 1402. West Conshohocken, PA: American Society for Testing and Materials, 1995.

109. American Society for Testing and Materials. ASTM 1588-95: standard guide for gunshot residue analysis by scanning electron microscopy/energy-dispersive spectroscopy. West Conshohocken, PA: American Society for Testing and Materials, 2001.

110. Zeichner A, Levin N. More on the uniqueness of gunshot residue (GSR) particles. J Forensic Sci 1997;42(6):1027–8.

111. Garofano L, Capra M, Ferrari F, Bizzaro GP, Di Tullio D, Dell'Olio M, et al. Gunshot residue: further studies on particles of environmental and occupational origin. Forensic Sci Int 1999;103(1):1–21.

112. Lloyd JBF. Diphenylamine traces in hand swabs and clothing debris: cleanup and liquid chromatography with sequential oxidative and reductive electrochemical detection. Anal Chem 1987;59(10):1401–4.

113. Hawley GG. The condensed chemical dictionary, 10th edn. New York, NY: Van Nostrand-Reinhold, 1981.

114. Leggett LS, Lott PF. Gunshot residue analysis via organic stabilizers and nitrocellulose. Microchem J 1989;39(1):76–85.

115. Harris A. Analysis of primer residue from CCL Blazer (R) lead-free ammunition by scanning electron-microscopy energy dispersive X-ray. J Forensic Sci 1995;40(1):27–30.

116. Oommen Z, Pierce SM. Lead-free primer residues: a qualitative characterization of Winchester WinClean (TM), Remington/UMC Lead-Less (TM), Federal BallistiClean (TM), and Speer Lawman CleanFire (TM) handgun ammunition. J Forensic Sci 2006;51(3):509–19.

117. Brozek-Mucha Z. Examinations of various features of GSR collected from target in the dependence on the shooting distance. Forensic Sci Int 2003;136(Suppl. 1):156.

118. Zeichner A, Levin N, Springer E. Gunshot residue particles formed by using different types of ammunition in the same firearm. J Forensic Sci 1991;36(4):1020–6.

119. MacCrehan WA, Patierno ER, Duewer DL, Reardon MR. Investigating the effect of changing ammunition on the composition of organic additives in gunshot residue (OGSR). J Forensic Sci 2001;46(1):57–62.

120. Carreras LF, Palma LAM. Ejection patterns of shot residues made from 9mm Parabellum gun, 9mm short gun, .38 revolver and 7.62mm Cetme rifle. Forensic Sci Int 1998;96(2–3):143–72.

121. Schlesinger HL, Lukens HR, Guinn VP, Hackleman RP, Korts R. Special report on gunshot residue measures by neutron activation analysis. San Diego, CA: Gulf General Atomic Inc, US atomic energy commission, 1990.

122. Fojtasek L, Vacinova J, Kolar P, Kotrly M. Distribution of GSR particles in the surroundings of shooting pistol. Forensic Sci Int 2003;132(2):99–105.

123. Wolten GM, Nesbitt RS, Calloway AR. Particle analysis for the detection of gunshot residues. III: case record. J Forensic Sci 1979;24(4):864–9.

124. Gerard RV, McVicar MJ, Lindsay E, Randall ED, Smaglinski C, Harvey EA. Long-range deposition of gunshot residue and the mechanism of its transportation. Scanning 2006;28(2):106–7.

125. Bergman P, Enzel P, Springer E. The detection of gunshot residue (GSR) particles on the bottom of discharged bullets. J Forensic Sci 1988;33(4):960–8.

126. Schwoeble AJ, Harrison LG. A study of gunshot residue particle air suspension to deposition time following discharge of a weapon. Scanning 2006;28(2):107.

127. Harrison HC, Gilroy R. Firearms discharge residues. J Forensic Sci 1959;4(2):184–99.

128. Andrasko J, Maehly AC. Detection of gunshot residues on hands by scanning electron microscopy. J Forensic Sci 1977;22(2):279–87.

129. Heard BJ. Handbook of firearms and forensic ballistics. Chichester: John Wiley and Sons, 1997.

130. Jalanti T, Henchoz P, Gallusser A, Bonfanti MS. The persistence of gunshot residue on shooters' hands. Sci Justice 1999;39(1):48–52.

131. Murdock J. The collection of gunshot discharge residues. AFTE Journal 1984;16(3):136–41.

132. Krishnan SS. Detection of gunshot residue on the hands by trace element analysis. J Forensic Sci 1977;22(2):304–24.

133. Nesbitt RS, Wessel JE, Wolten GM, Jones PF. Evaluation of a photoluminescence technique for the detection of gunshot residue. J Forensic Sci 1977;22(2):288–303.

134. Douse JMF, Smith RN. Trace analysis of explosives and firearm discharge residues in the metropolitan police forensic science laboratory. J Eng Mater 1986;4:169–86.

135. Jane I, Brooks PG, Douse JMF, O'Callaghan KA, editors. Detection of gunshot residues via analysis of their organic constituents. International symposium on the analysis and detection of explosives. Washington, DC: U.S. Government Printing Office, 1983.

136. Vinokurov A, Zeichner A, Glattstein B, Koffman A, Levin N, Rosengarten A. Machine washing or brushing of clothing and its influence an shooting distance estimation. J Forensic Sci 2001;46(4):928–33.

137. Gialamas DM, Rhodes EF, Sugarman LA. Officers, their weapons and their hands: an empirical study of GSR on the hands of non-shooting police officers. J Forensic Sci 1995;40(6):1086–9.

138. Pettersson S. What conclusions can be drawn from the presence of gunshot residues. Forensic Sci Int 2003;136(Suppl. 1):158.

139. Berk RE, Rochowicz SA, Wong M, Kopina MA. Gunshot residue in Chicago police vehicles and facilities: an empirical study. J Forensic Sci 2007;52(4):838–41.

140. Degaetano D, Siegel JA. Survey of gunshot residue analysis in forensic science laboratories. J Forensic Sci 1990;35(5):1087–95.

141. Singer RL, Davis D, Houck MM. A survey of gunshot residue analysis methods. J Forensic Sci 1996;41(2):195–8.

142. MacCrehan WA, Smith KD, Rowe WF. Sampling protocols for the detection of smokeless powder residues using capillary electrophoresis. J Forensic Sci 1998;43(1):119–24.

143. Beijer R. Experiences with Zincon, a useful reagent for the determination of firing range with respect to lead free ammunition. J Forensic Sci 1994;39(4):981–7.

144. Walker JT. Bullet holes and chemical residues in shooting cases. J Crim Law Criminol 1940;31:497–521.

145. Steinburg M, Leist Y, Goldschmidt P, Tassa M. Spectrophotometric determination of nitrites in gunpowder residues on shooters' hands. J Forensic Sci 1984;29(2):464–70.

146. Tschirhart DL, Noguchi TT, Klatt EC. A simple histochemical technique for the identification of gunshot residue. J Forensic Sci 1991;36(2):543–7.

147. Dahl LK. A simple and sensitive histochemical method for calcium. Proc Soc Exp Biol Med 1952;80(3):474–9.

148. Bartsch MR, Kobus HJ, Wainwright KP. An update on the use of the sodium rhodizonate test for the detection of lead originating from firearm discharges. J Forensic Sci 1996;41(6):1046–51.

Additional information and reprint requests:
Jason W. Birkett, Ph.D.
Nottingham Trent University
School of Science and Technology
Clifton Lane, Nottingham NG11 8NS
United Kingdom
E-mail: jason.birkett@ntu.ac.uk

PAPER
CRIMINALISTICS

Jakub Szumera,[1] *M.Sc.; Mirołsaw Wełniak,*[1] *Ph.D.; Andrzej Olejniczak,*[1] *M.Sc.; and*
Jerzy P. Lukaszewicz,[1] *Ph.D.*

Transfer of Triazine-iron(II) Chromic Complexes Left by Iron Items on Textile Background and Human Skin

ABSTRACT: The research is focused on the detection and transfer of iron traces left by iron items on clothing and human skin. The method is based on the formation of colored complexes between ferrous ions and five synthesized, mostly new triazines. Iron traces originally were left by iron rings on slightly wetted (artificial sweat) cotton fabrics and subsequently transferred to a separate textile substrate. Prior to the use of trazines the contact spots were treated with a new inorganic reducing agent (Sn^{2+}) to reduce Fe^{3+} to Fe^{2+}. The method is sensitive to detect iron traces on wetted canvas after 10 min contact with iron items. More spectacular results were obtained for traces left on human palm even after very short contact (10 sec). The new iron-trace-transfer method eliminated the contact of triazines solutions with human skin. Transmission visible spectra of Fe(II)–triazine complexes were determined.

KEYWORDS: forensic science, triazine synthesis, ferrous ion complexes, iron traces detection, iron traces transfer

Criminal acts are often associated with the use of metal, particularly steel, tools like knifes, guns, etc. The usual way in which investigations proceed in such cases involves collecting, identifications, and storage of fingerprints left on the tools used for crime commitment. Little attention is paid to searching and identification of traces of metal objects left on clothing and skin of an offender. However, some ideas on how to solve the problem have been presented in literature and/or have been announced at conferences and meetings (1). In general, metal traces left on some solid backgrounds can be detected by chromogenic effects resulting from the chelation of metal ions (including iron ions) with properly selected ligands. Iron ions (particularly Fe^{2+}) easily form chromatic complexes with organic molecules belonging to the group of triazines (2). The reaction of other organic molecules with iron ions may lead to the formation of colorful complexes, too (3). Detection limit of the metal–ligand reaction (minimal concentration of iron ions that is necessary to cause a visible chromatic effect with the excess of ligand) may be regarded as a key factor in selection of ligands for practical application. It is particularly important regarding the expected low concentration of iron derivatives on the surface of human skin or clothing left upon the contact with an iron item.

In general, the concentration of iron traces on skin and clothing depends strongly on many factors like: chemical properties of iron of which the item is made, time of contact with an iron item, time elapsed after the contact, temperature, sweat presence on skin and clothing, sweat composition, washing of skin or clothing, etc. Known affinity of human skin (4,5) and some natural fibers (virgin and modified) (6–8) to metal ions results in the immobilization of

metal ions (ferrous and ferric ions) in form of iron-based derivatives on skin and clothing. It hinders (for some time) the removal of metal traces by washing. This effect should help to detect iron traces left on skin and clothing. However, the mentioned metal (iron) ion immobilization may hinder further reaction of the ions with organic ligands which should lead to visible chromogenic effects.

Avissar, Almog et al. (9–11) demonstrated that several triazines can be useful for the detection of firearm traces due to the formation of colored complexes with ferrous ions. According to these studies the applicability of triazines can be optimized regarding the structure of triazines, considering the time factor with particular respect to the state of human skin. The reaction between the triazines and Fe^{2+} ions is suitable to detect the ions which concentration is at level of single μM. Some triazine synthesis pathways may be associated to the application of hydrazine, which is widely known as a harmful substance (12). Therefore, in the current study hydrazine was replaced by hydrazine hydrate as in the case of other studies (2).

Up to now, almost all of the disclosed results on the triazine-based detection of iron traces were focused on *in situ* detection on human skin. Since some fears on toxicity of triazine-based reagents exist and must be taken into account, the elaboration of an efficient procedure for transferring iron traces onto independent background is of particular importance. Such a method can avoid a prolonged contact between triazine containing reagents and human skin. To our knowledge, only one study has dealt in part with the problem of iron traces transfer from skin or clothing to a separate background (13). However, the presented photographic documentation proved that the transfer was moderately spectacular. The color intensity of transferred traces left by iron objects was low. Additionally, the shape of colored and transferred traces was hardly recognizable. The primarily investigated transfer method did not eliminate direct and intensive wetting of skin with triazine

[1]Faculty of Chemistry, Nicholas Copernicus University, 87-100 Torun, Poland.

solutions, which we attempted to avoid (or reduce) in the current study. Additionally, this study aims at the increase of Fe^{2+} concentration by the use of a new, efficient inorganic reducing agent, i.e., Sn^{2+}. The mentioned transfer of iron traces to a separate background is somehow similar to the collection of fingerprints left by suspects on critical spots. Fingerprints are routinely transferred to a new background, protected, and stored in an archive.

In the current study, the most attention is paid to the eye observable colored effects beside the results of instrumental method investigations. Therefore, photographic documentation of the obtained results is of high importance.

Beside the main target of the study, i.e., the development of transfer method, some side targets are planned to achieve:

- Synthesis of a series of new triazines which have not been tested yet as a chelating agent for the *in situ* formation of colored complexes with ferrous ions.
- Elimination of pure hydrazine from the synthesis pathway of triazines and/or reduction of hydrazine (and hydrazine derivatives) consumption during synthesis.
- Optimization of the synthesis conditions (i.e., type of solvent, temperature, duration, extraction procedure) in order to achieve the highest possible product yield.
- Determination of spectral properties of Fe(II)– triazine complexes in solution.

Experimental

General Procedure for the Synthesis of Triazines

We intended to synthesize a series of new triazines differing from those investigated so far (1,2,9–11) but applying some important elements of the applied synthesis pathway. New triazines were obtained to demonstrate that not only so far investigated triazines are applicable to iron trace detection and confirm that number of useful triazines is in general high. In the past and in the current study, 2-pyridylhydrazidine was considered as the main substrate for the synthesis of a series of triazines. The latter compound was synthesized from 2-Cyanopyridine following the original method proposed by Case et al. (14) (Fig. 1) modified by the application of 80% hydrazine hydrate solution as in former studies (2).

At first, the yield of reaction was not satisfying, therefore several modifications were imposed to the original (15,16) pathway (type of solvent, temperature, duration, extraction procedure) to obtain effectively pyridine-4-hydrazidine. Finally, we proposed to run the synthesis in ethanol at 50°C for 4 h using equimolar proportions between hydrazine hydrate and 2-Cyanopyridine. The proposed method was characteristic because of high yield of c. 65%. The method of pyridine-2-hydrazidine synthesis excludes the molar excess of hydrazine hydrate, which reduces the cost of experiment and is valuable regarding environmental and health issues.

As mentioned, hydrazine hydrate was applied which is easier to handle and less irritant and harmful than pure hydrazine. Benzil, benzil derivatives, furil, and pyridyl served as the second substrate.
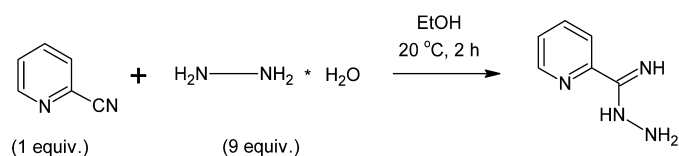
FIG. 1—*Synthesis of pyridine-2-hydrazine according to the concept of Case et al. (14).*

The general concept of the synthesis of five triazines is summarized in Table 1. In this way yellowish powders of triazines were obtained.

Visible Spectroscopy Studies

The 0.001 M solution of each triazine in ethanol (spectroscopic grade) was prepared by continuous stirring at room temperature for 48 h. For spectral studies 5 mL of each triazine solution (0.001 M) was transferred to a 10-mL volumetric flask, mixed with different amounts of Fe^{2+} ions (5, 10, 15, 20 μL of 0.010 M solution) and diluted to necessary volume with water. The final concentration of appropriate Fe(II)-triazine complexes was equal to 5, 10, 15, and 20 μM respectively. The applied proportions provided considerable molar excess of ligands (triazines I-V) ensuring the complete formation of appropriate Fe(II)-triazine complexes. The visible spectra of such obtained complexes in water-ethanol solution were recorded between 400 and 800 nm with an ultraviolet-visible spectrophotometer.

Detection of Iron Traces Left by Iron Items
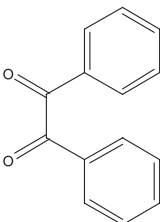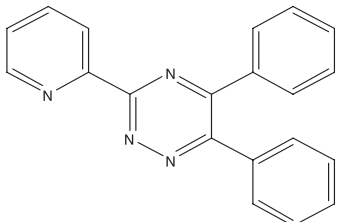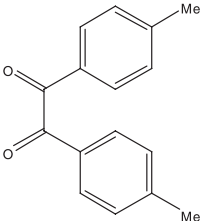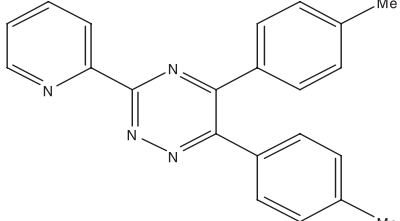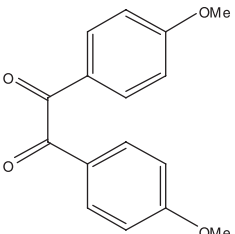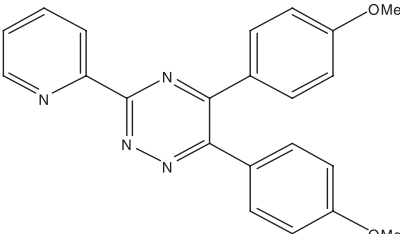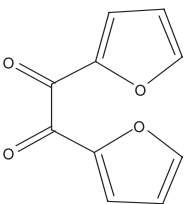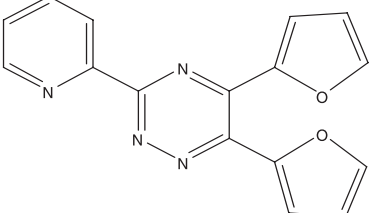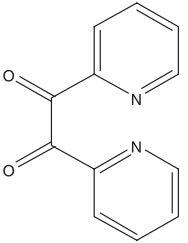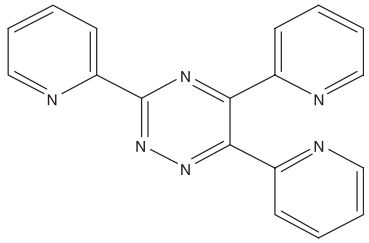
The detection of iron traces left during contact of some iron-made items with a background was tested in a series of experiments performed using the ethanol solutions of triazines (0.001 M). First, one tried to mimic in laboratory conditions in which iron item could leave traces on clothing (and skin). Therefore, an artificial sweat was used to simulate the action of the natural one on the surface of iron items. Basing on some published reports (17) one came to the conclusion that the presence of Cl^- ions in natural human sweat is a key component of human sweat which must be present in the liquid being a substitute of natural sweat. The concentration of Cl^- ions in human sweat is not constant and depends on personal inclinations, emotional state, ambient temperature, current physical activity of a human. For example it may change from 20 to 140 mM (18–22). Therefore we assumed that 100 mM concentration of Cl^- ions is an average and typical value. Additionally, we ascribed a minor role of other components of natural sweat like fatty acids in the dissolution of iron and its transfer to any background. Thus, artificial sweat used in the study was just a 100 mM solution of Cl^- in distilled water.

In the first stage of the study, we employed an adsorbing background in the form of pure cotton fabrics as a substitute of clothing (and human skin). The cotton backgrounds were saturated with the artificial sweat and carefully squeezed. The cotton backgrounds remained slightly wet during the contact with iron-made items. Iron-made rings (37 mm outer diameter and 20 mm inner diameter, 99.3% weigh content of iron) were used as test metal items which contact with background we intended to detect by the formation of colored complexes of ferrous ions with triazines. The iron rings were put on the surface of wet cotton background and left for 10, 20, 30 min. Figure 2 presents the placement of iron rings on cotton stripes wetted with the Cl^- solution. The columns from left to right corresponded to the contact time of 10, 20, and 30 min respectively. The lines in up to down order corresponded to the experiments with triazines from I to V, respectively.

After the mentioned contact time, the iron rings were removed and the contact spot was subjected to the "chromogenic" treatment including the use of dissolved trazines and appropriate reducing agent to remove Fe^{3+} ions (undesired side effect). A similar experiment was performed with iron rings put on inner part of human palm.

Beside the modifications of triazine synthesis, the basic novelty of the study consists in the transfer of iron traces left on wet textile

TABLE 1—*Synthesis of triazines.*

Second Substrate Beside Pyridine-2-Hydrazine	Solvent/Time/ Temperature	Product—Triazine
	Triazine I EtOH 24 h Room temperature	
	Triazine II EtOH 15 min Boiling temperature	
	Triazine III EtOH 15 min Boiling temperature	
	Triazine IV i-PrOH 30 min 60°C	
	Triazine V EtOH 24 h Room temperature	

background (and/or on human hand) to an independent and movable substrate and subsequent colorization of them. In this way we were going to avoid direct colorization of iron original traces on skin (and clothing). The independent substrate with transferred and colored iron traces should be much more durable than the traces on human skin or clothing which may by removed by washing. The movable and durable substrate with transferred iron-offprints may (theoretically) serve as long lasting evidence of crime commitment.

Results

The obtained triazines were characterized by the determination of melting points and recording of NMR spectra. Melting points were measured on a Boetius apparatus and were uncorrected. ^{1}H

and ^{13}C NMR spectra were recorded in CDCl$_3$ solutions with a Bruker Avance 300 spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany).

2-Pyridylhydrazidine

2-Cyanopyridine (5.2 g, 50 mmol), 80% hydrazine hydrate (3.2 mL, 50 mmol) in 10 mL of ethanol were stirred at 50°C for 4 h. The precipitated product was filtered off and a filtrate was extracted with Et$_2$O (2 × 25 mL). The etheral extract was evaporated and the combined solid products were crystallized from toluene to give 4.28 g (62.9%, m.p. 95–96°C (14), m.p. 95–96°C).

^{1}H NMR (DMSO, 300 MHz): $\delta = 8.51$ (dq, $J = 4.9$ Hz, $J = 0.9$ Hz, NCH$_{ar}$), 8.00 (dt, $J = 8.1$ Hz, $J = 1.2$ Hz, CH$_{ar}$), 7.69

FIG. 2—*Planned placement iron rings on cotton stripes (a) and real experimental setup (b).*

(ddd, $J = 8.1$ Hz, $J = 7.5$ Hz, $J = 1.8$ Hz, CH$_{ar}$), 7.26 (m, CH$_{ar}$), 5.26 (s, NH$_2$), 4.58 (s, 2 NH).
^{13}C NMR (CDCl$_3$, 50 MHz) δ: 150.7, 148.5, 147.8, 136.2, 123.6, 119.5.

5,6-Diphenyl-3-(2-pyridyl)-1,2,4-triazine (Triazine I)

2-Pyridylhydrazidine (2.72 g; 20 mmol), benzil (4.20 g, 20 mmol) in ethanol were stirred at ambient temperature for 24 h. The precipitated product was crystallized from ethanol to receive 3.47 g (55.9%, m.p. 192–194°C (14), m.p. 189–190°C).
^{1}H NMR (CDCl$_3$, 300 MHz): $\delta = 8.92$ (ddd, $J = 4.5$ Hz, $J = 1.8$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 8.72 (dt, $J = 7.8$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 7.94 (td, $J = 7.5$ Hz, $J = 1.8$ Hz, 1H, NCH$_{ar}$), 7.72 (m, 4H, CH$_{ar}$), 7.42 (m, 8H, CH$_{ar}$).
^{13}C NMR (CDCl$_3$, 50 MHz) δ: 160.8, 156.5, 156.4, 152.9, 150.5, 137.0, 135.6, 135.3, 130.7, 130.0, 129.7, 129.6, 128.6, 128.5, 125.4, 124.1.

5,6-Bis-(4-methylphenyl)-3-(2-pyridyl)-1,2,4-triazine (Triazine II)

2-Pyridylhydrazidine (1.36 g, 10 mmol), 4,4′-dimethylbenzil (2.38 g, 10 mmol) in 100 mL of ethanol was stirred under reflux for 15 min. The volume of the reaction mixture was reduced to one-third of capacity, 150 mL of hot hexane was added and left

for overnight. The yellow product was filtered off to obtain 2.26 g (66.8%, m.p. 137–139°C (16), m.p. 130–132°C).
^{1}H NMR (CDCl$_3$, 300 MHz): $\delta = 8.92$ (ddd, $J = 4.8$ Hz, $J = 1.8$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 8.69 (dt, $J = 7.8$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 7.93 (td, $J = 7.8$ Hz, $J = 1.8$ Hz, 1H, NCH$_{ar}$), 7.58 (m, 4H, CH$_{ar}$), 7.48 (ddd, $J = 7.8$ Hz, $J = 4.8$ Hz, $J = 1.2$ Hz, 1H, CH$_{ar}$), 7.18 (dd, 4H, CH$_{ar}$, $J = 9.0$ Hz, $J = 7.8$ Hz), 2.40 (s, 3H, CH$_3$), 3.38 (s, 3H, CH$_3$).
^{13}C NMR (CDCl$_3$, 500 MHz) δ: 160.5, 156.3, 156.1, 153.1, 150.4, 141.1, 139.9, 137.0, 133.0, 132.6, 129.9, 129.4, 129.3, 129.3, 125.2, 124.0, 21.5, 21.4.

5,6-Bis-(4-methoxyphenyl)-3-(2-pyridyl)-1,2,4-triazine (Triazine III)

2-Pyridylhydrazidine (1.36 g, 10 mmol), 4,4′-dimethoxybenzil (2.70 g, 10 mmol) in 100 mL of ethanol was stirred under reflux for 15 min. The mixture was condensed to about 30 mL, 150 mL of hot hexane was added and left for 24 h. The yellow solid product was filtered off and crystallized from EtOH to obtain 2.68 g (72.3%, m.p. 154–159°C (16), m.p. 143–155°C).
^{1}H NMR (CDCl$_3$, 300 MHz): $\delta = 8.88$ (ddd, $J = 4.8$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 8.62 (dt, $J = 8.1$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 7.89 (td, $J = 7.8$ Hz, $J = 1.8$ Hz, 1H, NCH$_{ar}$), 7.65 (m, 4H CH$_{ar}$), 7.62 (ddd, $J = 7.5$ Hz, $J = 4.5$ Hz, $J = 0.9$ Hz, 1H, CH$_{ar}$), 6.92 (m, 3H, CH$_3$), 3.8 (s, 3H, CH$_3$).
^{13}C NMR (CDCl$_3$, 50 MHz) δ: 161.7, 160.8, 160.1, 155.6, 155.2, 153.0, 150.3, 136.9, 131.6, 130.8, 127.9, 127.8, 125.1, 123.8, 114.0, 113.9, 55.3, 55.2.

5,6-Bis-(2-furyl)-3-(2-pyridyl)-1,2,4-triazine (Triazine IV)

2-Pyridylhydrazidine (3.40 g, 25 mmol), α-furil (4.75 g, 25 mmol) in 40 mL of I-propanol were stirred at 60°C for 0.5 h. The formed precipitate was filtered off, washed with water and then transferred to 40 mL of 0.06 M aq. HCl. Mixture was stirred at ambient temperature for 2.5 h. The yellow product was filtered off, washed with water and hexane to obtain 5.06 g (69.7%, m.p. 159–162°C (23), m.p. 155–157°C).
^{1}H NMR (CDCl$_3$, 300 MHz): $\delta = 8.92$ (ddd, $J = 4.8$ Hz, $J = 1.8$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 8.67 (dd, $J = 8.1$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 7.94 (dt, $J = 7.8$ Hz, $J = 1.5$ Hz, 1H, NCH$_{ar}$) and 7.92 dt, $J = 7.8$, $J = 1.5$ Hz, 1H, NCH$_{ar}$), 7.7 (dd, $J = 1.5$ Hz, $J = 0.6$ Hz, 1H, CH$_{furil}$) and 7.65 (dd, $J = 1.5$ Hz, $J = 0.6$ Hz, 1H, CH$_{furil}$), 7.45 (ddd, $J = 7.5$ Hz, $J = 4.8$ Hz, $J = 0.1$ Hz, 1H, NCH$_{ar}$), 7.45 (ddd, $J = 7.5$ Hz, $J = 4.8$ Hz, $J = 0.1$ Hz, NCH$_{ar}$), 7.22 (d, $J = 3.6$ Hz, CH$_{furil}$), 7.20 (d, 1H CH$_{furil}$, $J = 3.6$ Hz), 7.00 (d, $J = 3.6$ Hz, CH$_{furil}$), 6.65 (dd, $J = 3.6$ Hz, $J = 1.8$ Hz, 1H, CH$_{furil}$) and 6.60 (dd, $J = 3.6$ Hz, $J = 1.8$ Hz, 1H, CH$_{furil}$).
^{13}C NMR (CDCl$_3$, 500 MHz) δ: 160.3, 152.5, 150.6, 148.4, 148.3, 146.9, 145.1, 144.6, 144.6, 137.0, 125.4, 124.2, 118.1, 113.8, 112.7, 112.4.

3,5,6-Tri-(2-pyridyl)-1,2,4-triazine (Triazine V)

2-Pyridylhydrazidine (2.72 g, 20 mmol), 2,2′-pyridyl (4.24 g, 20 mmol) in 20 mL of ethanol were stirred at ambient temperature for 24 h. The light-yellow product was filtered off and crystallized from 2-ethoxyethanol. The yellow-brown product was filtered off to obtain 2.95 g (47.2%, m.p. 191–193°C (14), m.p. 186–187°C).
^{1}H NMR (CDCl$_3$, 300 MHz): $\delta = 8.90$ (ddd, $J = 4.8$ Hz, $J = 1.8$ Hz, $J = 0.9$ Hz, 1H, NCH$_{ar}$), 8.70 (dt, $J = 7.8$ Hz,

J = 1.2 Hz, 1H, NCH$_{ar}$), 8.91 (m, 2H, NCH$_{ar}$), 8.23 (dt. J = 7.5 Hz, J = 0.9 Hz, 2H, NCH$_{ar}$), 7.92 (td, J = 7.8 Hz, J = 1.8 Hz, 1H, NCH$_{ar}$), 7.87 (m, 3H, NCH$_{ar}$), 7.46 (ddd, J = 7.8 Hz, J = 4.8 Hz, J = 1.2 Hz, 1H, NCH$_{ar}$), 7.26 (m, 2H, CH$_{ar}$).

^{13}C NMR (CDCl$_3$, 500 MHz) δ: 161.1, 156.1, 156.0, 154.6, 154.4, 152.4, 150.5, 148.6, 137.0, 136.9, 136.8, 125.5, 124.7, 124.3, 124.3, 124.1, 123.6.

The five synthesized compounds are respectively denoted as triazine I, II, III, IV, and V. As mentioned, the triazines were supposed to form colored complexes with iron(II) ions. The contact of iron items with water and/or water solution of chloride ions (textile backgrounds saturated with artificial sweat) triggers iron corrosion which consists in the release of iron ions from ring surface and subsequent transfer of the ions to the textile (wet) background. In the first step the formation of Fe^{2+} ions occurs (and ferrous derivatives), which later undergo oxidation to Fe^{3+} ions (and ferric derivatives) under the influence of atmospheric oxygen. The following set of reactions is frequently quoted:

$$2Fe^0 \rightarrow 2Fe^{2+} + 4e \, (\text{iron oxidation})$$

$$\underline{O_2 + 2H_2O + 4e \rightarrow 4OH^- \, (\text{oxygen reduction})}$$

$$\text{In summary} : 2Fe^0 + O_2 + 2H_2O \rightarrow 2Fe(OH)_2$$
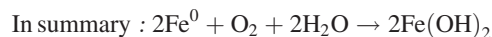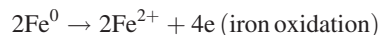
Iron(II) hydroxide is water insoluble, greenish in color, and hardly observable in room conditions since it usually undergoes oxidation by atmospheric oxygen:

$$2Fe(OH)_2 + \frac{1}{2}O_2 + H_2O \rightarrow 2Fe(OH)_3$$
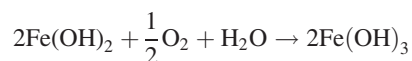
Thus, a mixture of Fe(II) and Fe(III) derivatives is expected to be present in the wet textile background. Fe(OH)$_3$ as brownish traces was visible on the textile backgrounds (Fig. 3) in fact. The intensity of brownish coloring increased upon the prolonged time of contact of iron ring with the wetted textile background. The formation of Fe(III) containing derivatives is an undesired effect since Fe^{2+} ions are capable of forming colored complexes with triazines. Therefore one applied a new inorganic reducing agent, i.e., Sn^{2+} ions in water solution prior to the colorization of traces by means of dissolved triazines. The advantages of Sn^{2+} ions are twofold. First and foremost, the difference between the standard oxidation potentials of Sn^{4+}/Sn^{2+} (equal to c. +0.15 V) and Fe^{3+}/Fe^{2+} (equal to c. +0.77 V) is high enough to provide effective reduction of Fe^{3+} ions to Fe^{2+} ions. Second, the molar excess of Sn^{2+} ions may disturb the complexation equilibrium between Fe^{2+} and skin proteins, and, finally improves the transfer of iron(II) ions form the skin. However, in the last time Sn(II) compounds are supposed to irritate skin especially in the case of prolonged contact and if skin is broken (24). Fortunately, the human body can quickly rid itself
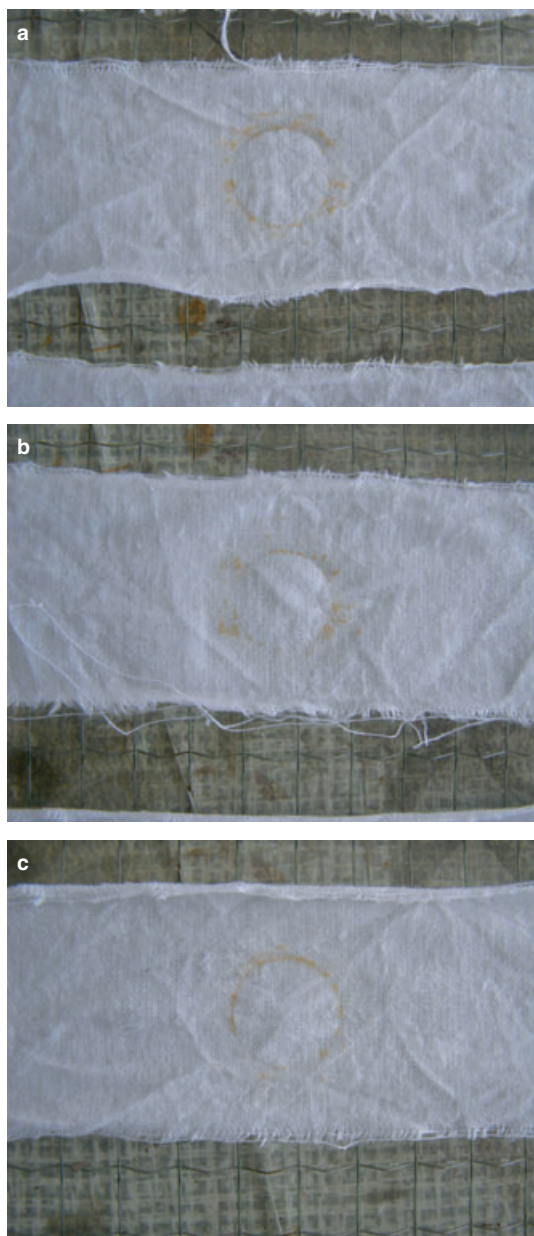
FIG. 3—A series (a, b, c) of example brownish traces left by iron rings on wetted ([Cl−] = 0.1 M) cotton backgrounds. Contact time 20 min, room temperature.

FIG. 4—A complete set of independent tampons with colored traces (left by iron rings) transferred from cotton backgrounds to the tampons. The subsequent rows are corresponding to triazines I to V, respectively. The original contact spots (iron ring-wetted cotton background) on cotton fabrics treated with reducing solution of Sn2+ prior to the transfer. Room temperature. The single upper sample corresponds to a blind test (no contact with iron ring, treated with same chemical reagents as other spots).

of most inorganic tin in weeks, at least for 2–3 months (very small amounts of tin stay for longer time in some tissues like the bones).

Thus, a diluted water solution of $SnCl_2$ (0.001 M) was used to saturate the textile background after the removal of iron rings. 2–3 min later we put an independent substrate (a textile tampon saturated with ethanol solution of the investigated triazines) on each contact spot. In this way we tried to avoid direct spraying of a triazine solution onto contact spots. The tampons were pressed firmly for 10 sec to the wet textile backgrounds supposed to contain iron (mainly ferrous derivatives) traces left by iron rings. Then, the tampons were turned upside-down and put beside the original contact spots (Fig. 4). It is visible that the original contact spots were not intensively colored since ferrous ions were transferred to tampons and then colored by triazines (I-V). To confirm that

FIG. 5—*A complete set of independent tampons with colored traces (left by iron rings) transferred from cotton backgrounds to the rampons. The original contact spots (iron ring-wetted cotton background) on cotton fabrics treated with reducing solution of Sn2+ and additionally saturated with solutions of traizines. Room temperature. The single upper sample corresponds to a blind test (no contact with iron ring, treated with same chemical reagents as other spots). Minor increase of color intensity after wetting of the original contact spots with triazine solutions.*
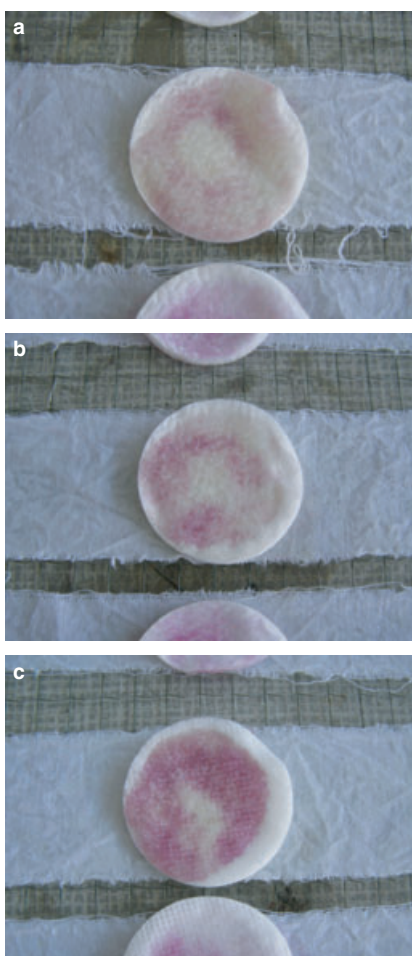
FIG. 6—*A set of textile tampons with colored traces of iron rings transferred from cotton backgrounds. Triazine II, room temperature, contact time: 10 (a), 20 (b), 30 (c) min.*

FIG. 7—*Results of an experiment involving triazine II and human palm. Mechanical print on skin just after the removal of iron ring (a). Colored iron traces transferred to independent textile substrate (b). Residual iron traces reduced and colored by triazine II solution on skin (c).*

suggestion, we saturated the original contact spots on cotton stripes first with the reducing solution and then with triazines in ethanol solution (the triazines are not water soluble). Then, dry tampons were put on the contact spots. Also in this case the color intensity of chelated iron traces on tampon was higher than the intensity of chromic effect related to residual iron traces at the original contact spots (Fig. 5). It confirms the occurrence of three phenomena:

- the applied experimental conditions (including reaction/contact time) are adequate to start the oxidative dissolution of iron (corrosion) and transfer iron ions to the wetted cotton background,
- transfer of iron traces to independent tampons is efficient and only minor amounts of iron ions remain on the cotton backgrounds,
- the detection limit of the proposed method (iron traces detection via the formation of triazine complexes) is low and chromic effects are visible without any instrumentation support.

The intensity of tampon coloring increases with the increasing contact time of iron ring with the wet textile background (Fig. 6). The colored spots on tampons are resembling the size and shape of iron rings being in contact with the textile background (Figs. 4–6).

The Fe(II)-triazine complexes transferred to the tampons differ from each other regarding their color. One may notice that ring-looking offprints on tampons are red-pink for trazines I, II, III and dark blue for triazines IV, V. It is consistent with visible transmission spectra recorded for all Fe(II)-triazine complexes. The noticed transfer of iron traces from wet cotton background (possibly from human skin or clothes) to new independent and movable substrate (tampon) is a unique effect which has not been exploited intensively. Despite the first published transfer attempt, researchers' attention has been for long paid only to on-skin coloring of iron traces left by metal items. In an early study Lee (13) demonstrated that such a transfer from skin to a piece of filter paper (wetted with distilled water) was possible. However, that study differs at some points from the current one. All necessary regents including triazine solutions were directly applied on skin, which resulted in colorization of skin areas that were in contact with iron object. Moreover, the enclosed photographic documentation proved that the transfer of Fe(II)-triazine complexes was less spectacular that the effects observed in the current study.

The cotton tampons preserved the original color attributed to the formation of Fe(II)-triazine complexes after drying. No significant change of intensity of colored contact spots was observed even after 3 months. The possibility of the transfer of iron item traces left on suspect's skin or clothing to an independent and movable background is potentially of great importance regarding the collection of crime evidences. Such a background with transferred traces
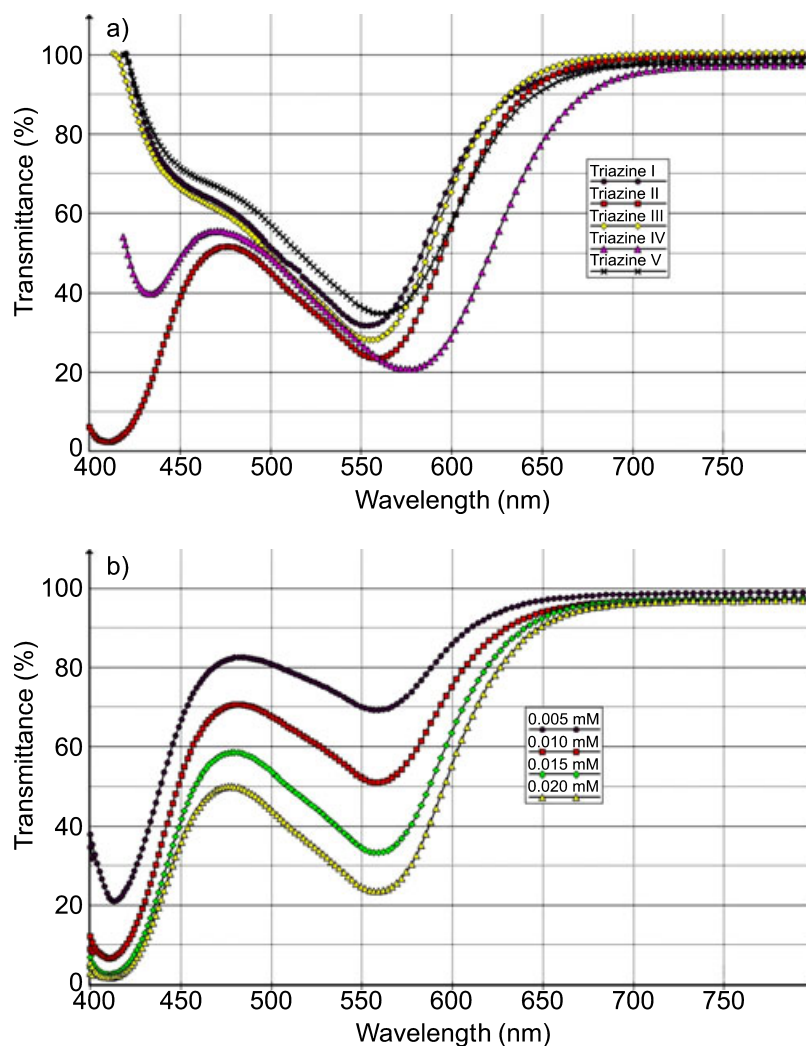
FIG. 8—*Transmission visible spectra recorded for Fe(II)-triazine complexes in water-ethanol solution: 0.020 M solution of triazine (I, II, III, IV, V)-based complexes (a) and concentration dependent changes in UV-VIS spectra of Fe(II)-triazine II complex (b).*

may in general serve as evidence stored in archive instead of (or parallel to) a photographic documentation of colored iron traces left on skin of a suspect.

The so far reported results dealt with the iron traces primarily left on canvas background saturated with artificial sweat. The textile background was made of pure cotton. A question may arise if a similar transfer is possible in the case of human skin and iron traces left on it. Some concerns may be attributed to the different chemical nature of the textile background used in the study and the chemical composition of human skin. Generally, polysaccharides (cellulose) are the main component of the textile while proteins of different kind are the most important constituent of human skin.

To confirm the hypothesis, we performed a similar experiment with a human palm. In the first step of the experiment, we pressed an iron ring against the surface of a human palm for considerably shorter time than mentioned in parts of the study (Fig. 7a, room temperature of 21–22°C, 45% relative humidity, 10 sec). In these conditions, the palm of one of the authors, according to his subjective impression, was not particularly intensively sweated. After 10 sec only, the ring was removed and the contact spot was wetted with the mentioned water solution of Sn^{2+} i.e., a cotton tampon saturated with the reducing solution was put on the contact spot and hold for the next 10 sec. Then, we put the next cotton tampon saturated with ethanol solution of triazine II. As formerly, the tampon was turned upside down and put on the palm beside the original trace left by iron ring on the skin (Fig. 7b). Also in this case we observed the formation of intensively colored trace on the tampon. It means that Fe^{2+} ions present on the palm (left by iron ring) were transferred from skin to the textile tampon. Additionally, we treated the contact spot on the palm with the trazine II solution and observed only a weak colorization of the original trace left on the palm (Fig. 7c). Thus, the extraction of iron ions from human skin is intensive enough to form color complexes with triazines.

Figure 8 presents some transmission visible spectra recorded for the Fe(II)- triazine complexes. From the run of spectra one can conclude that all the investigated complexes are characteristic because of broad absorption peak in the range c. 550–580 nm (Fig. 8a). The complexes with traizines II and IV exhibit additional absorption peaks at the edge of the examined range (c. 410–440 nm). For the tested complexes the intensity of absorption peak in the first mentioned region depends strongly on the complex concentration as it is demonstrated in Fig. 8b for Fe(II)–triazine II complex.

Summary

A series of triazines was synthesized in the way involving limited amounts of hydrazine hydrate. The structure of the triazines was confirmed by NMR studies proving that a series of mostly not tested triazines was successfully obtained. The trazines hardly dissolve in ethanol but after dissolution may be applied to the detection of iron traces left by iron items on absorbing textile background saturated with artificial sweat. The application of described procedure leads to the transfer of iron traces to an independent and movable substrate. The image of the original metal item (iron ring) is visible due to the formation of colored Fe(II)-triazine complexes on the movable substrate. Similar effects were recorded in parallel experiments performed with a human palm. Traces left by an iron ring on a palm can be successfully transferred to a separate substrate. Spectral data show that the reaction between the synthesized triazines and ferrous ions was very sensitive to the presence of the ions. The experiments confirm that ferrous/ferric ions bonding to the components of human skin and cotton is not particularly durable. The ions may be extracted and transferred to independent backgrounds and stored for a long time.

References

1. Almog J, Avissar YY, Sagiv A, Mandler D. Detection of recent handling of firearms. Proceedings of the American Academy of Forensic Sciences; 2005 Feb 21-26; New Orleans, LA. Colorado Springs, CO: American Academy of Forensic Sciences, 2005;115.
2. Almog J, Hirshfeld A, Glattstein B, Sterling J, Goren Z. Chromogenic reagents for iron(II): studies in the 1,2,4-triazine series. Anal Chim Acta 1996;322:203–8.
3. Cheng GW, Crisosto CH. Iron-polyphenol complex formation and skin discoloration in peaches and nectarines. J Am Soc Hortic Sci 1997;122:95–9.
4. Hostynek J, Maibach H. Skin penetration by metal compounds with special reference to copper. Toxicol Mech Methods 2006;16:245–65.
5. Hostynek J, Hinz RS, Guy RH. Metals and the skin: topical effects and systemic absorption. New York, NY: Marcel Dekker, 1999.
6. Lee BG, Rowell RM. Removal of heavy metal ions from aqueous solutions using lignocellulosic fibers. J Nat Fibers 2004;1:97–108.
7. Okieimen FE, Ogbeifun DE, Nwala GN, Kumsah CA. Binding of cadmium, copper, and lead ions by modified cellulosic materials. Bull Environ Contam Toxicol 1985;34:866–70.
8. Lee HJ, Lee BG, Shin DY, Park H. Effect of different chemical and physical characteristic having lignocellulosic fibers on heavy metal ion removal from aqueous solution. Mater Sci Forum 2008;569:285–8.
9. Leifer A, Avissar Y, Berger S, Wax H, Donchin Y, Almog J. Detection of firearm imprints on the hands of suspects: effectiveness of PDT reaction. J Forensic Sci 2001;46:1442–6.
10. Avissar YY, Sagiv AE, Mandler D, Almog J. Identification of firearms holders by the [Fe(PDT)3]+2 complex. Quantitative determination of iron transfer to the hand and its dependence on palmar moisture levels. J Forensic Sci 2004;49:1215–9.
11. Avissar YY, Sagiv AE, Mandler D, Almog J. Identification of firearms handling by the [Fe(PDT)3]2+ complex: chemical and time-dependent factors. Talanta 2005;67:328–33.
12. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry. Hydrazine CAS # 302-01-2U.S., Sept (1997). Washington, DC: US Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry, 1997.
13. Lee C-W. The detection of iron traces on hands by ferrozine sprays: a report on the sensitivity and interference of the method and recommended procedure in forensic science investigations. J Forensic Sci 1986;31:920–30.
14. Case FH. The preparation of hydrazidines and as-triazines related to substituted 2-cyanopyridines. J Org Chem 1965;30:931–3.
15. Shintou T, Ikeuchi F, Kuwabara H, Umihara K, Itoh S. Synthesis of 2-pyridylpyridines via aza-Diels–Alder reactions between 3-pyridyl-1,2,4-triazines and some vinyl alkanoates. Chem Lett 2005;34:836–7.
16. Stephen WI, Islam MA. 3,5,6-Trisubstituted 1,2,4-triazines as analytical reagents: Part I. Compounds containing the ferroin functional group or iron(II)-methine chromophore. Anal Chim Acta 1993;247:335–46.
17. Burton JL, Pye RJ, Brookes DB. Metal corrosion by chloride in sweat. The problem of 'rusters' in industry. Br J Dermatol 1976;95:417–22.
18. Costa F, Galloway DH, Margen S. Regional and total body sweat composition of men fed controlled diets. Am J Clin Nutr 1969;22:52–8.
19. Przepiorka M. Mineral concentrations in human sweat during exercise. Med Sportiva 2001;5:E97–105.
20. Verde T, Shephard RJ, Corey P, Moore R. Sweat composition in exercise and in heat. J Appl Physiol 1982;53:1540–5.
21. Morgan RM, Patterson MJ, Nimmo MA. Acute effects of dehydration on sweat composition in men during prolonged exercise in the heat. Acta Physiol 2004;182:37–43.
22. van Heyningen R, Weiner JS. The effect of arterial occlusion on sweat composition. J Physiol 1952;116:404–13.
23. Hennessy DJ, Reid GR, Smith FE, Thompson SL. Ferene—a new spectrophotometric reagent for iron. Can J Chem 1984;62:721–4.
24. US Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry. Toxicological profile for tin and tin compounds, Sept (2005). Washington, DC: US Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry, 2005.

Additional information and reprint requests:
Jerzy P. Lukaszewicz, Ph.D.
Associate Professor
Research Group for Modeling and Synthesis of Novel Materials
Faculty of Chemistry
Nicolaus Copernicus University
Gagarina 7, 87-100 Toruń
Poland
E-mail: lukaszju@chem.uni.torun.pl

PAPER

CRIMINALISTICS

L. Scott Chumbley,[1] *Ph.D.; Max D. Morris,*[1] *Ph.D.; M. James Kreiser,*[2] *B.S.; Charles Fisher,*[1] *B.S.; Jeremy Craft,*[1] *M.S.; Lawrence J. Genalo,*[1] *Ph.D.; Stephen Davis,*[1] *B.S.; David Faden,*[1] *B.S.; and Julie Kidd,*[1] *M.S.*

Validation of Tool Mark Comparisons Obtained Using a Quantitative, Comparative, Statistical Algorithm

ABSTRACT: A statistical analysis and computational algorithm for comparing pairs of tool marks via profilometry data is described. Empirical validation of the method is established through experiments based on tool marks made at selected fixed angles from 50 sequentially manufactured screwdriver tips. Results obtained from three different comparison scenarios are presented and are in agreement with experiential knowledge possessed by practicing examiners. Further comparisons between scores produced by the algorithm and visual assessments of the same tool mark pairs by professional tool mark examiners in a blind study in general show good agreement between the algorithm and human experts. In specific instances where the algorithm had difficulty in assessing a particular comparison pair, results obtained during the collaborative study with professional examiners suggest ways in which algorithm performance may be improved. It is concluded that the addition of contextual information when inputting data into the algorithm should result in better performance.

KEYWORDS: forensic science, tool mark comparison, comparison microscope, screwdriver, statistics, striae

In the fifteen years since the 1993 *Daubert versus State of Florida* decision, increasing attacks have been aimed at firearm and tool mark examiners by defense attorneys via motions to exclude evidence based on expert testimony. Such motions claim that the study of tool marks has no scientific basis, that error rates are unknown and incalculable, and that comparisons are subjective and prejudicial. Often persuasive, these motions skillfully blend truth with unsupported assertions or assumptions in a number of ways. First, the claim that scientific evidence is lacking in tool mark examinations ignores the numerous studies that have been conducted, especially in the area of firearms (1–4), to investigate the reproducibility and durability of markings. These studies have shown time and again that while matching of cartridges cannot be universally applied to all makes and models of guns using all types of ammunition, the characteristic markings produced are often quite durable and a high percentage can be successfully identified using optical microscopy. Second, the claims that error rates are unknown, and that the probability of different guns having identical markings has not been established, are true. However, it must be understood that establishing error rates and probabilities in the area of tool marks is fundamentally different than in an area such as genetic matching involving DNA. When considering genetic matching, all the variables and parameters of a DNA strand are known and error rates can be calculated with a high degree of accuracy. This is not the case in tool marks where the variables of force,

angle of attack, motion of the tool, surface finish of the tool, past history of use, etc. are not known or cannot be determined, and the possibility for variation is always increasing as the population under study continues to increase and change. For practical purposes, this may indeed mean that realistic error rates cannot be completely characterized, but experiments based on sequentially manufactured tools may lead to useful approximations and/or bounds.

Finally, it is also true that an examiner necessarily offers a subjective opinion when rendering a decision. However, the pattern on which that decision is based consists of striations that can be characterized and quantified in an objective, mathematical manner. The proposition that tool marks must necessarily have a quantifiable basis is the principle upon which the Integrated Ballistics Imaging System (IBIS) developed and manufactured by Forensic Technology, Inc. for bullets and cartridge cases operates. IBIS uses fixed lighting and an image capture system to obtain a standard digital image file of the bullet or cartridge case. The contrast displayed in the image is reduced to a digital signal that can then be used for rapid comparisons to other files in a search mode. The latest version of IBIS uses the actual surface roughness as measured by a confocal microscope to generate a comparison file. The results are displayed in a manner analogous to a web search engine, where possibilities are listed in order with numbers associated with each possibility. An experienced tool mark examiner must then review the list of possibilities to make a judgment as to whether a match does, in fact, exist. In instances where a match is declared, it is quite common for the match not to be the first possibility displayed by IBIS, but to be further down the list. In other words, while the analysis/algorithm employed by FTI produces the numbers associated with each match, these numbers carry no clear statistical

relevance or interpretation related to the quality or probability-of-match of any given comparison (5). However, as the marks under investigation can be quantified, there appears to be a significant potential for advancement in analyses of such data. An objective method of analysis should be possible for any given type of tool mark, and (at least in principle) an error rate established for comparisons made between any given subset of marks within a larger population of similar marks.

Researchers at Iowa State University have developed a computer-based data analysis technique that allows rapid comparison of large numbers of data files of the type that might be produced when studying striated tool marks A major aim of the research reported here is to construct well-defined numerical indices, based upon the information contained within the tool mark itself, that are useful in establishing error rates for objective tool mark matching. While this error rate may only be practically achievable for a particular set of experimental conditions, it should serve as a benchmark error rate for subsequent studies. Initial results (6) indicated that simple statistics computed from the quantitative data produced by a surface profilometer, namely, maximized data correlations over short data segments, supported the empirical assertions of forensic examiners concerning comparisons of tool marks generated on lead plates by consecutively manufactured screwdriver tips. One drawback in using maximized correlations is that there is no clear standard against which they can be objectively compared. In some cases, maximized correlations may be high, implying a high degree of linear agreement between data pairs, but not necessarily implying strong similarity between the tool mark patterns. In others, the linear correlations over short data segments may be smaller, but the overall tool mark patterns are convincingly similar and would be declared a positive identification by a practicing examiner. One situation in which this shortcoming is especially troublesome is in poorly marked samples where striations may not be present across the entire surface of the lead plates used for making the tool marks. For example, consider the possibility where two dissimilar tools are used to mark two plates. Suppose that in both cases the screwdriver tip does not adequately mark the surface. In such cases the similar unmarked sections of the plates may produce very high correlation values, even though the marked sections are entirely dissimilar. For these and many other reasons, a simple maximized correlation coefficient is not a reliable index of match quality.

This article presents a description of a matching analysis and algorithm that overcomes many of these difficulties and summarizes experimental data collected to characterize algorithm performance. The index produced by the algorithm provides a more statistically meaningful comparison than maximized correlation. Experiments involving comparisons of samples obtained from a single tool to each other, and to samples produced from other similar sequentially manufactured tools, show that the analysis can fairly reliably separate sample pairs that are known matches from the same tool from pairs obtained from different tools. Additionally, the index provides a means of calculating estimates of error rates within the narrow and specific setting of this study.

For the sake of clarity, a brief summary of how the algorithm operates and the assumptions upon which it is based is given later. This discussion is necessary to understand the algorithm results in comparison with those obtained by human subjects. Agreement between algorithm results and examiner evaluations was assessed at the 2008 Association of Firearms and Tool Mark Examiners Training Meeting held in Honolulu, Hawaii. Results obtained from this blind study in which practicing tool mark examiners were asked to compare the same samples will be presented. Comparison of the results obtained by human examiners to those of the algorithm

provides interesting insights that will lead to algorithm performance improvements.

Statistics

An earlier work (7) described a statistical analysis and algorithm for comparing two-dimensional images of tool marks. The algorithm described here is similar in construction, although it is restricted only to matching along one-dimensional profilometer data traces, and so is lacking some of the steps required to deal with two-dimensional data arrays. The data examined in this analysis are of the type collected by a surface profilometer that records surface height (z) as a function of distance (x) along a linear trace taken perpendicular to the striations present in a typical tool mark. Some important assumptions in the analysis are that the values of z are reported at equal increments of distance along the trace and that the traces are taken as nearly perpendicular to the striations as possible. The algorithm then allows comparison of two such linear traces.

The first step taken by the algorithm, referred to as Optimization, is to identify a region of best agreement in each of the two datasets for the specified size of the comparison window (which is user-defined). This is determined by the maximum correlation statistic, hereafter referenced as an "R-value," and described in (6). By way of illustration, two different possibilities are shown in Fig. 1. The schematic of Fig. 1a shows the comparison of a true match, i.e., profilometer recordings from two specimens made with the same tool, while Fig. 1b shows data from a true nonmatch pair of specimens (i.e., two marks from two different tools). In each case, the matched regions marked with solid rectangles are the comparison windows denoting the trace segments over which the ordinary linear correlation coefficient is largest. Note that in both cases the R-value returned is very close to 1, the largest numerical value a correlation coefficient can take. In the first instance this is so because a match does in fact exist, and the algorithm has succeeded in finding trace segments that were made by a common section of the tool surface. In the second case, the large R-value is primarily a result of the very large number of correlations calculated in finding the best match. Even for true nonmatches, there will be short trace segments that will be very similar, and it is almost inevitable that the algorithm will find at least one pair of such segments when computing the R-value. It is primarily for this reason that the R-values cannot be interpreted in the same way that simple correlations are generally evaluated in most statistical settings.

For the reasons described earlier, the algorithm now conducts a second step in the comparison process called Validation. In this step a series of corresponding windows of equal size are selected at randomly chosen, but common distances from the previously identified regions of best fit. For example, a randomly determined shift of 326 pixels to the left, corresponding to the dashed rectangles in Fig. 1a, might be selected. The correlation for this pair of corresponding regions is now determined. Note that this correlation must be lower than the R-value, because the latter has already been determined as being the largest of all possible correlations determined in the Optimization step. The assumption behind the Validation step is that if a match truly does exist, correlations between these shifted window pairs will also be reasonably large because they will correspond to common sections of the tool surface. In other words, if a match exists at one point along the scan length (high R-value), there should be fairly large correlations between corresponding pairs of windows along their entire length. However, if a high R-value is found between the comparison windows of two
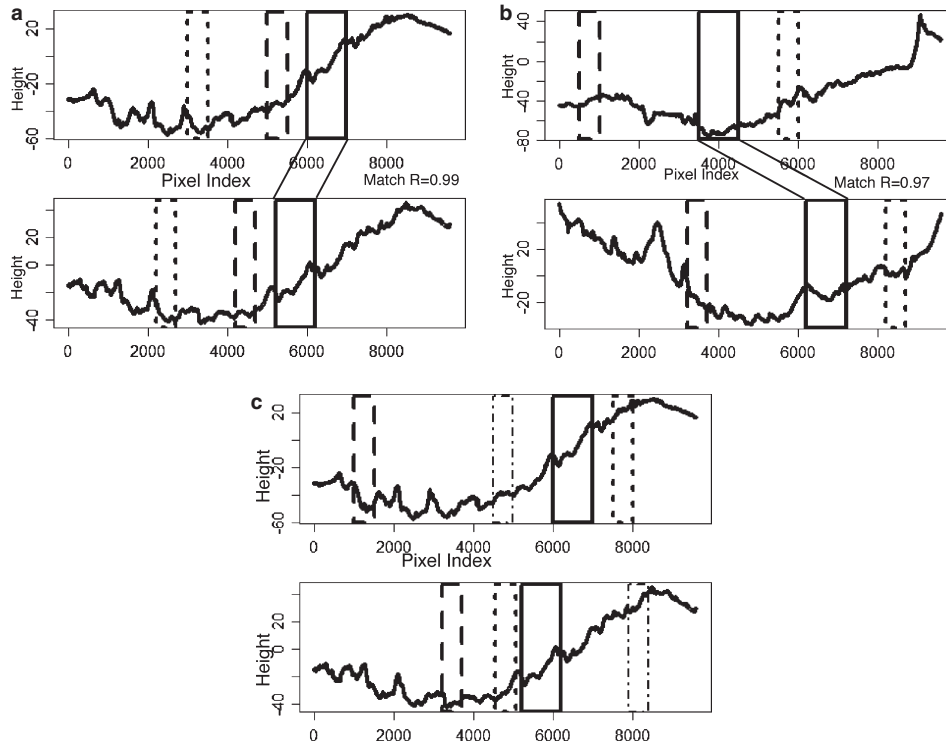
FIG. 1—(a) Comparison pair showing a true match. Best region of fit shown in solid rectangle with corresponding R-value. Note the similarity of the regions within the two possible sets of validation windows (dashed and dotted rectangles). (b) Comparison pair showing a true nonmatch. While a high R-value is still found between "Match" segments, the validation windows are distinctly different from one another. (c) Validation windows (dashed, dotted, and dot-and-dash rectangles) selected at random for the comparison pair shown in (a) to establish a baseline value.

nonmatch samples simply by accident, there is no reason to believe that the accidental match will hold up at other points along the scan length. In this case rigid-shift pairs of windows will likely not result in especially large correlation values.

During the Validation step a fixed number of such segment pairs are identified, corresponding to a number of different randomly drawn shifts, and the correlation coefficient for each pair is computed. Dotted and dashed rectangles displayed in Fig. 1 illustrate schematically the selection of two such pairs of shifted data segments; in actual operation the algorithm chooses many such pairs. In the case of the true match the regions within the corresponding dashed windows of Fig. 1a do appear somewhat similar and can be expected to return fairly large correlation values. However, when similar corresponding pairs of windows are taken from the nonmatch comparison of Fig. 1b, the shape of the scans within the windows is seen to differ drastically. Lower correlation values will be obtained in this case.

The correlation values computed from these segment pairs can be judged to be "large" or "small" only if a baseline can be established for each of the sample comparisons. This is achieved by identifying a second set of paired windows (i.e., data segments), again randomly selected along the length of each trace, but in this case, without the constraint that they represent equal rigid-shifts from their respective regions of best fit. In other words, for this second set of comparisons the shifts are selected at random and independently from each other—any segment of the selected length from one specimen has an equal probability of being compared to any segment from the other. This is illustrated in Fig. 1c for three pairs of windows, denoted by the dashed rectangles, the dotted rectangles, and the dot-and-dash rectangles.

The Validation step concludes with a comparison of the two sets of correlation values just described, one set from windows of

common random rigid-shifts from their respective regions of best agreement, and one set from the independently selected windows. If the assumption of similarity between corresponding points for a match is true, the correlation values of the first set of windows should tend to be larger than those in the second. In other words, the rigid-shift window pairs should result in higher correlation values than the independently selected, totally random pairs. In the case of a nonmatch, as the identification of a region of best agreement is simply a random event and there truly is no similarity between corresponding points along the trace, the correlations in the two comparison sets should be very similar.

A nonparametric Mann–Whitney U-statistic (referred to in this article as T1), computed from the joint ranks of all correlations computed from both samples, is generated for the comparison. Where the correlation values of the two comparison sets are similar, T1 takes values near zero, supporting a null hypothesis of "no match." If the correlations from the first rigid-shift sample are systematically larger than the independently selected shifts, the resulting values of T1 are larger, supporting an alternative hypothesis of "match."

Method

The test set for this study is the same as described in (6), namely, a series of 50 sequentially manufactured screwdriver tips were obtained and used to make tool marks at angles of 30°, 60°, and 85° on flat lead plates. The surface roughness of the resultant striae was measured using a surface profilometer and the measurements saved as a series of data files detailing z height as a function of x direction. All details of data collection are given in (6).

To compare the effectiveness of the algorithm to human examiners, and potentially identify areas where the algorithm might be
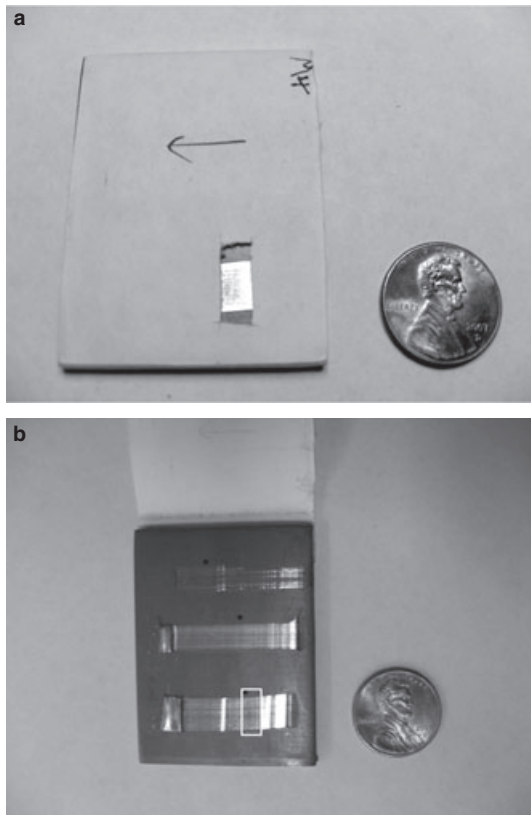
FIG. 2—*Image of a tool marked plate with (a) a blinder in place; (b) the blinder removed, a white rectangle shows the visible portion when the blinder is in place.*

enhanced or improved, a double-blind study was conducted during the 2008 Association of Forearms and Tool Mark Examiners Training Seminar. During the course of this meeting, 50 different volunteers rendered over 250 opinions on some of the sample pairs used for this study and evaluated by the algorithm.

A series of 20 comparison pairs covering a range of T1 values from low to high were selected from the tool marks produced at the 85° comparison angle. Of the 20 comparison pairs, five were from samples where the algorithm correctly identified a matched set (high T1); five were correctly eliminated nonmatch comparisons (low T1); five were incorrectly eliminated matched sets (T1 values in the low or inconclusive range); and five were incorrectly identified nonmatches (intermediate or high T1). Examiners were asked to assess each pair of samples twice. For the initial observation, paper blinders were placed on the samples so that examiners were restricted in their view to the same general area where the profilometer data were collected (Fig. 2). After making an initial assessment, the blinders were removed and the examiner was given the opportunity to make a second assessment based on a view of the entire sample. In each case, examiners were asked to render an opinion as to whether they were viewing a positive identification, a positive elimination, or inconclusive, for reasons that will become apparent.

Names of examiners were not recorded, although demographic data were collected concerning the experience and training of the volunteers. Of the 50 volunteers, all except five were court qualified firearm and tool mark examiners. Of the remaining five, two were firearms (but not tool mark) qualified, two were in training, and one was a foreign national where a court qualification rating does not exist. Volunteers were required to do a minimum of two

comparison pairs and could do as many as they wished. Several chose to do the maximum number of comparisons possible. Numbers were assigned to identify each volunteer during data collection; afterward the ID numbers were randomly mixed to preserve anonymity.

Examiners were asked to use whatever methodology they employed in their respective laboratories. This caused some confusion initially and placed constraints on the volunteers because some laboratories never use the term "positive elimination," while others are reluctant to use the term "positive identification" unless the examiner personally either makes the marks or knows more information about them than what could be supplied in this study. After understanding this, the examiners were told the direction of the tool when making the mark and that the tool marks were all made at the same angle from similar, sequentially made, flat blade screwdriver tips. Also, examiners were told that for the study they could consider the terms of "positive elimination" or "inconclusive" to be essentially interchangeable.

Results and Discussion

Algorithm Performance

The data obtained from the profilometer were used to test a series of hypotheses that are held as being true by tool mark examiners (Fig. 3). The first and most fundamental assumption, that all tool marks are unique, was tested by a comparison of marks made by different screwdriver tips at the angles of 30°, 60°, and 85° with respect to horizontal. The T1 statistic values are shown in Fig. 4 as a function of angular comparison. The data are plotted as box plots, the boxes indicating where 50% of the data falls with the spread of the outlying 25% at each end of the distribution shown as dashed lines. As stated previously, when using a T1 statistic a value relatively close to 0 indicates that there is essentially no evidence in the data to support a relationship between markings. For pairs of samples made with different screwdrivers (Fig. 4) the majority of the index T1 values produced by the algorithm fall near the 0 value; only three outlier comparisons had a T1 value greater than ±4.

In comparison, Fig. 5 displays indices computed using the algorithm from profilometer scans of marks made by the same side of the same tool and compared as a function of angle. While marks made at different angles still produce index values near 0, the T1 statistic jumps dramatically when marks made at similar angles are considered. Clear separation is seen between the 50% boxes, although overlap still exists when the outliers are considered.

Taken together, Figs. 4 and 5 support Hypotheses 1 and 2. When comparing tool marks made at similar angles with different tools, the resulting T1 values cluster near zero (Fig. 4), but when the same tool is used to make marks at similar angles, the T1 distributions are on substantially larger values, giving support for Hypothesis 1. Support for Hypothesis 2 is demonstrated by Fig. 5 alone, because even among the same tool marks, only those made at the same angle produce large T1 values.

Hypothesis 1:	*The 50 sequentially produced screwdrivers examined in this study all produce uniquely identifiable tool marks*
Hypothesis 2:	*In order to be identifiable, tool marks from an individual screwdriver must be compared at similar angles.*
Hypothesis 3:	*Different sides of a flat-bladed screwdriver produce different uniquely identifiable marks.*

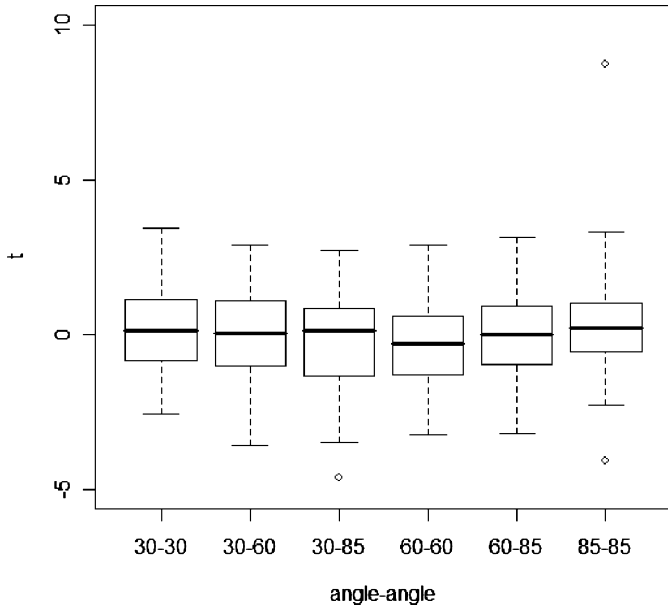FIG. 3—*Summary of hypothesis tested in this study.*

FIG. 4—*Box plots showing T1 results when comparing marks from different screwdrivers.*
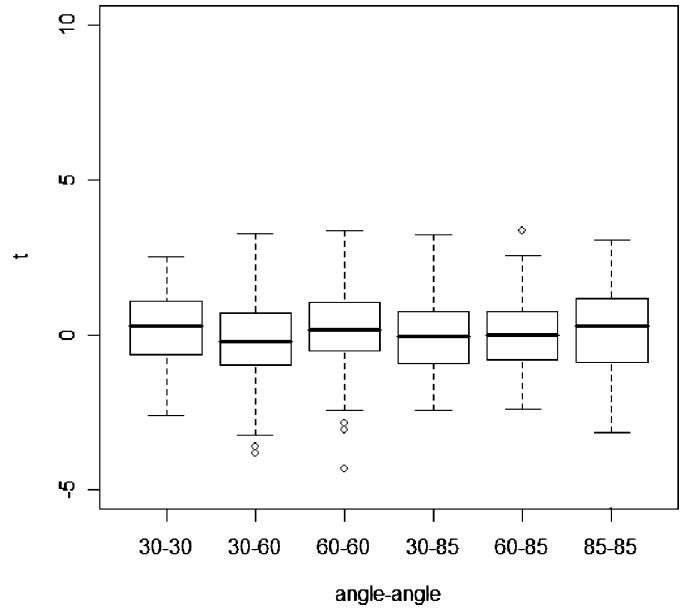

FIG. 6—*Box plots showing T1 results when comparing marks made from different sides of the same screwdrivers.*
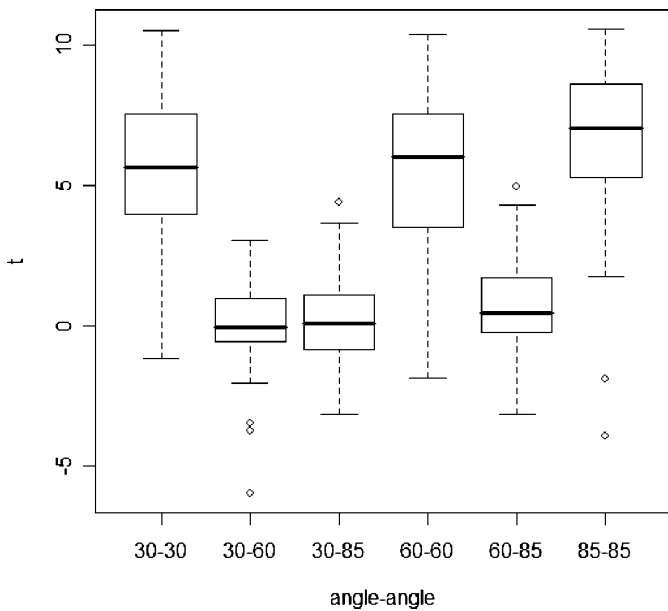

FIG. 5—*Box plots showing T1 results when comparing marks obtained from the same side of the same screwdrivers.*

The last hypothesis considered was that when comparing tool marks made from screwdriver tips, the marks must be made from the same side of the screwdriver; marks made using different sides of the screwdriver appear as if they have come from two different screwdrivers. These results are shown in Fig. 6. The hypothesis is again supported because, as in Fig. 4, the T1 values cluster around 0 regardless of the angles used in making the marks, indicating no relationship between the samples.

The T1 values summarized in Figs. 4 and 5 are individually replotted in Fig. 7, with the *y*-axis randomly varied (known as jittering) to create an artificial vertical separation that makes it easier to view the data points. Known comparisons that should match and

produce high T1 values are shown in black. Known "nonmatches" that should have T1 values near zero are shown in gray.

Examination of these plots indicates that the algorithm operates best using data obtained at higher angles than lower angles, i.e., the spread of black and gray spots is more defined for the 85° data than, for example, the 30° data. This is believed related to the quality of the mark. As the angle of attack of the screwdriver with the plate increased, the quality of the mark increased. It was common to obtain marks that represented the entire screwdriver tip at high angles, while marks at lower angles were often incomplete (5). Algorithm performance also appears more efficient at reducing false positives than it does in eliminating false negatives. At all angles known matches were found with very low T1 values, while nonmatches with high T1 values were very limited.

While T1 is a much more stable index of match quality than *R*-value, problems still remain in establishing an effective, objective standard for separating true matches from nonmatches. Ideally, when employing standard *U*-statistic theory the critical T1 values separating the regions of known matches (black data points) and known nonmatches (gray data points) should remain constant for all datasets. Examination of Fig. 7 shows that this is not the case. For example, reasonable separation for the 30° and 60° data appears to be somewhere around a T1 value <5, but rises to approximately 7 for the 85° data. This variation is most likely attributed to the lack of complete independence among the correlations computed in each sample, arising from the finite length of each profilometer trace.

For the reasons discussed earlier, assigned threshold values indicating "Positive ID" and "Positive Elimination," and denoted by black lines on the graphs of Fig. 7, were chosen based on a *K*-fold cross-validation using 95% one-sided Bayes credible intervals. Specifically, the lower threshold is a lower 95% bound on the 5th percentile of T1 values associated with nonmatching specimen pairs, and the upper threshold is an upper 95% bound on the 95th percentile of T1 values associated with matching specimen pairs. The region between these two threshold values is labeled "Inconclusive." A Markov Chain-Monte Carlo simulation was used to determine potential error rates.
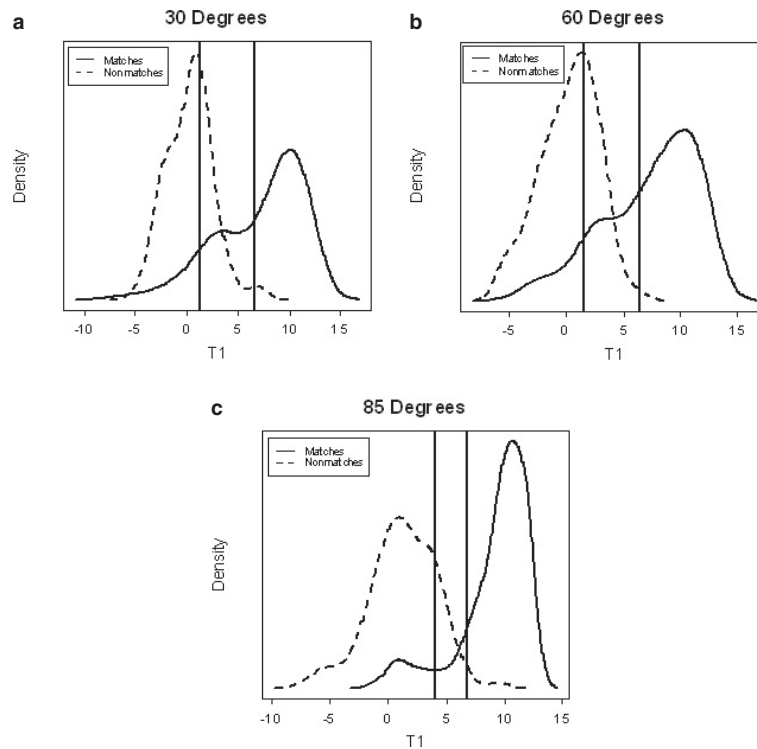
FIG. 7—*Summation of the T1 values from comparisons made at (a) 30°; (b) 60°; (c) 85°.*

Using this method the estimated error rates are as follows. For comparisons made at 30° the estimated probability of a false positive (i.e., a high T1 value for a known nonmatch comparison) is 0.023. In other words there is a possibility of slightly over two false positives for approximately every 100 comparisons. The estimated probability of a false negative is 0.089, or almost nine true matches having a low T1 value per every 100 comparisons. The cross-validation method used ensures that all the data have similar error rates, and the rates found for the 60° and 85° data are approximately 0.01 and 0.09 for false positives and false negatives, respectively. What is most noticeable is that the T1 lower threshold value for the 85° data is much larger than for the 30° and 60° data, being 4.10 versus 1.34 and 1.51, respectively. This suggests that a more distinct difference is required to classify nonmatches for the 30° and 60° cases than is true for the 85° case. This, in turn, results in a corresponding increase for the estimated inconclusive error rates, which are 0.103, 0.298, and 0.295 for the 85°, 60°, and 30° data, respectively. It would, of course, be possible to shift these error rates, i.e., produce fewer false negatives at the expense of more false positives, by altering the percentiles used in our estimation procedure.

Association of Firearm and Tool Mark Examiners Study

Results of the computerized analysis of specimen pairs were compared to expert evaluations of the same samples made by volunteer examiners at the 2008 Association of Firearm and Tool Mark Examiners seminar. However, before the algorithm performance can be discussed in comparison with the data obtained at the Association of Firearm and Tool Mark Examiners seminar using human volunteers, a brief consideration of the constraints experienced by the examiners is in order. First, it should be recognized that the conditions under which the examiners rendered an opinion would ordinarily be regarded as restrictive or even professionally unacceptable. Without having the tool in hand, or without being permitted to make the actual mark for comparison, tool mark examiners were forced to make assumptions they would not make in an actual investigation. For example, without having the screwdriver tip in hand the examiners did not know whether the mark they observed represented the entire width or only a portion of the screwdriver blade. Second, given this uncertainty about how the specimen was made, examiners tended to be more conservative in their willingness to declare a positive identification or elimination. During the course of the Association of Firearm and Tool Mark Examiners study, several examiners commented that typical lab protocol would require them to have physical access to the subject tool before rendering a "positive identification" judgment. Finally, examiners do not typically employ the terms used to denote the three regions identified for error analysis. Thus, while privately saying they felt a comparison was a "positive elimination" (given their knowledge of the test being conducted), lab protocol required an opinion of "inconclusive" to be rendered. Such policies are put in place because the signature of a tool may so change during use that a mark made at one point in time may not resemble a mark made with the same tool at a different point in time, e.g., after the tip has been broken and/or reground. In such cases positive elimination is only allowed if the class characteristics of the marks are different.

When viewed in light of these constraints, some interesting observations concerning the algorithm performance are apparent. In a small number of cases (12 out of 252 comparisons), when examining the entire tool mark after first viewing only the restricted area where the profilometer scans were obtained, examiners changed their opinion from inconclusive to either positive ID or positive elimination. This indicates that algorithm performance might be improved simply by increasing the amount of data processed. This may be achieved, for example, by ensuring that the profilometer scans span the entire width of the mark or possibly by considering

a number of scans taken at locations dispersed along the entire length of the available mark.

In a slightly smaller number of cases, comparisons between specimens made by the same screwdriver that were not conclusively identified as such by the algorithm also presented problems for the examiners. Five true matches that received low T1 values and were classified as a positive elimination by the algorithm were examined during the Association of Firearm and Tool Mark Examiners study. Three of the five were given ratings of "inconclusive" or "positive elimination" on one occasion, and one particular comparison sample (designated MW4) was rated this way seven times. Thus, while examiners in general were vastly superior to the algorithm in picking out the matches, both the algorithm and the examiners had more trouble with some true matches than with others.

Close examination of the sample that was most often problematic for examiners (i.e., MW4) was conducted, and the images obtained are shown in Fig. 8. Figure 8a shows the side-by-side comparison of the marks, where no match is seen. Note that the mark width matches extremely well, and the entire mark seems to be present. Figure 8b shows the samples positioned where the true match is evident. It can be seen that each mark only represents a portion of the screwdriver blade width, predominantly from the two sides of the tip. A match is only possible if the marks are offset, allowing the opposing "edge" sections (which actually were

FIG. 8—*Sample MW4 with (a) tool marks placed so that assumed edges align; (b) correct placement required for positive identification.*

produced by the middle of the screwdriver blade) to be viewed side-by-side.

This sample points out weaknesses in the study conducted at the Association of Firearm and Tool Mark Examiners as well as in the laboratory tests of the algorithm. In a screwdriver mark comparison it is common for examiners to use the edges of the marks as initial registration points for the start of an examination. As examiners make the comparison marks themselves they are well aware of the edge markings, if not for the evidence marks, at least for the marks they produced. In the Association of Firearm and Tool Mark Examiners study, such information was not provided and may have led to some false assumptions. For example, in the majority of cases the volunteers were under some pressure to quickly conduct a comparison before, e.g., the next meeting session started, or so that another examiner could use the equipment, etc. Because of these time constraints, samples often were placed on the stages of the comparison microscope for the volunteer, giving the examiner little or no time to observe the macroscopic appearance of the mark. Without the benefit of seeing the size of the entire mark, and given the identical widths of the two partial marks for sample MW4 when initially viewed using the comparison microscope, the assumption that the entire width of the screwdriver blade was represented would be a natural one. However, such an assumption could easily lead to an inconclusive or positive elimination conclusion, especially if the examiner was being conservative because of the lack of information concerning the sample.

The problem described earlier essentially relates to the examiners having a lack of a point of reference or registry of the mark for the comparison. The same could be said of the algorithm and the manner in which it performs, because no point of registry exists to indicate when the data being acquired is actually coming from a tool-marked region or from the unmarked plate. All of the profilometer scans analyzed by the algorithm were run using the same set of sampling parameters. However, the initial positioning of the stylus was inexact. For incomplete marks, large regions of the unaffected lead plate were also scanned to keep the file sizes consistent, and this lack of registry could have affected algorithm performance. This is not immediately evident if one examines the raw profilometer traces (Fig. 9). In this figure the top and bottom traces show the entire scans while the two middle traces show the matched details found within the two corresponding solid rectangles superimposed on the top and bottom traces. At first sight the two scans do appear quite different, as the offset in the scans, revealed during examination at the Association of Firearm and Tool Mark Examiners, is not immediately evident in the data files. Given observation of Fig. 8, one can mark the approximate location of the region that is common between the two traces; this is shown in Fig. 9 by the shaded rectangles. In this case, paired validation windows displaced equal amounts in either direction may return a low T1 value because the majority of either scan is not held in common with the other. In other words, there is a high probability that the validation windows fall in regions where no correspondence between plates exists (see Fig. 8b). Thus, what should be a match is rated as a nonmatch.

A somewhat different problem is revealed when traces from true nonmatch samples are examined (Fig. 10). In these instances, the optimization step may identify windows at extreme edges of the two traces as being most similar. Given the nearness of the match to the ends of the traces, the random selection of paired, rigid-shift windows during the validation step is severely constrained. For the example shown in Fig. 10a, the match region (denoted by solid rectangles) falls at the extreme right ends of the data files. This means that the rigid translations taken for each pair of verification
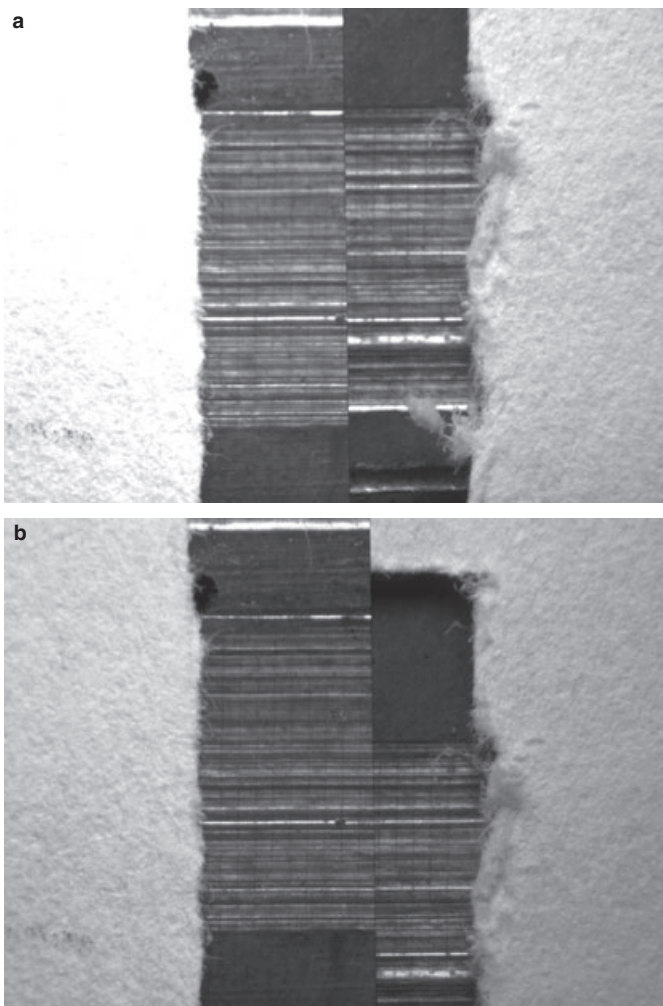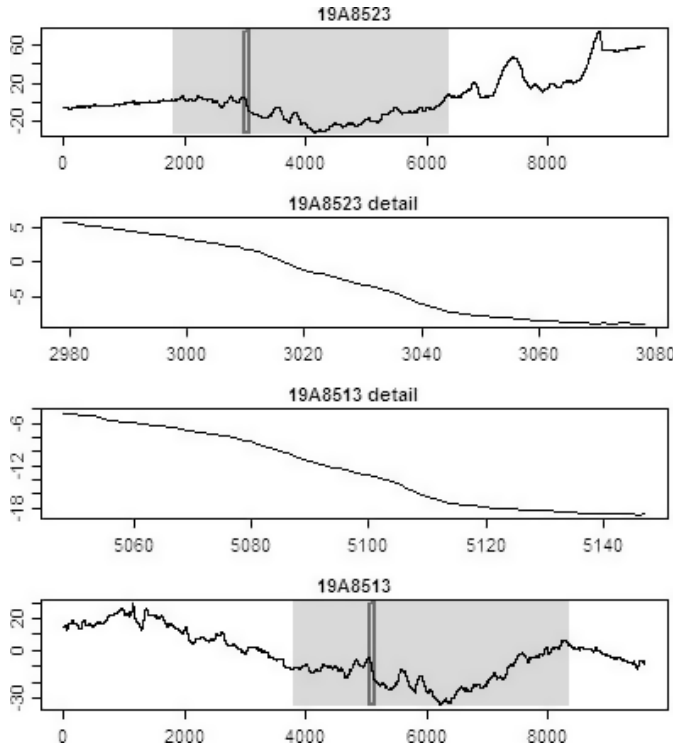
FIG. 9—*Profilometer data showing results from comparison MW4. The match region is shown by the solid rectangle.*

windows must always fall to the left of the match region. While this may be less than desirable, the entire mark is still available for validation and a large number of rigid-shift windows spaced across the entire length of the file should be sufficient to produce good separation between this accidental match and the T1 values of a true match. However, this is not true for the true nonmatch shown in Fig. 10*b*. In this case the windows identified in the optimization step as being most similar are at opposite ends of the compared data traces. The distances of possible rigid translations are constrained to a short distance to the left of the top profile and a short distance to the right of the bottom profile. Thus, the majority of the mark cannot be used in the validation step for this accidental match. If the regions in the immediate vicinity of the accidental match are also similar, high T1 values may be returned because of the constrained sampling parameters, giving results that cannot be separated from a true match.

The earlier discussion suggests that further development of the algorithm to incorporate additional data concerning the region of the profilometer trace that is actually tool-marked and/or the location of the tool edge might improve its performance. While tool mark examiners do not directly use features such as these as a basis for identification, they do use it indirectly in establishing a context for the comparison. Such information, routinely and quickly noted by a human examiner, is unavailable to the current algorithm. The algorithm treats all possible pairs of trace windows the same way and functions under the assumption that all marks analyzed are essentially the same, i.e., it assumes the screwdriver tip has completely marked the lead plate and that no unmarked regions exist. This clearly is not the case. At this time it appears the best way to enhance algorithm performance is to ensure that all comparison windows (i.e., Match and Validation) are taken from regions representing the true marked surface of the lead, and that most similar windows found at the trace edges are used as a basis for match
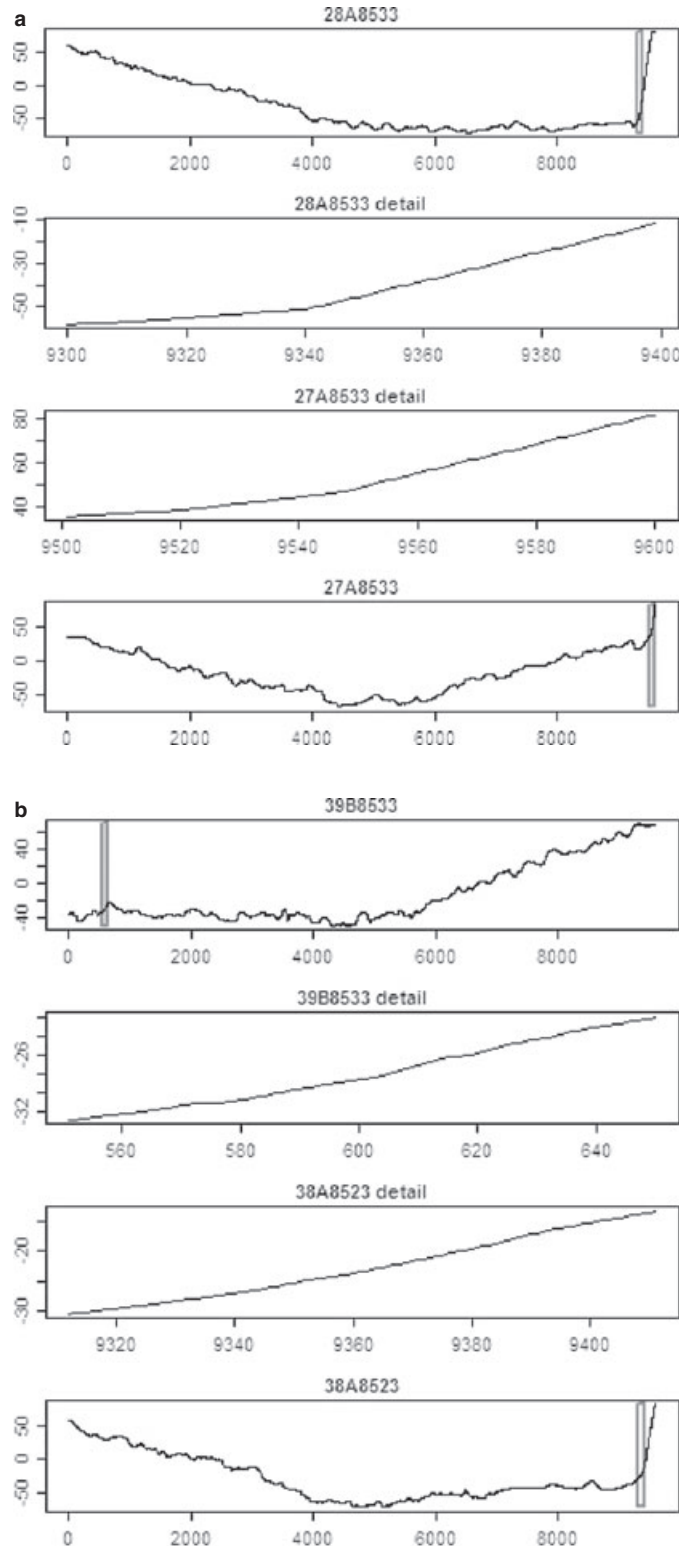
FIG. 10—*Comparison of traces obtained from four different screwdrivers that were rated as possible matches by the algorithm. (a) Good agreement found at similar ends of the traces resulted in high T1 numbers for known nonmatching pairs. (b) Good agreement found at opposite ends of the traces resulted in high T1 numbers for known nonmatches.*

identification only if they are found at the same end of their respective traces.

As a final comment, it should be noted that all types of volunteers (practicing examiners, trainees, retired examiners) were

involved in the study, with records kept as to the experience of the participant. Examination of the demographic data in relation to the results showed no significant difference between experienced examiners and rather newly qualified examiners or those in training; all performed equally well.

Summary and Conclusions

The analysis described here for comparing two tool-marked plates is a substantial improvement over simply identifying regions of highest correlation. It does this by producing a nonparametric Mann–Whitney statistic, here called T1, obtained through an optimization step followed by a validation step as a measure of evidence for tool mark matching. When used in evaluating the three hypotheses tested, namely, the uniqueness of tool marks, the necessity of comparing marks at similar angles, and the uniqueness of different sides of screwdriver blades, the T1 statistic results constitute support for the experiential knowledge of tool mark examiners. Analysis of algorithm performance in light of actual examiner results reveals deficiencies in algorithm performance that can now be addressed. Increasing the data input, possibly by including more scans spread over a larger tool mark area and incorporating contextual information normally available to examiners, such as the presence of partial scans and reference points from tool edges, may lead to performance improvements. Such changes should, for example, prohibit the identification of opposite-end windows in the optimization step as potential match segments.

It is clear that examiner performance was much better than the algorithm. While the 20 samples examined at the Association of Firearm and Tool Mark Examiners represent only a subset of the total comparisons examined using the algorithm, they did contain those samples that were most definitively misclassified by the algorithm. For example, of the 20 true match pairs shown to the Association of Firearm and Tool Mark Examiners volunteers, the algorithm correctly identified 10 of the 20 samples unambiguously; the remaining 10 were listed either as inconclusive or misidentified as a false negative. In comparison, only 11 out of 126 volunteer examinations resulted in false negative classification of true match pairs, primarily from sample MW4 (7 out of 11). Further, the Association of Firearm and Tool Mark Examiners volunteers reported no false positives at all. (N.B. The caveat must be added that the terminology used in the previous statements regarding errors is not entirely consistent with examiner protocols and should not be construed by the reader to suggest that the examiners erred. Examiners are trained to render an opinion of positive identification only when no doubt exists in their minds. Thus, a false negative only suggests that the examiner was not fully persuaded.)

References

1. Bisotti A. A statistical study of the individual characteristics of fired bullets. J Forensic Sci 1959;4(1):34–50.
2. Bonfanti MS, DeKinder J. The influence of the use of firearms on their characteristic marks. AFTE Journal 1999;31(3):318–23.
3. Bonafanti MS, De Kinder J. The influence of manufacturing processes on the identification of bullets and cartridge cases—a review of the literature. Sci Justice 1999;39:3–10.
4. Bisotti A, Murdock J. Criteria for identification or state of the art of firearm and tool mark identification. AFTE Journal 1984;16(4):16–24.
5. Committee to Assess the Feasibility, Accuracy and Technical Capability of a National Ballistics Database, National Research Council for the National Academy of Sciences. Ballistic imaging. Cork DL, Rolph JE, Meieran ES, Petrie CV, editors. Washington, DC: National Academies Press, 2008.
6. Faden D, Kidd J, Craft J, Chumbley LS, Morris M, Genalo L, et al. Statistical confirmation of empirical observations concerning tool mark striae. AFTE Journal 2007;39(3):205–14.
7. Baldwin D, Morris M, Bajic S, Zhou Z, Kreiser MJ. Statistical tools for forensic analysis of tool marks. Ames (IA): Ames Laboratory Technical Report, 2004; IS-5160, http://www.osti.gov/bridge/product.biblio.jsp?osti_id=825030.

Additional information and reprint requests:
L. Scott Chumbley, Ph.D.
Materials Science and Engineering Department
Iowa State University
2220 Hoover
Ames, IA 50011
E-mail: chumbley@iastate.edu

PAPER

CRIMINALISTICS

*Giorgia De Paoli,[1] Ph.D; Samuel A. Lewis Sr.,[2] M.S.; Ellyn L. Schuette,[1] M.S.;
Linda A. Lewis,[1] Ph.D.; Raynella M. Connatser,[2] Ph.D.; and Tivadar Farkas,[3] Ph.D.*

Photo- and Thermal-Degradation Studies of Select Eccrine Fingerprint Constituents

ABSTRACT: Photo- and thermal-degradation studies on eccrine fingerprint components are presented herein. Dilute distinct solutions of urea, lactic acid, and seven amino acids were deposited on steel coupons and Teflon® disks, exposed to artificial sunlight or heat, extracted, and analyzed. This aim of this study was to determine whether the investigated eccrine components, previously determined to be Raman active for a parallel study, experienced photo- or thermally induced degradation, and if so, to determine the rate and identify any detectable products. Neither the amino acids nor urea exhibited photo-degradation; however, when heated for a period of three minutes, the onset of thermal-degradation was initiated at 100°C for the amino acids and 100°C for urea. Lactic acid, the major polymerization initiator of superglue fuming, showed photochemical and thermal-degradation. These results could be used for future development of new latent fingerprint visualization methods, especially when lactic acid is degraded.

KEYWORDS: forensic science, latent fingerprint, fingerprint chemistry, artificial eccrine fingerprint, photo-degradation, thermal-degradation, liquid chromatography-electrospray mass spectrometry, atmospheric pressure chemical ionization-mass spectrometry

Fingerprint evidence offers great value to criminal investigations; it is no surprise, then, that fingerprint development remains an active area of research within the forensic science community. Over the years, great diversity in latent fingerprint development techniques has unfolded for use with the various print types and surface matrices that are encountered. Yet, despite the many advances, situations still remain in which print development is inconsistent or completely unsuccessful. One such situation involves the exposure of deposited prints to high temperatures, such as those created by an IED detonation, fired ammunition, or evidence from an arson site. There is evidence for thermal decomposition of fingerprint constituents (1,2), but a thorough characterization of the decomposition processes and products is lacking. Another problematic situation involves the aging of prints in the presence of light. Recent work by Lewis identified sodium lactate as the major initiator of the cyanoacrylate polymerization that occurs during the popular superglue fuming technique, and then went on to describe the degradation (and subsequent failure of print development) that resulted when the lactate in a fingerprint deposit was exposed to light (Lewis LA, Devault G. Technology development, enhanced latent fingerprint detection in missing and exploited children investigations. Final Report to Department of Energy; work for others program [NIJ], 2004. Personal communication). How many other instances of a print development method performing poorly as prints age could be the result of the thermal- or photo-degradation of the key reactant within the fingerprint deposit? This discovery underscores the need for a thorough study of possible light and/or heat degradation of other fingerprint components. Once such thermal- and photo-degradation information is available, scientists would be in a much better position to develop successful fingerprint visualization techniques by targeting the stable and degradation products. Thus, the current study was embarked upon to provide that characterization of photo- and thermal-degradation of fingerprint materials.

A deposited fingerprint can be classified as one of two general print types based on composition: an eccrine (clean) print or a sebaceous (oily) print. Eccrine prints contain only those materials secreted by the eccrine glands, which are found over the entire surface of the skin and in especially high concentrations on the palms of the hands and the soles of the feet (collectively known as the volar surfaces). More than 98% water, the balance of eccrine print composition includes urea, sodium lactate, inorganic salts, free amino acids, free organic acids, and some mucin-like glycoproteins (3–5). Sebaceous prints include both eccrine secretions and the sebum produced by the sebaceous glands, an oily secretion comprised free fatty acids, wax esters, squalene, cholesterol esters, and cholesterol (3,4). The volar surfaces do not contain sebaceous glands, which are located by hair follicles. However, sebaceous prints are prevalent because of the contamination of the fingertips by touching other areas of the body. Regardless of print type, the residue comprising a fingerprint deposit is a complex mixture of compounds. One study alone identified over 300 unique components in a sebaceous print (6). To foster a more manageable study, the scope of the degradation experiments was limited to eccrine fingerprint, which contains the major initiators responsible for superglue polymerization. Whereas a great deal of research has been carried out on sebaceous secretions (4,7), there are substantially fewer publications concerning eccrine secretions (Mong GM, Peterson CE, Clauss TRW. Advanced fingerprint analysis project.

[1]Oak Ridge National Laboratory, Bethel Valley Road, Building 4500N, Rm. F56, MS 6120, Oak Ridge, TN 37831.
[2]Oak Ridge National Laboratory, NTRC Building MS 6159, 2360 Cherahala Blvd. Knoxville, TN 37932.
[3]Phenomenex, Inc., 411 Madrid Avenue, Torrance, CA 90501.

Fingerprint constituents. PNNL Report 13019, 1999; Personal communication). This data pool shrinks still further when focusing on quantitation. To create an artificial eccrine fingerprint solution, the authors relied on a study of amino acids in an eccrine thumbprint (8), supplemented by the eccrine sweat entry in the Geigy Scientific Tables (5). Specifically, Raman active eccrine components were targeted in support of a Raman imaging technique being developed at ORNL in parallel with this degradation study.

The decision to study eccrine materials was made not only to simplify the composition of the artificial fingerprint solution, but also because eccrine prints are better models of children's fingerprints (9). Children do not begin to secrete sebum until they are between 7 and 10 years old (3), the usual age range in which adrenal androgens are first produced. Therefore, lessons learned in the study of eccrine print degradation could be applied to the development of fingerprints of both adults and children.

Materials and Methods

Artificial Eccrine Fingerprint Standards

Hamilton's quantitative data of the individual amino acids plus urea found in an eccrine-only thumbprint (8) was combined with reference data from the Geigy Scientific Tables on eccrine sweat secretions (5) to create a 44-component artificial eccrine fingerprint solution (Table 1). After screening the individual component chemicals for ease of detection by surface enhanced raman spectroscopy (SERS) (10) for a correlating degradation studies discussed in another publication, the following SERS-active components were chosen for further study: aspartic acid, glutamic acid, glycine, histidine, ornithine, serine, threonine, urea, and lactic acid. Aqueous fingerprint standards were made of each amino acid at five times its single-print concentration, a mixture of all seven amino acids at five times their single-print concentrations, urea at 100 times its single-print concentration, and lactic acid at 100 times its single-print concentration. All amino acids (DL-aspartic acid, DL-glutamic acid monohydrate, glycine, DL-histidine, L-ornithine HCl, DL-serine, DL-threonine) and urea were purchased from Aldrich (Milwaukee, WI) at purity levels of greater than 98%; lactic acid (an 88% racemic mix in water) was purchased from Mallinckrodt (Phillipsburg, NJ). To prevent bacterial growth contamination of these standards, all solutions were biologically sterilized by filtering through 0.1-μm pore size, 32-mm-diameter acrodisc sterile syringe filters with Supor membrane (Pall Corporation, East Hills, NY). The solution containers were flame sterilized. All sterilization procedures were carried out in a biological hood.

Photo-Degradation Studies

Using gas-tight syringes for precise volume control, 28 μL (equivalent to the standard volume of material in one fingerprint, based on calculations from data in [1,2] of the individual 5× amino acid standards was deposited on separate sets of seven Teflon® disks (22 mm diameter) and seven steel coupons (19.6 mm × 15.8 mm × 4.4 mm), for a total of 98 amino acid samples. The samples were allowed to dry overnight in the dark, before being exposed to a 200–500 Watt Xe/HgXe Arc Lamp intense light source (Oriel Corporation Model 87301, Stratford, CT) for periods of time equivalent to 0, 2, 5, 7, 14, 28, and 56 days. The light source emits visible, UV-A, UV-B, and UV-C wavelengths, simulating 365 days of sunlight exposure in 24 h of lamp use (see Table 2 for the lamp exposure times that

TABLE 1—*Artificial eccrine fingerprint solution.*

Components	mg/L = (μg/mL) for a Single Print	% Composition
Amino acids		
Serine	399.2	7.16
Glycine	191.00	3.42
Ornithine (as Ornithine HCl)	205.46	3.68
Aspartic acid	109.71	1.97
Histidine	100.08	1.79
Alanine	92.59	1.66
Threonine	76.84	1.38
Lysine	57.63	1.03
Valine	54.58	0.98
Leucine	51.71	0.93
Glutamic acid (as monohydrate)	53.26	0.95
Proline	45.38	0.81
Phenylalanine	41.44	0.74
Tyrosine	38.96	0.70
Isoleucine	37.61	0.67
Arginine	31.21	0.56
Citrulline	25.11	0.45
Methionine	10.69	0.19
Taurine	4.48	0.08
Cystine	0.72	0.01
α-Amino-n-butyric acid	0.36	0.01
Fatty acids		
Acetic acid	7.69	0.14
Propionic acid	0.26	0.005
Butyric acid	0.21	0.004
Isovaleric acid	0.11	0.002
Hexanoic acid	0.10	0.002
Isobutyric acid	0.07	0.001
Urea	1180.0	21.15
Sodium chloride (NaCl)	1138.46	20.41
Potassium chloride (KCl)	640.74	11.49
Lactic acid	616.00	11.04
Ammonium hydroxide (NH_4OH)	107.01	1.92
Calcium sulfate ($CaSO_4$)	98.50	1.77
Glucose	70.00	1.25
Pyruvic acid	40.00	0.72
Sodium sulfate (Na_2SO_4)	24.92	0.45
Magnesium sulfate ($MgSO_4$)	15.85	0.28
Creatinine	4.60	0.08
Zinc chloride ($ZnCl_2$)	2.40	0.04
Uric acid	2.02	0.04
Sodium bicarbonate ($NaCO_3H$)	1.45	0.03
Acetylcholine iodide	0.1214	0.00218
Histamine	0.0135	0.00024
Sodium iodide (NaI)	0.0045	0.00008

TABLE 2—*Sample time under Xe/XeHg lamp to simulate given amount of sunlight exposure.*

Time under Xe/XeHg Lamp	Time Exposed to Sunlight (days)
0 h 0 min 0 sec	0
0 h 7 min 53 sec	2
0 h 19 min 44 sec	5
0 h 27 min 37 sec	7
0 h 55 min 14 sec	14
1 h 50 min 28 sec	28
3 h 40 min 56 sec	56

correlated with the desired duration of continuous sunlight exposure). The same sample preparation procedure was repeated with the 100× urea and 100× lactic acid standards, for an additional 28 samples. Teflon® disks were reused throughout the study, with a multi-step cleaning with lab soap, distilled water, acetonitrile, and methanol between each use. Steel coupons were used only once.

Thermal-Degradation Studies

Thin galvanized steel coupons (15.7 mm × 15.7 mm × 0.5 mm) were cleaned with acetonitrile and methanol to remove lubrication oils used in the coupon-cutting process, sanded with 4000-grit (or better) sandpaper to remove any possible surface oxidation, and wiped again with acetonitrile and methanol. Using gas-tight syringes for precise volume control, 28 μL of the 5× standard amino acid mixture was deposited on four steel coupons. The samples were allowed to dry overnight in the dark. A commercially available 120-V, 1200-W heat gun (Black&Decker Model No. 9756, Hunt Valley, MD) was turned on to the "high" setting for at least 3 min. A vise-grip clamp was used to grasp a steel coupon along its edge and suspend it directly over the heat gun, with the nondeposited side facing the heat source. A thermocouple (Omega Engineering, Stamford, CT) in contact with the steel coupon measured the temperature experienced by the side spotted with the fingerprint sample. Once the desired temperature was achieved an estimated time of exposure should be given to demonstrate the attempt of rapid heating (50°C, 100°C, or 150°C), the steel coupon was immediately removed from the heat source and allowed to cool on a wire rack for 3 min. The fourth coupon was not heated, to be used as the ambient reference. The thermal-degradation procedure was repeated with the 100× urea and 100× lactic acid standards.

Amino Acid Sample Analysis

The amino acid samples were placed face-down in glass vials containing 1 mL of a 50 mM aqueous solution of sodium dodecyl sulfate (Fluka, Seelze, Germany); then, the vials were placed in an ultrasonic bath for 30 min to promote extraction. The solvent was recovered and taken through the "EZ:faast™ for Free Physiological Amino Acid Analysis by LC–MS" derivatization kit (Phenomenex, Torrance, CA). This EZ:faast™ sample derivatization consisted of the introduction of a mixture of three internal standards and a solid-phase extraction utilizing amino acid-trapping sorbent particles, followed by the proprietary derivatization and a liquid/liquid extraction. The organic liquid layer was evaporated to dryness under nitrogen then reconstituted in 100 μL of a 2:1 (v/v) mixture of 10 mM ammonium formate (Fluka) in MeOH and 10 mM ammonium formate in H_2O. Samples were analyzed by liquid chromatography coupled to positive-ion electrospray mass spectrometry using an Agilent (Wilmington, DE) 1100 Series LC instrument and Bruker Daltonics (Billerica, MA) Esquire-LC MS instrument with settings as presented in Table 3.

Urea and Lactic Acid Sample Analysis

The urea and lactic acid samples were placed face-down in glass vials containing 5 mL of MeOH:H_2O (1:1, v/v) with 0.15% formic acid (JTBaker Inc., Phillipsburg, NJ). The vials were placed in an ultrasonic bath for 30 min to promote extraction. The solvent was removed and filtered through 0.1-μm Supor membrane sterilized filters (Pall Corporation, East Hills, NY). Urea samples used an internal standard of 10 ppm of melamine (as melamine 99%, Aldrich, Milwaukee, WI); lactic acid samples used an internal standard of 10 ppm of cyanuric acid (Aldrich, Milwaukee, WI). Melamine and cyanuric acid were dissolved in MeOH:H_2O (1:1, v/v) with 0.15% formic acid at a final concentration of 20 ppm. For urea samples, 100 μL of the extracted solution was spiked with 1 mL of internal standard solution and 900 μL of blank extraction solvent. For lactic acid samples, because of the low intensity of the

TABLE 3—*Instrument settings for LC–MS analysis of amino acid photo-degradation and thermal-degradation samples.*

Liquid Chromatography			
Column	Phenomenex® AAA-MS		
Column length	250 mm		
Internal diameter	2.0 mm		
Flow rate	0.25 mL/min		
Column temperature	35°C		
Injection volume	20 μL		
Solvent A	10 mM Ammonium formate in MeOH		
Solvent B	10 mM Ammonium formate in H_2O		
Solvent gradient	Time (min)	Solvent A (%)	Solvent B (%)
	0.00	68	32
	13.00	83	17
	13.01	68	32
Run time	24 min		
Postrun flush	68% solvent A		
Postrun time	4 min		

Mass Spectrometry			
Mode	Positive ion	Skim 1	13.0 V
Capillary voltage	−2700 V	Cp exit offset	60.0 V
End Plate offset	−500 V	Octopole 1 DC	2.70 V
Nebulizing gas	20.0 psi N_2	Octopole 2 DC	2.40 V
Drying gas	5.00 L/min N_2	Trap drive	45.6
Source temperature	365°C		
Target	20,000	Skim 2	8.0 V
Accumulation time	0.25 ms	Cap. exit	73.0 V
Scan range	100–600 m/z	Oct RF	150 V_{pp}
Averages	30	Lens 1	−5.6 V
		Lens 2	−88.0 V

signal, 200 μL of the extracted solution was spiked with 1 mL of internal standard solution and 800 μL of blank extraction solvent. The samples were then analyzed by direct infusion at 2181 μL/h into an atmospheric pressure chemical ionization mass spectrometer (APCI-MS) using a Bruker Daltonics Esquire-LC MS instrument with settings as presented in Table 4.

Because of the high background of the cyanuric acid for the lactic acid sample determination, the analysis was conducted adopting a 4-port switching valve. For each analysis, the sample was injected twice for 24 sec. An infusion of blank solvent extraction at a rate of 4 mL/h was constantly used between sample injection and each run.

Results

A great deal of time was spent developing and optimizing the sample extraction and HPLC analysis protocol. Selections for the HPLC method began with a C18 column and a trifluoroacetic acid (TFA)-based mobile phase gradient. When used with an aqueous mixture of amino acids, co-elution was a major problem. References were found that touted the successful separation, identification, and quantitation of numerous amino acids and related molecules with the C18 column using a tridecafluoroheptanoic acid (TDFHA) mobile phase gradient (11–13). However, co-elution remained an issue during sample runs. Next, a silica column with an ACN:H_2O mobile phase system was tested with the amino acids. Although separation efficiency was improved, problems remained: significant peak broadening and tailing was noted; at 40 min, runs were longer than desired; and several amino acids were not detected (despite increased concentration levels). Finally, the EZ:faast™ free amino acid analysis kit (which utilized a

TABLE 4—*Instrument settings for MS analysis of urea and lactic acid photo-degradation and thermal-degradation samples.*

Mass Spectrometry of Urea Samples

Mode	Positive Ion	Skim 1	13.0 V
Capillary voltage	−3500 V	Cp exit offset	18.0 V
End plate offset	0 V	Octopole 1 DC	2.70 V
Corona	+1400 nA	Octopole 2 DC	2.40 V
Nebulizing gas	25 psi N_2	Trap drive	45.6
Drying gas	4.00 L/min N_2		
Dry temperature	300°C	Skim 2	8.0 V
APCI source temperature	325°C	Cap. exit	31.0 V
Target	20,000	Oct RF	150.0 V_{PP}
Accumulation time	1.00 ms	Lens 1	−5.6 V
Scan range	50–200 m/z	Lens 2	−88V
Averages	100		

Mass Spectrometry of Lactic Acid Samples

Mode	Negative Ion	Skim 1	−40 V
Capillary voltage	+500 V	Cp exit offset	−100 V
End plate offset	−500 V	Octopole 1 DC	−12.0 V
Corona	−1700 nA	Octopole 2 DC	−1.70 V
Nebulizing gas	20 psi N_2	Trap drive	30.2
Drying gas	4.00 L/min N_2		
Dry temperature	300°C	Skim 2	8.0 V
APCI source temperature	325°C	Cap. Exit	−100.0 V
Target	20,000	Oct RF	109.2 V_{PP}
Accumulation time	5.00 ms	Lens 1	5.0 V
Scan range	50–200 m/z	Lens 2	60.0 V
Averages	40		

Phenomenex® AAA-MS column with an COOHNH$_4$ in MeOH:COOHNH$_4$ in H$_2$O mobile phase gradient) was tested and resulted in appropriate resolution, sensitivity, peak shapes, and elution times for the components of the amino acid mixture. Although researchers had initially sought a method of separating and detecting the components as they existed in fingerprint material, the success of the derivatization kit led them to adopt this protocol.

Attention next shifted to optimizing the extraction of fingerprint material from substrates. Numerous extraction solvents/solutions (Table 5) were variously tested until a boiling 50 mM sodium dodecyl sulfate (SDS) aqueous solution was found to consistently extract the amino acids, even at low concentrations, for successful postderivatization LC–MS detection. The complete sample processing and analysis protocol for the amino acid portion of eccrine fingerprints was provided in the preceding section.

In pursuit of one protocol capable of analyzing all constituents of eccrine fingerprints, the SDS extraction and EZ:faast derivatization/LC–MS analysis was attempted with the nonamino acids (i.e., urea and lactic acid). One approach involved the collection and subsequent derivatization of the wash from the solid-phase

extraction step utilizing amino acid-trapping sorbent particles; however, no constituents were detected during LC–MS analysis. A second approach involved the use of specially ordered solid-phase extraction tips designed to trap nonamino organic acids. This approach also failed. Instead, researchers concluded a separate protocol was needed for analysis of the nonamino acid portion of eccrine fingerprints. The starting point was a dinitrophenylhydrazine (DNPH) derivatization followed by LC–MS analysis using a C18 column, ACN:H$_2$O mobile phase system, and electrospray ionization MS. DNPH is a commonly used method of derivatization of LC–MS analysis of aldehydes. Although the targeted constituents were detected, reproducible quantitation was difficult. A more favorable approach was chosen that utilizes an extraction solution of MeOH:H$_2$O (1:1, v/v) with 0.15% formic acid, followed by direct infusion and APCI-MS analysis. The complete sample processing and analysis protocol for the nonamino acid portion of eccrine fingerprints was provided in the preceding section. Before the implementation of the urea and lactic acid analysis, different types of surfaces (Teflon® and galvanized steel) spiked with known amount of analytes were extracted to verify and compare the efficiency of the recovery. After the extraction, lactic acid and urea recovery was greater than 80%.

Photo-Degradation Studies—Amino Acids

The goal of these studies was to determine whether any of the tested amino acids experienced photo-induced degradation, and if so, to determine the rate of degradation and identify any detectable degradation products. This was accomplished by monitoring the relative ratio of the peak area of each amino acid to the peak area of the closest-eluting internal standard. Data processing used BRUKER DALTONICS software (Billerica, MA), and consisted of choosing the extracted ion chromatograms for the amino acid/internal standard, subtracting the baseline, applying 3-cycle smoothing of the chromatogram, and manually integrating the peak area. The literature accompanying the derivatization kit provided the m/z values to monitor for the amino acids and internal standards (Table 6), as well as the instructions to use the first internal standard (homoarginine, HARG) for comparison with SER, GLY, and THR and the second internal standard (methionine-d_3, Met-d_3) for comparison with ORN, ASP, HIS, and GLU. For each of the seven amino acids, the relative ratio of the amino acid peak area to the internal standard peak area remained constant across the seven timepoints within the sample set (the data for ORN on Teflon® and GLY on steel, chosen to illustrate the amino acid data sets, are provided in Figs. 1 and 2). No notable photo-degradation of the amino acids studied was observed on either Teflon® or stainless steel surfaces.

TABLE 5—*Solvents/solutions tested for extraction of amino acids.*

Solvent/Solution	Chosen for Extraction?
Water	No
Formic acid	No
Ethanol	No
Acetonitrile	No
0.1 M Trifluoroacetic acid (aq)	No
1.0 M Ammonium acetate (aq)	No
1.0 M Ammonium acetate (aq) with 1% formic acid	No
1.0 M Ammonium acetate (aq) with 10% formic acid	No
50 mM Sodium dodecyl sulfate (aq) – (boiling)	Yes

TABLE 6—*Basic information on AA ions after derivatization kit.*

Chemical Name	Abbreviation	Reported t_R (min)	AA MW (g/mol)	Derivatized AA MW+1 (m/z)
Homoarginine	HARG	2.9	188.2	317
Serine	SER	3.7	105.1	234
Glycine	GLY	4.3	75.1	204
Threonine	THR	4.3	119.1	248
Ornithine	ORN	6.9	132.1	347
Methionine-d_3	Met-d_3	6.9	152.2	281
Aspartic acid	ASP	7.8	133.1	304
Histidine	HIS	7.8	155.1	370
Glutamic acid	GLU	8.3	147.1	318
Homophenylalanine	HPHE	12.2	179.2	308

Photo-Degradation Studies—Urea and Lactic Acid

The choice of extraction solvent (a formic acid-containing water/methanol mix) for the urea and lactic acid samples is compatible with direct infusion into the MS via an APCI source. Therefore, the collected data consists of spectra; but as with the amino acid, samples, peak area ratios are employed to monitor degradation. The urea samples were analyzed in positive-ion mode, with urea producing an M^+ ion at m/z 61, and the melamine used as an internal standard producing an M^+ ion at m/z 127. The lactic acid samples were analyzed in negative ion mode, with lactic acid producing an M^+ ion at m/z 89, and the cyanuric acid used as an internal standard producing an M^+ ion at m/z 128. Two sets of lactic acid and urea samples were exposed to the Xe/HgXe arc lamp for intervals simulating the following periods of time: 0, 2, 5, 7, 14, 28, and 56 days. Each sample was then injected twice for 24 sec. This approach provided sufficient data to yield statistically meaningful qualitative information about the physical and chemical characteristics of the analytes. As the amino acids experiments, the aim of these studies was to determine urea and lactic acid photo-induced degradation, the rate of degradation, and identify any degradation products. For each set of samples, the amount of lactic acid and urea recovered on Teflon® disks and steel coupons was

calculated. For pyruvic acid, the ingrowth was calculated only as its relative ratio (on Teflon® disks and steel coupons) to the internal standard peak area, because it was a degradation product and not an eligible component of this study (Figs. 3–8). According to the results from Figs. 3 and 4, urea exposed to light did not appear to undergo degradation, even for long exposure times. This set of experiments sought to confirm what is well known, urea does not undergo photo-degradation. A different result, however, appears in Figs. 5 and 7 for lactic acid samples exposed to light. The ingrowth of pyruvic acid was found to correlate with lactic acid photo-degradation. Light seems to play a fundamental role in the production of pyruvic acid as the oxidation product.

Currently, in the scientific literature, there are no studies published on lactic acid in fingerprints. However, one article by Pérez-Ruiz discussed the decomposition of lactate with an unidentified degradation product in presence of iron upon irradiation with UV or visible light (14). The pyruvic acid M^+ ion at m/z 87 was monitored in all the samples analyzed and correlated with the amount of lactic acid. The photo-degradation of lactic acid started in a relatively short period of time. According to our results, after only 2 days' equivalent of light exposure, the amount of lactic acid was reduced prominently from the initial concentration deposited. Comparing the slope of the curves, decay on steel is faster than on
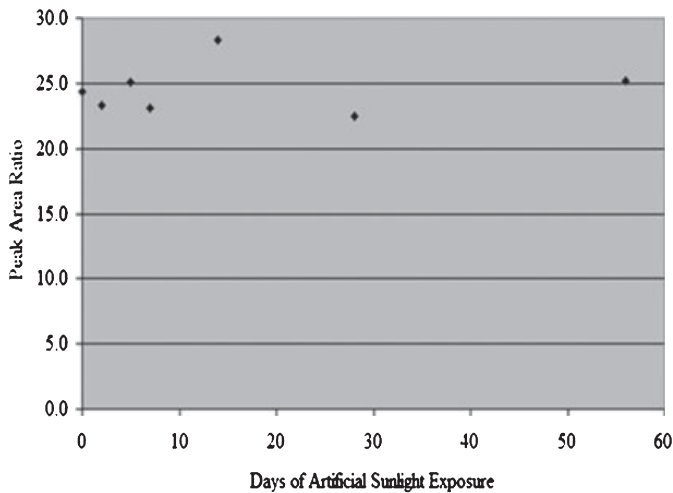
FIG. 1—*Photo-degradation experiments of ornithine on Teflon®.*
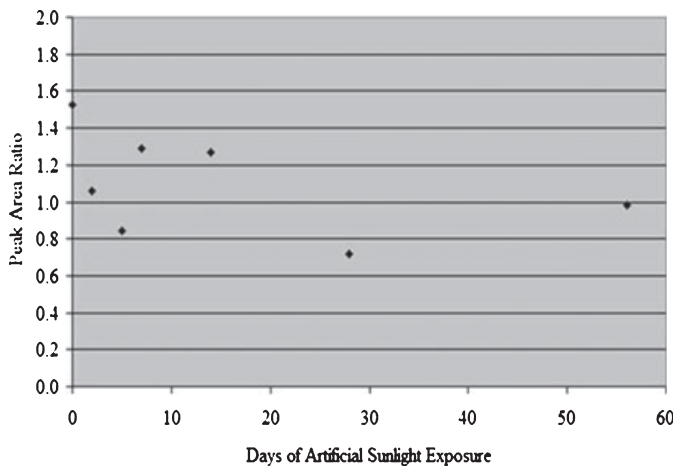
FIG. 2—*Photo-degradation experiments of glycine on steel.*
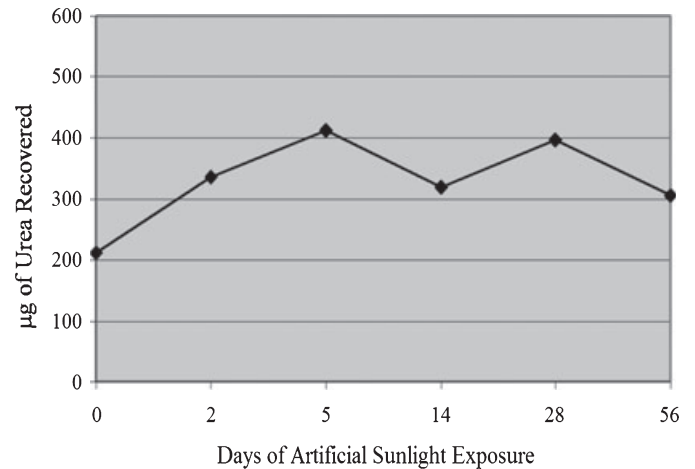
FIG. 3—*Photo-degradation experiments of urea on steel.*
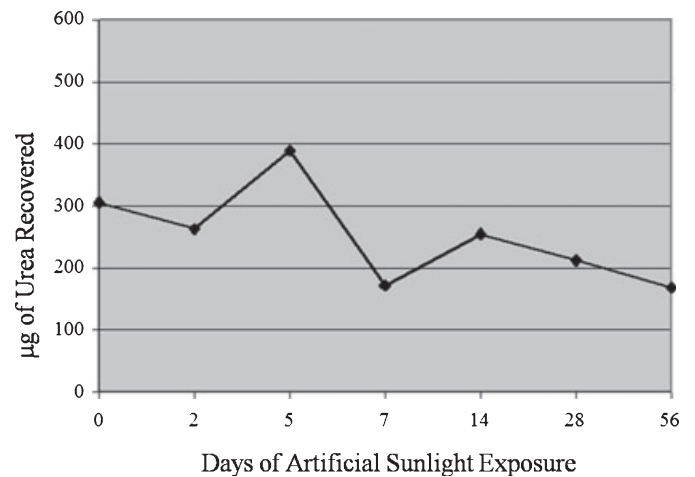
FIG. 4—*Photo-degradation experiments of urea on Teflon®.*
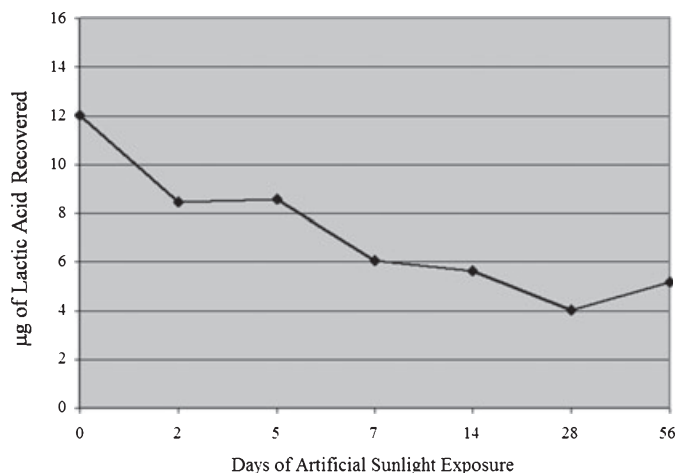
FIG. 5—*Photo-degradation experiments of lactic acid on steel. The equation of the curve is:* $y = 0.2279x^2 - 2.9743x + 14.465$.
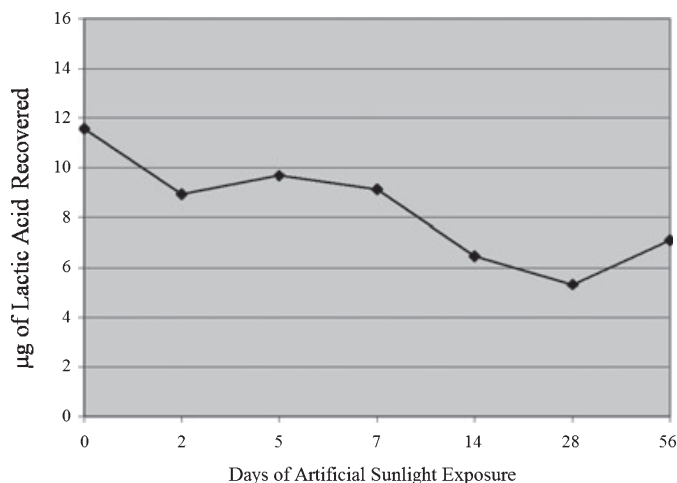
FIG. 7—*Photo-degradation experiments of lactic acid on Teflon[®]. The equation of the curve is:* $y = 0.1x^2 - 1.6458x + 12.898$.
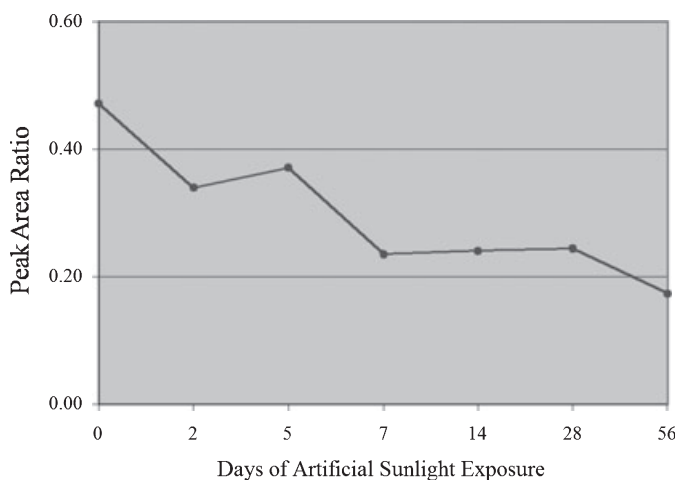
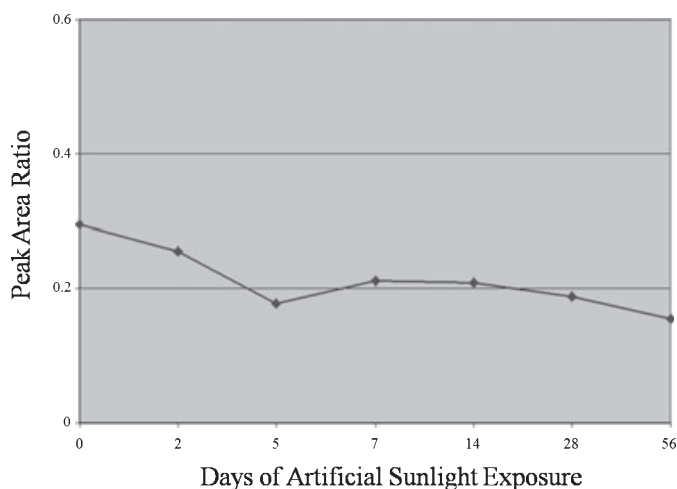FIG. 6—*Photo-degradation experiments of pyruvic acid on steel.*

FIG. 8—*Photo-degradation experiments of pyruvic acid on Teflon[®].*

Teflon[®]. With its high iron content, the steel itself is involved in the lactic acid degradation process, along with the light. The result of this cooperation would be an increase in the speed of the degradation reaction. On the contrary, Teflon[®] would act only as an inert platform for the experiments. After monitoring the pyruvic acid peak, it was observed that, along with the lactic acid decomposition, the pyruvic acid also degrades under exposure to light as well as air (Figs. 6 and 8). However, correlating the rate of degradation of pyruvic acid to lactic acid we noticed an inflection point around the 28th day of equivalent light exposure, indicating that something was inhibiting the rate of lactic acid degradation. The fact that the rate of degradation did not follow the same rate confirmed that many factors, both chemical and physical, were involved in the degradation process. We speculate that the rate could be influenced by a threshold amount of degraded lactic acid necessary to produce a detectable amount of pyruvic acid. To reach this amount, another parameter should be considered: the time required by the light to reach and degrade layers of lactic acid underneath the outer layer. This period of time could be around 4 weeks.

Another interesting result is the slightly earlier ingrowth of pyruvic acid on the Teflon[®] relative to steel. This difference may indicate an interaction between the lactate and components of the
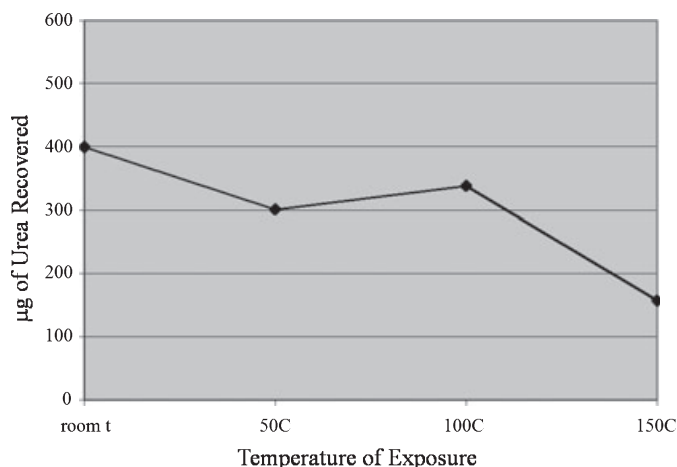
FIG. 9—*Thermal-degradation experiments of urea on steel.*

steel, which temporarily delays conversion to pyruvic acid relative to the inert Teflon[®], where the light acts alone to convert lactic to pyruvic.

FIG. 10—*Thermal-degradation experiments of lactic acid on steel.*
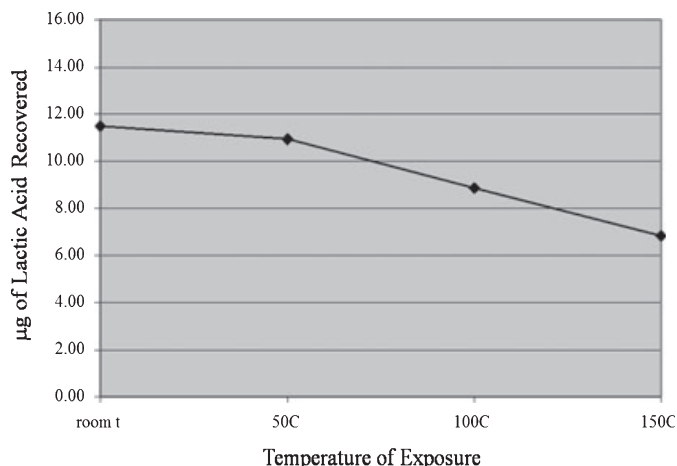
FIG. 11—*Thermal-degradation experiments of pyruvic acid on steel.*

Thermal-Degradation Studies—Amino Acids

All seven amino acids deposited on steel were degraded to some extent when exposed to heat. Serine and glycine data yielded a large-scale reduction in detected amino acid for the sample heated to 150°C, and no peak was present for threonine at that temperature. Ornithine, aspartic acid, histidine, and glutamic acid all experienced a similar degradation pattern: there was substantial amino acid loss incurred at the 100°C temperature point, and additional loss at the 150°C sample. Unfortunately, when the chromatograms were analyzed for ingrowth products, none could be found.

Thermal-Degradation Studies—Urea and Lactic Acid

Two sets of samples of lactic acid and urea on steel were analyzed at different temperatures: room temperature, 50°C, 100°C, and 150°C. Each sample was then injected twice for 24 sec. For the samples containing lactic acid, the M^+ peak of pyruvic acid at m/z 87 was also monitored. For each set of samples, the amount of lactic acid and urea recovered, and the relative ratio of pyruvic acid to the internal standard peak area on steel coupons was calculated (Figs. 10–11).

Already an object of studies (15), our results confirmed the fact that thermal decomposition of urea was initiated around 150°C.

In contrast, the lactic acid thermal-degradation process started between 50 and 100°C along with pyruvic acid generation and subsequent degradation.

Discussion

Photo-Degradation Experiments

APCI mass spectrometry of urea and HPLC of amino acids showed that these components were not affected by the action of light (even for long times of exposure) or interaction with Teflon® and/or steel. In contrast, photo-degradation experiments on lactic acid indicated a significant reduction of the initial concentration of the acid itself after just 2 days of sunlight exposure, with a concomitant degradation of its ingrowth product, pyruvic acid. However, for lactic acid samples on steel and Teflon®, the degradation rate of pyruvic acid ingrowths to the lactic acid present, slowed down around the 28th day of light exposure.
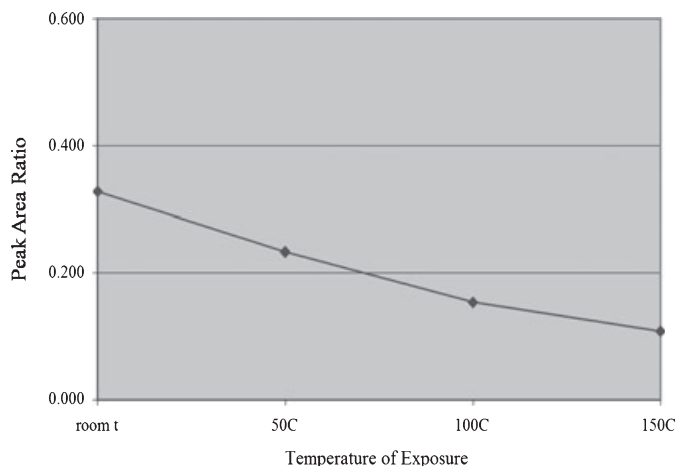
Thermal-Degradation Experiments

Analysis of amino acids exposed to heat established the temperature tolerated prior to the complete decomposition process for this class of components. All seven amino acids deposited on Teflon® disks and steel coupons exposed to 150°C were undetectable in the analysis with HPLC. However, no degradation product was found.

Similar results were obtained for urea, indicating the onset of degradation at 150°C.

Both lactic acid and its photo-oxidation product pyruvic acid are affected by heat exposure. The beginning of their decomposition process was marked at 50°C.

Results achieved in this work are expected to be useful to the future development of fingerprint visualization methods. Specifically, the target of these methods would be latent fingerprints arisen from typical scenarios where the application of the superglue fuming technique could be fallible because of the degradation of the components that initiate polymerization.

References

1. Richmond-Aylor A, Bell S, Callery P, Morris K. Thermal degradation analysis of amino acids in fingerprint residue by pyrolysis GC–MS to develop new latent fingerprint developing reagents. J Forensic Sci 2007;52(2):380–2.
2. Chiavari G, Galletti GC. Pyrolysis-gas chromatography/mass spectrometry of amino acids. J Anal Appl Pyrolysis 1992;24:123–37.
3. Lewis LA, Smithwick RW, Devault GL, Bolinger B, Lewis SA Sr. Processes involved in the development of latent fingerprints using the cyanoacrylate fuming method. J Forensic Sci 2001;46(2):241–6.
4. Ramotowski RS. Composition of latent print residue. In: Lee HC, Gaensslen RE, editors. Advances in fingerprint technology, 2nd edn. Washington, DC: CRC Press, 2001;63–104.
5. Lentner C, editor. Geigy scientific tables. Volume I: Units of measurement, body fluids, composition of the body, nutrition. 8th rev. edn. West Caldwell, NJ: Ciba-Geigy Limited, 1981;108–12.
6. Bernier UR, Kline DL, Barnard DR, Schreck CE, Yost RA. Analysis of human skin emanations by gas chromatography/mass spectrometry. 2 Identification of volatile compounds that are candidate attractants for the yellow fever mosquito (Aedes aegypti). Anal Chem 2000;72(4):747–56.
7. Asano KG, Bayne CK, Horsman KM, Buchanan MV. Chemical composition of fingerprints for gender determination. J Forensic Sci 2002; 47(4):805–7.

8. Hamilton PB. Amino-acids on hands. Nature 1965;205(4968):284–5.
9. Noble D. Vanished into thin air: the search for children's fingerprints. Anal Chem 1995;67:A435–A438.
10. Connatser RM, Prokes SM, Glembocki OJ, Gardner CW, Schuler RL, Lewis SA, et al. Toward surface enhanced Raman imaging of latent fingerprints. J Forensic Sci. In press.
11. Piraud M, Vianey-Seban C, Petritis K, Elfakir C, Steghens JP, Moria A, et al. ESI-MS/MS analysis of underivatized amino acids: a new tool for the diagnosis of inherited disorders of amino acid metabolism. Fragmentation study of 79 molecules of biological interest in positive and negative ionization mode. Rapid Commun Mass Spectrom 2003;17:1297–311.
12. Piraud M, Vianey-Saban C, Petritis K, Elfakir C, Steghens J-P, Bouchu D. Ion-paring reversed-phase liquid chromatography/electrospray ionization mass spectrometric analysis of 76 underivatized amino acids of biological interest: a new tool for the diagnosis of inherited disorders of amino acid metabolism. Rapid Commun Mass Spectrom 2005;19:1587–602.
13. Piraud M, Vianey-Seban C, Bourdin C, Acquaviva-Bourdin C, Boyer S, Elfakir C, et al. A new reversed-phase liquid chromatographic/tandem mass spectrometric method for analysis of underivatised amino acids: evaluation for the diagnosis and the management of inherited disorders of amino acid metabolism. Rapid Commun Mass Spectrom 2005;19:3287–97.
14. Pérez-Ruiz T, Martínez-Lozano C, Tomás V, Martín J. Flow injection determination of lactate based on a photochemical reaction using photometric and chemiluminescence detection. The Analyst 1999;124:1517–21.
15. Fang HL, DaCosta FH, (Cummins Inc.). Thermolysis characterization of urea-SCR. Proceeding of the 8th Annual Conference of the Engine-Efficiency and Emissions Research (DEER); 2002 Aug 25–29; San Diego (CA), USA: Department of Energy, 2002.

Additional information and reprint requests:
Samuel A. Lewis, Sr., M.S.
Oak Ridge National Laboratory
NTRC Building MS 6159
2360 Cherahala Blvd.
Knoxville, TN 37932
E-mail: lewissasr@ornl.gov

PAPER

GENERAL; DIGITAL & MULTIMEDIA SCIENCES

Abhishek Gupta,[1] *M.Sc., M.Phil and Raul Sutton,*[1] *Ph.D.*

Pore Sub-Features Reproducibility in Direct Microscopic and Livescan Images—Their Reliability in Personal Identification

ABSTRACT: Third level features have been reported to have equal discriminatory power as second level details in establishing personal identification. Pore area, as an extended set third level sub-feature, has been studied by minimizing possible factors that could affect pore size. The reproducibility of pore surface area has been studied using direct microscopic and 500 ppi Livescan images. Direct microscopic pore area measurements indicated that the day on which the pore area was measured had a significant impact on the measured pore area. Pore area measurement was shown to be difficult to estimate in 500 ppi Livescan measurements owing to lack of resolution. It is not possible to reliably use pore area as an identifying feature in fingerprint examination.

KEYWORDS: forensic science, biometrics, fingerprints, third level detail, poroscopy, personal identification, livescans, direct microscopic images

Forensic science concerns the individualization of material found at scenes of crime (1). Among these, latent fingerprint marks have been shown to be capable of great discriminatory power and are therefore capable of individualizing with high resolution. Latent fingermarks left at the scene of crime are sometimes partial depositions and are not often protected from environment, thus can get affected by atmospheric agents. This may result in recovery of fragmented and partial fingermarks. Also, fragmented fingermarks are recovered from scene of crime where friction skin ridge comes in contact with folded surfaces or when only part of the finger contacts the surface. Comparing these partial fingermarks against the prints in a database can pose several problems when there is insufficient second level detail (ridge pattern analysis), unspecified orientation, and nonlinear distortion of these partial fingermarks, which reduces the discriminatory power. Thus, the forensic science community has been active in extending the usable features in fingermarks to increase the individualizing power of latent fingermarks. One of these is to use third level detail such as the information provided by pores and ridges in personal identification (2).

The science of poroscopy is based on the fact that the pores are permanent, immutable, and variable from one individual to another (3). Thus, an identification based on poroscopy can be as accurate and reliable as that based on ridge endings (4). Various attempts have been made (4–13) to ascertain which third level features can be used in personal identification and how reliable these features can be in establishing personal identity.

Most of the recent studies (8–10,13,14) have suggested that a Livescan resolution of 500 ppi is not sufficient enough to extract enough pore details for comparison. However, there is a discrepancy in the minimum scanning resolution required to study third

level features. Different feature extraction models have been applied by various researchers to scanned images to extract third level features. Most of these are focused on pore extraction to utilize pore location, and we are unaware of any automated feature extraction methods that attempt to match using pore area or shape.

Nonlinear (elastic) distortion is one of the major challenges in fingerprint matching system. Researchers have attempted to determine and resolve distortion introduced when a three-dimensional finger surface touches a two-dimensional image acquisition surface (15–17). As pores are located on fingerprint ridges, they are equally affected by this nonlinear distortion. In the present study, estimation of pore area in direct microscopic images is a step forward, eliminating nonlinear distortion.

Electronic Livescan digital images are replacing the traditional inked fingerprints as reference prints. In the United Kingdom, Livescan images are routinely captured at 500 ppi (18), and this scanning resolution is adequate to extract first and second level detail. Most studies considering third level features have proposed that scanning resolution should be higher to study third level details (4,6,8,19). One recent study has suggested that pores can be extracted in Livescan images captured using 500 ppi optical livescanner (20). The present study is aimed to determine whether pore size can be measured in Livescans captured at 500 ppi.

Pore area has been reported to be nonreproducible in inked prints (11) and latent fingermarks (12). In previous studies, attempts were made to minimize the causes (5), which could lead to nonreproducibility of pore area. In the present study, the reproducibility of third level detail (pore area) in direct microscopic images and Livescans has been investigated. In studying direct microscopic images of fingerprint pores, this study will evaluate the causes of variability in previously published work (11,12). The variability in latent developed fingermarks and inked reference prints could be because of the deposition method or actual variation in the size of the pore owing to physiological factors. Direct microscopic images will

[1]School of Applied Sciences, University of Wolverhampton, Wolverhampton WV1 1LY, U.K.

allow discrimination between deposition and physiological effects in pore size variation. This is important to establish, as if the deposition method contributes to pore area variability, then alternative image capture methods that avoid print deposition may yield reliable data sets for using this feature of fingerprint pores in individualization. As far as we are aware, no prior research has compared reference inter-image variation of pore area in direct microscopic images. The present study investigated ways of capturing live images using microscope, and inter-image variability has been explored. The results obtained are compared to the relative pore area in inked and latent prints presented in previous work (11,12).

Materials and Methods

Direct Optical Microscopic Images of Fingermarks

Direct microscopic images were captured using a SMZ-2T Nikon optical microscope (focused at 20× magnification) coupled to a digital camera (Nikon Coolpix 4500 using an adaptor with the camera set at automatic mode). The left index finger was secured using blu-tack and the subject asked to relax the secured hand. The central area around the core of the left index finger's friction ridge skin was cleaned before being focused under the microscope. This direct microscopic image capture process was repeated on five different days over the period of 34 days, that took no account of the physical and psychological state of subject or environmental factors such as ambient temperature, humidity, time of the day. Twenty images were taken on each day at regular intervals over the period of 1 h. These images were then calibrated using a reference marker (1 by 1 mm grid) and the "Spatial Calibration" option in Image Pro Plus.

Four pores, from the central portion near to the core of the friction ridge skin of distal phalanx of left index finger, were randomly selected to study pore size variation. The precision of the measurement method was validated before taking area measurements. Area of these pores was estimated taking best fit of the pore by drawing the boundaries around the margin of the pore with the aid of Adobe Photoshop (Fig. 1).

Livescan Images

Images of both right and left index fingers were captured at Home Office Scientific Development Branch (HOSDB) using an L Scan Guardian scanner manufactured by Cross Match Technologies. These images were collected at 500 ppi resolution by pressing the finger against the platen of the scanner. Livescan images were collected and analyzed using Image Pro Plus software (version 4.5, Media Cybernetics Inc., Bethesda, MD) and Adobe Photoshop (version Adobe CS2, Adobe Systems Inc., San Jose, CA). Images were captured at low, medium, and high pressure judged qualitatively, and results are shown in Fig. 2a–c.

Livescan images of both the right and left index fingers were analyzed using Image Pro Plus software as described previously (12) as well as Adobe Photoshop. Pore detail was studied in these images to see the reproducibility of pore area in Livescan images.

Results and Discussion

Results of Pore Area Reproducibility in Direct Microscopic Images

It was established that the measurement method is precise with %C.V within acceptable limits. One hundred images of the left index finger were captured and four pores (1–4) were chosen for measurement in these images (see Fig. 3a–e example images). The data obtained were analyzed statistically to study the reproducibility of pore area in direct microscopic images.

Pore area variability in direct microscopic images was studied by measuring each of the four selected pores in 100 images captured over a period of 5 days. Firstly, the variability of pore area was studied in direct microscopic images captured on the same day.
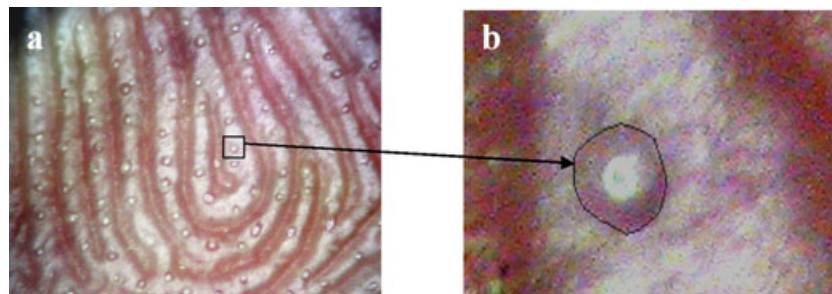

FIG. 1—(a) Direct microscopic image of left index finger at 20× magnification. (b) Enlarged view showing the measurement method of pore area estimation using Adobe Photoshop. As can be seen, the center of the pore (white bright spot) is filled with sweat.

FIG. 2—Results of Livescan images of right index finger at (a) low pressure, (b) medium pressure, (c) high pressure.
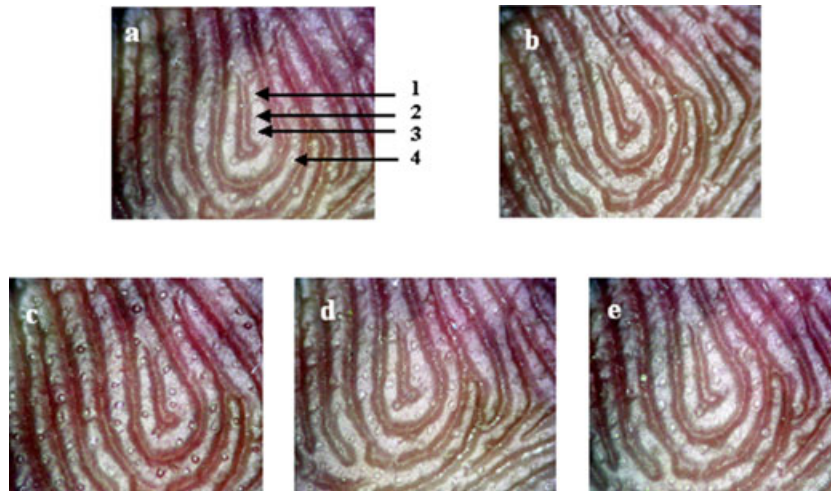
FIG. 3—*Exemplars of magnified (20×) direct images of left index finger (a) image showing pores 1–4 on day 1, (b) image captured on day 15, (c) image captured on day 30, (d) image on day 33, (e) image on day 34.*

The area of pore 1 was measured in each of the 20 images captured on day 1, and mean area and %C.V were calculated. This was repeated for pore 1 on days 15, 30, 33, and 34. The results of the study are shown in Table 1.

Analysis of data revealed that the pore area measured in images captured on day 1 varied between 56,418.0 µmsq and 66,498.6 µmsq with %C.V of 4.2%. The results show that %C.V for pore area estimated in images taken on same days lies within acceptable limits.

In the same manner, data were collected for pores 2–4 on day 1, day 15, day 30, day 33, and day 34, and mean of area and %C.V were calculated. The results of mean area and % C.V of pores 1–4 on different days are shown in Table 2.

Taking into account the value of %C.V (Table 2), which is nearer to the acceptable limit, it can be established that pore area is reproducible within a batch of direct images captured over a single hour on the same day.

Secondly, the reproducibility of pore area was studied by comparing images captured on different days. The mean of mean areas obtained on five different days for pore 1 was calculated to determine %C.V. This was repeated for pores 2–4. Results of this analysis are shown in Table 3.

The results show that %C.V lies over and above the normal 5% acceptable limits for analytical work, indicating that pore area is not reproducible in direct microscopic images captured on different days, which is clearly presented in Fig. 4.

When comparing the variability of pore area in direct microscopic images to inked (11) and latent prints (12) developed using cyanoacrylate and ninhydrin, it was noticed that there is much less variability in direct images. This indicates that the deposition and development method make significant contributions to the pore area variability previously observed.

TABLE 1—*Summary of mean area and % coefficient of variance (% C.V) of pore 1 measured in 20 images each on five different days.*

	Mean Area (μm^2)	Standard Deviation	% C.V
Day 1	62,261.9	2592.2	4.2
Day 15	42,128.1	2144.0	5.1
Day 30	58,966.7	2402.7	4.1
Day 33	68,796.5	2129.3	3.1
Day 34	68,946.3	2414.5	3.5

TABLE 2—*Summary of mean area of pores and % C.V of mean area of pores (1–4) measured in 20 images each on five different days.*

	Pore Area (μm^2)			
	Pore 1	Pore 2	Pore 3	Pore 4
Mean area of pores				
Day 1	62,261.9	58,140.5	49,100.1	54,426.9
Day 15	42,107.9	40,947.7	28,911.6	40,268.9
Day 30	58,966.7	53,183.5	48,956.3	67,214.3
Day 33	68,796.5	45,989.1	48,823.1	67,549.5
Day 34	68,946.3	58,214.2	70,945.8	76,097.8
%C.V of mean area of pores				
Day 1	4.2	4.0	3.6	6.2
Day 15	5.2	4.7	4.8	5.0
Day 30	4.1	5.1	3.8	3.9
Day 33	3.1	6.4	4.4	4.5
Day 34	3.5	4.1	2.9	2.5

TABLE 3—*Summary of mean of mean pore areas (μm^2) and % coefficient of variance (% C.V) of pores (1 – 4) measured in 100 images captured on five different days.*

Pore	Mean of Mean Area (μm^2)	Standard Deviation	% C.V
1	60,215.9	10,148.0	16.8
2	51,295.0	7290.8	14.2
3	49,347.4	13,496.0	27.3
4	61,111.5	12,843.8	21.0

Statistical Analysis

Data obtained were further analyzed using the statistical software package SPSS, version 12 for Windows (SPSS, Chicago, IL). Before analysis, Kolmogorov–Smirnov test (K–S test) for Normality and Levene's test for Homogeneity of Variance were applied to the data sets of pore 1, to determine whether they met the criteria for one-way analysis of variance (ANOVA). The K–S test shows that the data were normally distributed. Using Levene's test, the value of $p > 0.05$ thus, the assumption of homogeneity was met. ANOVA was applied using "area" as the dependent variable and the results are presented in Table 4. The results show that the pore area on five different days was significantly variable ($F_{4, 95} = 437.086$, $p < 0.001$).
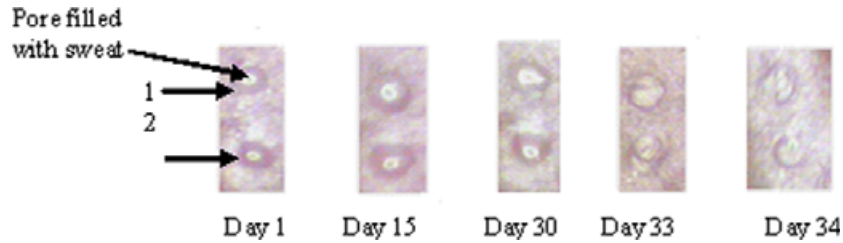
FIG. 4—*Picture showing pore 1 and 2 in magnified direct images of left index finger captured on five different days.*

TABLE 4—*Summary of the results obtained using ANOVA. Tests of between-subjects effects. Dependent variable: area. R squared = .948 (adjusted R squared = .946).*

Source	Type III Sum of Squares	Df	Mean Square	F	Significance
Corrected model	9669864055.018(a)	4	2417466013.754	437.086	0.000
Intercept	362595077977.994	1	362595077977.994	65558.419	0.000
Days	9669864055.018	4	2417466013.754	437.086	0.000
Error	525432628.894	95	5530869.778		
Total	372790374661.906	100			
Corrected total	10195296683.912	99			

Livescan Images

Figure 2a represents Livescan image captured by pressing the right index finger against scanner platen with low pressure. Figures 2b and c represent images captured applying medium and high pressure, respectively. It was noticed that pores were more visible in images deposited at low and medium pressure when compared to the ones captured using high pressure (Fig. 5).

The pores, although visible in images captured using low and medium pressure, were not measurable using Image Pro Plus. When these images were magnified to measure the area occupied by a pore, it was found that image was highly pixellated with pores occupying 4 pixels. This meant that accurate pore margins could not be differentiated. As can be seen in Fig. 6b for an example of magnified image of pore 1, the area occupied by a pore is 4 pixels but that also is not very defined which shows that pore size estimation in these images could not be done accurately. The pore area measurement was also tried using Adobe Photoshop but pores were still not measurable.

Conclusion

The reliability of pore area, as a tool in personal identification when using direct microscopic images and Livescan images at 500 ppi as reference prints, was studied in this experiment. Little variation in pore area was noted when pores were examined in
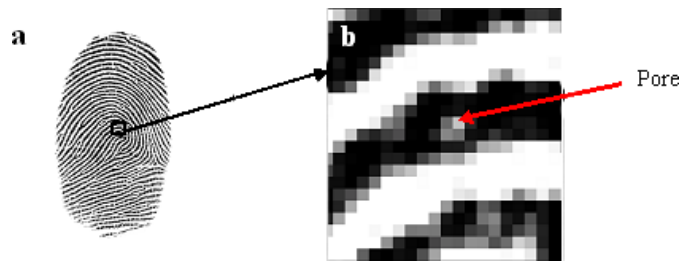
FIG. 6—*(a) Livescan at medium pressure. (b) Enlarged view in Adobe Photoshop showing that pores are not measurable.*

direct microscopic images captured at regular intervals over the period of 1 h on same day with %C.V lying near to or below acceptable values in all cases. Further study of the same pore area measured in direct images taken on different days revealed much more variation and high %C.V value, which is well in excess of 5%, putting in doubt the reliability of pore area in direct microscopic images as the means of personal identification.

Given the results observed for inked reference prints (11) and latent fingermarks (12), one may think that it is irrelevant to pursue the study of pore area direct microscopic and Livescan images, but in the present study, the possible causes mentioned by Ashbaugh (5), which could lead to nonreproducibility of pore area, have been minimized in direct microscopic images. In direct microscopic

FIG. 5—*Pore visibility demonstrated in magnified Livescan images captured at low, medium, and high pressure, respectively.*

images, we have made an attempt to capture images of friction ridge skin without any of the distortion factors inherent in deposition recording of fingerprint images.

Although it was established that pore detail can be studied in Livescan prints at the higher resolution of 1000 ppi (6–9), it has been proposed by Ray et al. (20) that pores can be extracted in Livescan images at 500 ppi by applying an algorithm of modified minimum square error approach to the images. In the present work, study of pore size reproducibility in Livescan images at 500 ppi is an attempt made to explore if pore area is detectable and measurable in scans at 500 ppi. This is not surprising when one attempts to calculate the area that an individual pixel would occupy at 500 ppi. The average pore area for all the pores measured in this article by direct microscopy is c. 61,000 μm^2. If one assumes that the pores are circular (although in our experience most pores appear to be elliptical in shape, this does not affect the conclusions reached here), then the diameter of the average pore can be calculated at 280 μm. Five hundred pixels per inch means that each pixel occupies 0.002 inches which is 0.051 mm or about 50 μm. This means that each pore would be about 5–6 pixels in diameter. Our Livescan-captured image (see Fig. 6b) suggests that a pore is c. 2 pixels in diameter. However, the direct microscopic images take no account of pressure distortion when taking Livescan images that may affect the size of pores images when captured using this method. The effect of this on pore area can be seen when you compare area estimates from inked or developed latent prints (see 11,12). In those articles, mean area was established as being closer to 9000 μm^2. Making the same assumptions as earlier in this paragraph gives a pore diameter of about 100 μm in diameter or 2 pixels. Given the resolution limits of frustrated total internal reflection (FTIR) Livescan image capture technology, our estimates of pore diameter (2 pixels) are very close to the observed value seen in Fig. 6.

Microscopes coupled to inbuilt photographic image capture and associated image analysis software were used to capture photomicrographs and to take measurements to avoid missing of pore detail. Capturing of direct microscopic images was an attempt to avoid any possible factor that may affect the pore size like substrate effect, over or under inking, smudging and smearing of ink, nonlinear distortion and effects because of development method.

Pore area reproducibility was studied in these images, and on statistical analysis, the value of $p \leq 0.000$ that is less than significance level of 0.01, indicating that there is highly significant interprint variation in pore area. Despite avoiding these factors, pore area reproducibility could not be established in direct microscopic images. But it has been found that direct microscopic images are a reliable method for pore image capture. While measuring the pores, the results showed that direct microscopic images give best size estimation as these are live images and there was no inking or development method involved that could lay some sort of impact on actual size of the pore. Also, it was found that some pores measure more than 220 μm in diameter that is not in accordance with the findings of Locard (3). They came up with results that the diameter of pores varies between 88 and 220 μm. Our study has found that pore diameter can be up to 265 μm. Initial attempts by Roddy and Stosz (7) to study the detection, size, and shape of pores in latent, inked, and Livescan prints concluded that pore size can be best estimated in inked prints. We challenge their conclusion, as already presented (12), pore area is not reproducible even in these images.

Although pores can be extracted in Livescans captured at 500 ppi (20), measuring pore area accurately in these scans is a big challenge. Thus, reproducibility of pore area could not be studied

in Livescan images because of the difficulty in extracting enough pore detail at 500 ppi. As can be seen in Fig. 6b, at 500 ppi, the area occupied by the pore is between 1 and 4 pixels, with no obvious marked boundaries. The amount of information is simply too sparse to give accurate pore area measurement for fingerprint-matching purposes. Our study supports extensive discussions that pore detail is not clear at 500 ppi. Observations made in case of Livescan images at 500 ppi are in accordance with those of Jain et al. (9). In their article, they stated "500 pixels per inch (ppi), which is inadequate to capture Level 3 features, such as pores." They suggested that the resolution should be extended from 500 to 1000 ppi to capture measurable pore detail. Revisiting the earlier calculations, even at 1000 ppi, Livescan pore images are expected to be 4 pixels in diameter, so even at this resolution, accurate pore area estimations would not be possible (12 pixel pore area mean that 1 pixel difference would mean a variation of 8% before other factors are considered).

This study further confirms that pore area in direct microscopic images is subject to too great a variation to be relied upon for making personal identification. By establishing nonreproducibility of pore area with statistical data, we have excluded use of pore area as an extended set feature in personal identification. Further research in third level features of fingerprints should be focussed elsewhere.

References

1. Kirk PL, Kingston CR. Evidence evaluation and problems in general criminalistics. J Forensic Sci 1964;9(4):434–44.
2. Ashbaugh DR. Quantitative-qualitative friction ridge analysis. Boca Raton, FL: CRC Press, 1999.
3. Locard E. Les pores et l'identification des criminels. Biologica: Revue Scientifique de Medicine 1912;2:357–65.
4. Jain A, Chen Y, Demirkus M. Pores and ridges: fingerprint matching using level 3 features. 18th International Conference on Pattern Recognition (ICPR'06); 2006 Aug 20–24; Hong Kong. IEEE Comput Soc 2006;(4):477–80.
5. Ashbaugh DR. Poroscopy. Identification news. 1982. Available at: http://www.scafo.org/library/110601.html (accessed February 26, 2006).
6. Stosz JD, Alyea LA. Automated system for fingerprint authentication using pores and ridge structure. SPIE—The International Society for Optical Engineering 1994;2277:210–23.
7. Roddy AR, Stosz JD. Fingerprint features—statistical analysis and system performance estimates. Proc IEEE 1997;85(9):1390–421.
8. Krzysztof KM, Morier P, Drygajlo A. Study of the distinctiveness of level 2 and level 3 features in fragmentary fingerprint comparison. Proceedings of the ECCV 2004 International Workshop on Biometric Authentication, Lect Notes Comput Sci; 2004 May 15; Prague, Czech Republic. Springer-Verlag 2004;3087:124–33.
9. Jain AK, Chen Y, Demirkus M. Pores and ridges: high resolution fingerprint matching using level 3 features. IEEE Trans Pattern Anal Mach Intell 2007;29(1):15–27.
10. Chen Y, Jain AK. Dots and incipients: extended features for partial fingerprint matching. The Biometric Consortium Conference (BCC); 2007 Sept 11–13; Baltimore, MD. IEEE Xplore, Biometrics Symposium 2007;1–6, doi 10.1109/BCC.2007.4430538.
11. Gupta A, Buckley KA, Sutton R. The effect of substrate on the reproducibility of inked fingerprint pore dimensions examined using photomicrography. Fingerprint Whorld 2007;33(128):156–63.
12. Gupta A, Buckley KA, Sutton R. Latent fingermark pore area reproducibility. Forensic Sci Int 2008;179:172–5.
13. Zhao Q, Zhang L, Zhang D, Luo N. Adaptive pore model for fingerprint pore Extraction. IEEE 2008. Available at: http://figment.cse.usf.edu/~sfefilat/data/papers/TuAT6.3.pdf (accessed May 5, 2009).

14. International Biometric Group. Quantitative research on friction ridge patterns analysis of level 3 features at high resolutions phase II—final report 2008. Available at: http://www.biometricgroup.com/reports/public/NIJ_QRFRP_2007-DN-BX-K239_Final_Report.pdf (accessed April 27, 2009).

15. Ratha NK, Bolle RM. Effect of controlled acquisition on fingerprint matching. Proceedings of the 14th International Conference on Pattern Recognition. Vol. 2; 1998 Aug 16-20; Brisbane, Australia. Washington, DC. IEEE Comput Soc 1998;2:1659–61.

16. Watson C, Grother P, Cassasent D. Distortion-tolerant filter for elastic-distorted fingerprint matching. Optical Pattern Recognition XI Conference; 2000 April 26; Orlando, FL. Proc Soc Photo Opt Instrum Eng 2003; doi:10.1117/12.381591.

17. Ross A, Dass SC, Jain AK. Fingerprint warping using ridge curve correspondences. IEEE Trans Pattern Anal Mach Intell 2006;28(1):19–30.

18. Police Information Technology Organisation. Identification roadmap 2005–2020. United Kingdom: Police Information Technology Organisation, 2005.

19. NIST fingerprint data interchange workshop. Latent print discussion panel. 1998. Available at: http://www.itl.nist.gov/iad/894.03/fing/summary.html (accessed November 8, 2006).

20. Ray M, Meenen P, Adhami R. A novel approach to fingerprint pore extraction. Proceedings of the 37th Southeastern Symposium on System Theory, 2005;282–6.

Additional information and reprint requests:
Abhishek Gupta, M.Sc., M.Phil
School of Applied Sciences
University of Wolverhampton
Wolverhampton WV1 1LY
U.K.
E-mail: a.gupta@wlv.ac.uk

PAPER

ODONTOLOGY

Mary A. Bush,[1] *D.D.S.; Howard I. Cooper,*[2] *D.D.S.; and Robert B. J. Dorion,*[3] *D.D.S.*

Inquiry into the Scientific Basis for Bitemark Profiling and Arbitrary Distortion Compensation

ABSTRACT: Prediction of dental characteristics from a bitemark (bitemark profiling) and arbitrary photographic distortion compensation are two practices proposed in bitemark analysis. Recent research on the effect of inherent skin tension properties in bitemark analysis suggests that these practices are subject to review. A biting apparatus was used to create 66 bitemarks in human cadaver skin. The bitemarks were photographed, sized 1:1, and evaluated with Adobe Photoshop®. Metric/angular measurements and hollow volume dental overlays were employed. Distortion produced was calculated and assessed. Results showed distortional ranges were nonuniform both between bites, as well as within each bite. Thus, enlarging/decreasing the photograph uniformly would not correct the distortion that resulted. With regard to bitemark profiling, 38% of the bites created patterns that could be misleading if profiled. Features were present/absent that were inconsistent with the biter's dentition. Conclusions indicate bitemark profiling and arbitrary distortion compensation may be inadvisable.

KEYWORDS: forensic science, forensic odontology, bitemarks, bitemark research, anisotropy, skin tension

It is well known that distortion can occur in a bitemark (1). What may not be recognized is the limitation that this distortion can place on two potential tenets of bitemark analysis. The assumption that a bitemark photograph can be arbitrarily altered to account for the tissue distortion is one theory (2,3). In this situation, the 1:1 bitemark photographed may be enlarged/decreased by an arbitrary amount, while the overlay remains the same size. A second supposition is that a profile can be generated from a bitemark in an attempt to anticipate the dental configuration of a biter (1,4–6). Research on the properties of skin, however, has suggested that these may not be prudent practices in bitemark analysis (7–9).

A basic knowledge of human skin and its biomechanical properties is important in performing bitemark analysis. One of the properties of skin responsible for distortion is anisotropy, meaning that skin possesses different properties in different directions (10–12). Thus, in a bitemark, the transferred dental pattern can be distorted unequally in one direction, or another, because of the inherent pretension that exists in skin. Anisotropy itself can dictate the overall resultant configuration of a bitemark (7,9). Previous studies have shown that the degree of tightness of skin is one of the major variables that determine distortion (7,9).

Skin exists in a state of pre-tension (13–15). The direction of pre-tension is best described by skin tension lines, commonly referred to as Langer Lines after Karl Langer (16–18), one of the first to describe this property. Pre-tension is dictated by the mechanical demands of each part of the body, which will vary between individuals, and also differ with positional change (standing, lying, flexion, extension, rotation, supination, etc.) (17). Site-to-site variation is governed in part by movement, underlying tissue type, and joint articulation. Simple pinch tests can be used to indicate the pre-tension direction as the skin is easier to pinch perpendicular to tension lines and more difficult in the parallel direction (19).

Skin also exhibits a nonlinear response to stress that can be described in stress–strain curves (10,20). These curves illustrate the visco-elastic properties of skin. The *y*-axis represents stress, expressed in Megapascal units (MPa). The *x*-axis is strain (percent elongation). The initial portion of the curve depicts the rapid extension of skin under low stress (elastic phase). Under low stresses, the skin exhibits an elastic response to stress; however, as the stress increases, the skin becomes more viscous (stiffer) and elongation becomes limited. The curve thus begins to turn upward in the viscous stage, becoming almost linear in the final portion, illustrating the small amount of elongation possible in this phase.

Pre-tension in the skin dictates the curve's location on the *x/y* axis. The curve will shift to the left in tighter tissue types (7). As the curve shifts to the left, skin elongates less in the elastic phase given the same application of stress (20). Thus, it becomes stiffer more quickly (Fig. 1). Tissue type and tension direction can cause the bitemark to distort unevenly (7,21). Therefore, it would be inappropriate to apply a uniform enlargement or reduction to a bitemark photograph to create a better "fit" for a suspect dentition.

Also, because skin deforms, a bitemark may mimic a dentition other than the perpetrator's. This is particularly significant in cases dealing with an open population, children within the same family, or in which the dentition is either naturally well aligned or has received orthodontic treatment (1). In a study by Miller et al. (8) it was found that a number of dental casts "fit" the bitemarks better than the perpetrator's dentition. In some cases, gaps were noted between the individual tooth indentations in the bitemarks where there were no missing teeth or diastema in the perpetrator dentition.
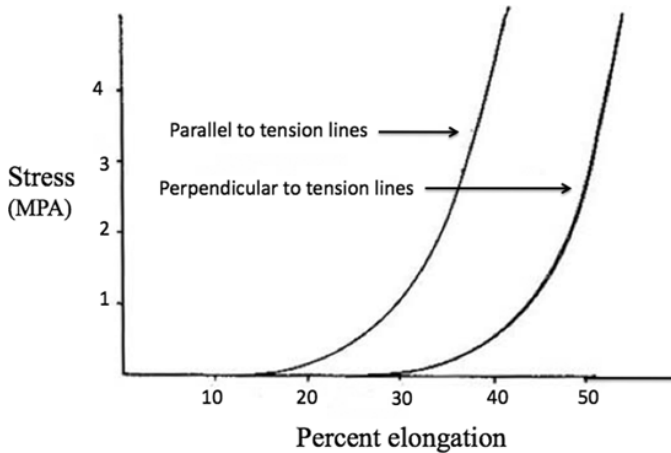
FIG. 1—*Stress–strain curves for skin parallel and perpendicular to skin tension lines. There is less elongation possible for skin parallel to tension lines.*

Features were noted in bitemarks such as lingual tooth displacement, tooth rotation, and arch flattening that were not present in the perpetrating dentition, but which could readily be found in an open population (8). In another study, several instances were found in which a missing tooth appeared to create an indentation (9). The discrepancies seen in these studies were mainly related to skin properties.

Consequently, unless the examiner is well acquainted with these potential problems, creating a biter profile may result in a significant misdirection for perpetrator identity. The distortion produced by the variables of skin may be of such magnitude that prediction of characteristics of the dentition from a bitemark would be imprudent.

The goals of this study, therefore, were to investigate the effect of arbitrary alteration of bitemark photographs and to determine limitations of profiling a biter from a bitemark photograph.

Materials and Methods

Human Subject Institutional Review Board (HSIRB) exemption was granted for all phases of this project. A total of 122 bites were created on 11 human cadavers. The cadavers were stored at 4°C and allowed to warm to room temperature prior to bite creation. The bites were inflicted during experimentation for several previous projects that were performed over a 2-year period (7–9). Bites were created with the dentition located directly above the skin. No bites were created with an angled approach. Sixty-six of these experimentally created bites were chosen for this project. The criteria for exclusion were bitemarks created to investigate issues such as postural distortion and laceration, as the distortion in these bites would have been more extreme.

TABLE 1—*Distortion pattern seen, and number of bites that depicted these changes.*

Type of Distortion Pattern	Number
Flattening of the arch	6
Constriction of the arch	5
Significant deviation in overall alignment	4
Missing tooth appears in the bite	3
Significant rotation of teeth	2
Significant buccal/lingual displacement of teeth	2
Questionable orientation of bitemark	2
Diastema appears when no diastema is present in biter's dentition	1

Polyvinylsiloxane (PVS) impressions were randomly collected from the patient population pool of the State University of New York at Buffalo School of Dental Medicine. The patient pool in the UB clinic represents a varied population of all ages, gender, race, and socio-economic status. PVS impressions were also taken of a single volunteer. Models were created from the PVS impressions.

Each set of models was scanned on a flatbed scanner at 300 dpi (Hewlett Packard 6100/CT), and using Adobe Photoshop®, the images were sized 1:1. Hollow volume overlays were produced and metric/angular analysis was performed on each model. The mesial to distal width of each tooth, the intercanine distance, and the relative angle between teeth was determined. The models were attached to a vise grip by a screw mechanism. The bites were then created with the hand-held vise grip.

The maximum anterior bite force capable of the vise grip was tested with a bite force transducer and found to be well within the range of maximum anterior human biting capacity. This range was
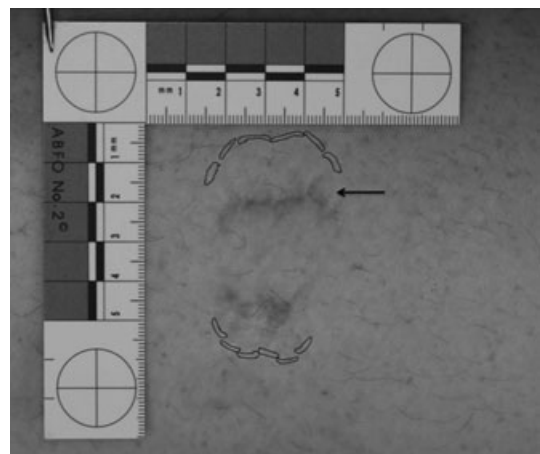

FIG. 2—*Both maxillary and mandibular impressions are far more constricted than the overlay of the biter. #7 (right lateral incisor) is labially positioned in the bite (arrow).*
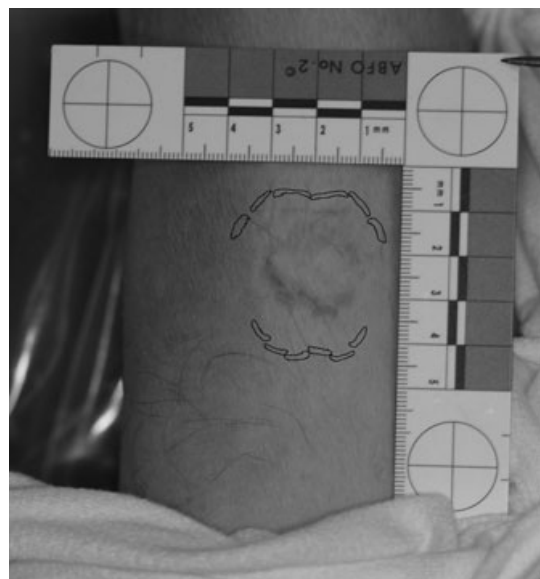

FIG. 3—*Notice significant constriction of the arch form in the bite compared to the dentition that created it.*

established by a volunteer's *in vivo* test biting on the bite force transducer giving an average of 190N. This range was also consistent with studies of mean maximum anterior bite force (22,23).

Photography was performed with a Canon Rebel XTi 10.1 Mp digital camera. An ABFO #2 scale was in place for each photograph. The maxillary and mandibular bitemarks were photographed separately, as needed, to minimize photographic distortion because of curvature in the surface of the bitten area. Using Adobe Photoshop®, each photograph was sized 1:1. For bitemark profiling, the photographs were analyzed and hollow volume overlays of the biter were created and compared to the resultant bitemark. Any bite pattern that had a deviation great enough from the dentition of the biter that could be misleading for an investigator was included in this study.

For arbitrary distortion compensation, three sets of three bitemarks (each set produced on the same body part) were created with the same dentition, and metric and angular measurements were made to calculate the distortion that resulted. The deviations for

angle between teeth, mesial to distal length, and intercanine diameter for the six anterior maxillary and mandibular teeth for each bite were tabulated. Hollow volume overlay comparison was also performed.

The experimental intra-observer measurement error was ±0.2 mm for the intercanine and mesial to distal distances, and ±2° for the rotational angle difference.

Results

Although some bite patterns reflected the biter's dental arrangement, in many instances, the bite pattern, if profiled, would misdirect an investigator to a person that had features not present in the perpetrator's dentition. Table 1 describes the type of distortion that resulted and the number of bites affected. As these bites were created for two previous studies and the deviations reported in those studies, metric/angular calculations were referenced (7,8). Of the 66 bites, 25 (38%) showed a change that could be misleading if profiled.

Figures 2–12 illustrate the changes that occurred. In Fig. 2, the mandibular dentition as seen in the bitemark appears far more
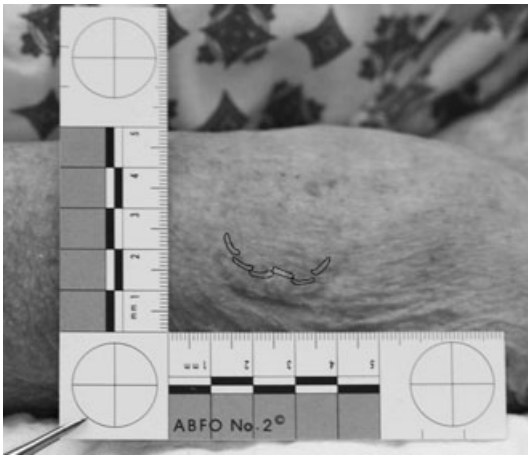
FIG. 4—*The bitemark appears to represent a relatively straight dentition. The overlay of the dentition that created it is above the bite. Note the lingual displacement of #25 in the overlay. This is not depicted in the bitemark.*
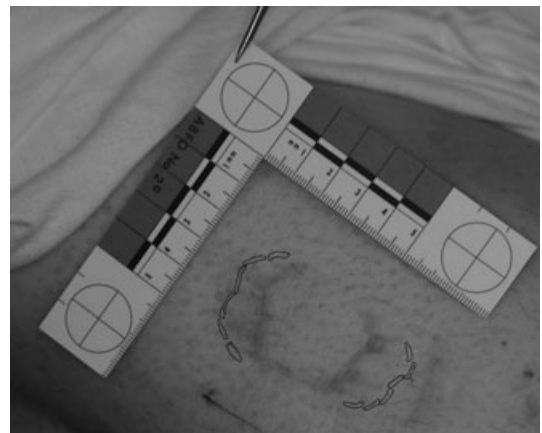

FIG. 6—*Notice discrepancy between the alignment pattern of the lower dentition in the bitemark compared to the dentition that created it.*
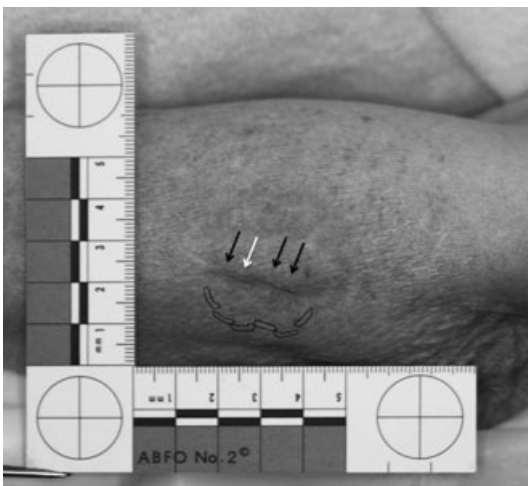

FIG. 5—*Again, there is significant straightening of the arch form. However, in this bite, #24 appears to be positioned slightly labially (white arrow), while #23, 25, and 26 are slightly lingual (black arrows). Also, #23 appears to have a disto-facial rotation (left black arrow). This is not the alignment pattern of the biter.*
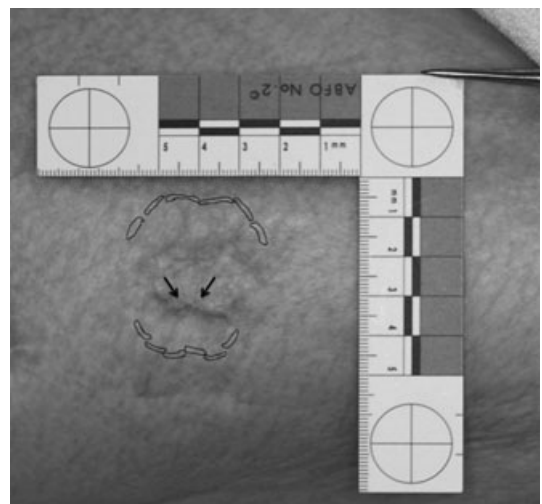

FIG. 7—*The mandibular incisors appear to have a mesio-lingual rotation. This feature is not present in the biting dentition. Also notice arch form constriction.*

constricted and mal-aligned than the dentition that created it. In this example, the maxillary dentition in the bitemark suggests a pattern in which #7 (right lateral incisor) is labially positioned. Bitemark arch width is constricted for the maxillary dentition as well. Figure 3 also illustrates a bitemark in which the arch shape is far more constricted than the biter's. Figures 4 and 5 show a flat, relatively straight bitemark impression. The dentition that created this bite is slightly mal-aligned with a lingually displaced #25. This dental feature is not present in the bitemark. The mandibular impression in Fig. 6 suggests a different alignment pattern than that of the biter. In Fig. 7, it appears as if the mandibular incisors may have a mesio-lingual rotation. The dentition that created this bite does not possess this feature. Figure 8 illustrates a significant deviation of the bitemark from the perpetrator's dentition. For Fig. 9, it appears as if #25 (lower right central incisor), a missing tooth, left an indentation. Figure 10a,b show two orientations of the same bitemark. This bite was created with three missing anterior teeth for the maxillary and mandibular dentition (upper and lower central
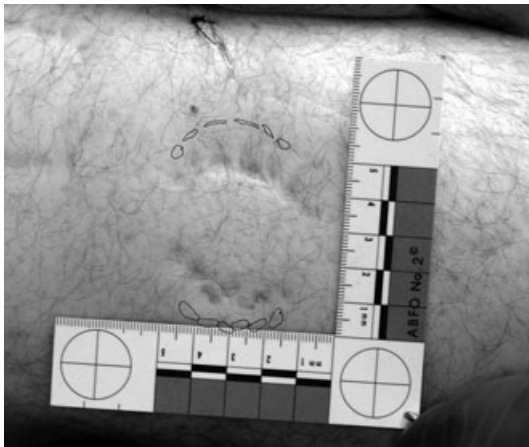
FIG. 8—*Notice the difference between the bitemark and the overlay of the dentition that created it. There is a significant overall discrepancy in the mandibular alignment pattern of the bite compared to the dentition that created it.*
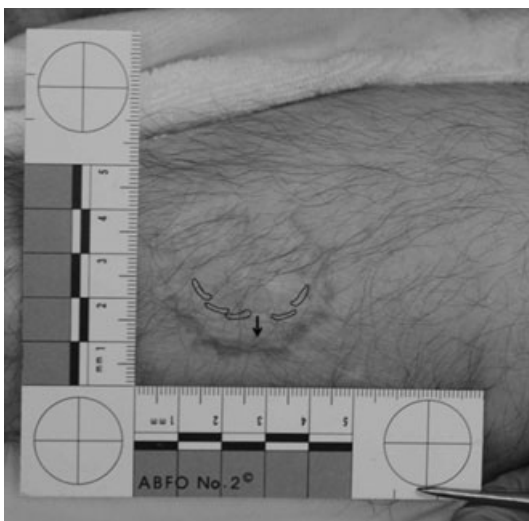
FIG. 9—*It appears as if there are no missing indentations in the bitemark. This bite was created with a dentition in which #25 (lower right central incisor) was missing (arrow).*
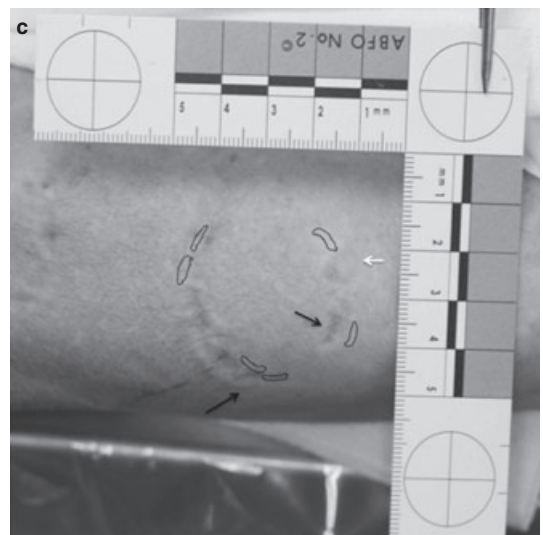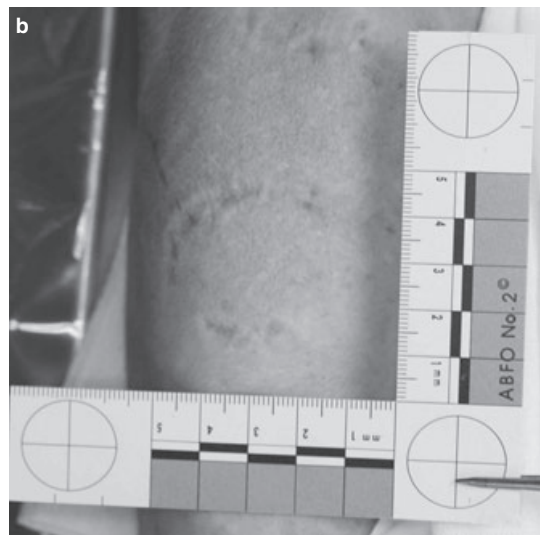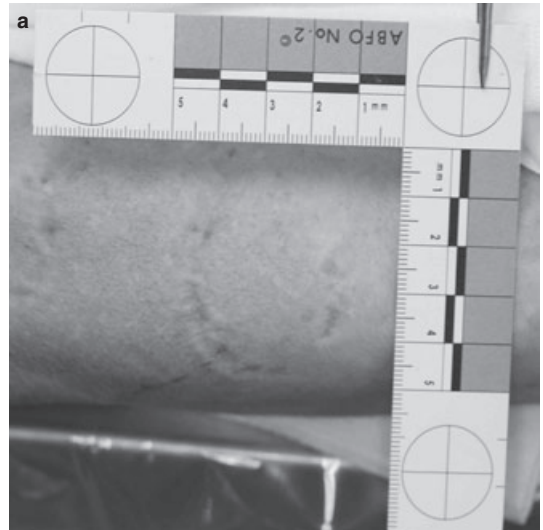

FIG. 10—*(a) The orientation of the bite may be questionable. This is the correct orientation of the bitemark. (b) Figure 10a rotated 90°. This figure suggests a more typical wider maxillary arch form. This is the incorrect orientation. (c) Overlay of the biter on the bitemark. Three teeth were missing on the maxillary and mandibular dentitions. Note the discrepancy of the upper right canine (white arrow). Also note the discrepancy of the lower overlay (black arrows).*
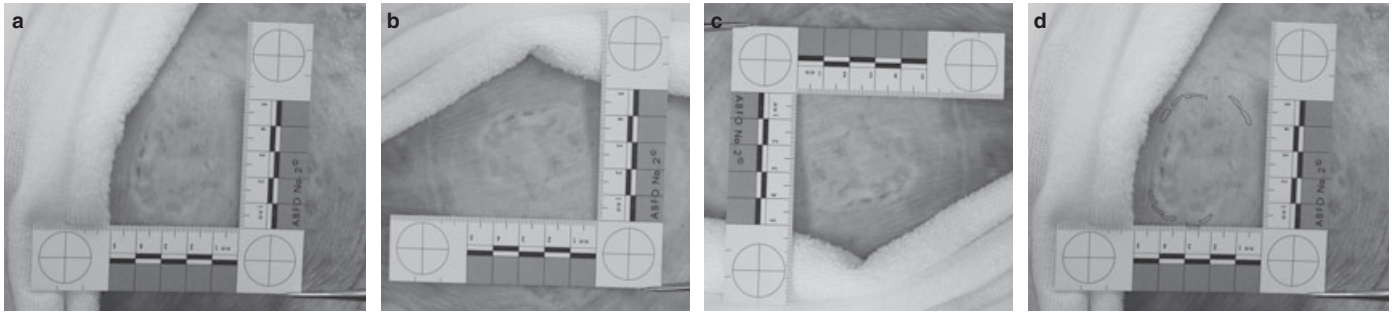
FIG. 11—*(a) Again, the orientation of this bitemark may be questionable. This is the correct orientation. (b) Rotation of Fig. 11a by 90° clockwise. The dentition appears to have a more typical wider maxillary arch form, however, these indentations appear to be small for maxillary teeth. (c) Rotation of Fig. 11a by 90° counter-clockwise. The dentition appears to have a more characteristic arch form and the small size of the indentations may be more typical of the mandibular teeth. (d) Overlay of the biting dentition placed around the bite in the correct orientation. Both maxillary and mandibular dentitions are missing one tooth.*
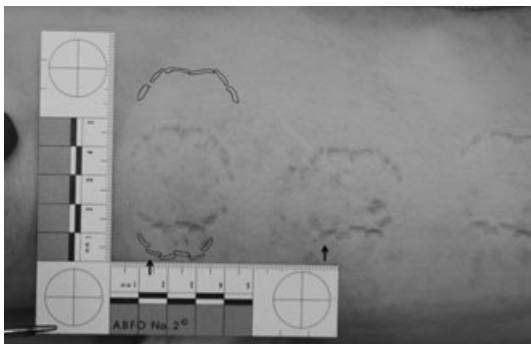
FIG. 12—*The overlay of the biter is placed around the bite on the left. This dentition created all three bites shown (Fig. 13 depicts the bite on the right completely). Notice variations in arch shape and form. Also note the change in orientation of tooth #23, in the middle bite, from the dentition that created it (arrows).*

TABLE 2—*Angle deviation between teeth in degrees for each of three series of bites created.*

| | Angle Deviation Between Teeth in Degrees | | | | | | | | |
| | Cadaver 1 | | | Cadaver 2 | | | Cadaver 3 | | |
Teeth	Bite 1	Bite 2	Bite 3	Bite 1	Bite 2	Bite 3	Bite 1	Bite 2	Bite 3
6–7	+0.5	+8.2	−1.6	+9.3	+0.7	−5.8	X	+10.4	+4.2
7–8	−11.8	−15.8	−2.7	−10.1	−9.5	−6	−29.4	−21.3	−11.2
8–9	+8.7	+13.3	+7.6	+8.6	+21.6	−5.7	+14.1	+10.2	+5.7
9–10	−2.9	−16.5	−0.5	−13.1	−7.3	−5.2	−7.8	−20.2	−13.2
10–11	−9.6	+13.1	+7.2	+12.6	+0.1	+9	−0.3	X	X
22–23	−58.5	−56.4	−54.8	−54	+79.4	−67	−56.2	−70.4	−46.2
23–24	+19.1	+19.5	+25	+27.8	+30.5	+25.5	+25.1	+30.1	+8.2
24–25	+4.8	−1.3	−3.1	−2.9	−4.1	−5.6	−6.5	−9.7	+5.1
25–26	+7.6	+12.3	+13	+23.9	+17.1	+15.9	+13.9	+18.7	+19.6
26–27	−8.2	+6.7	+8.8	+7.3	+5.1	+19.2	+13.2	+3.8	−4.9

Note discrepancies not only between bites, but also between teeth within the same bite. X denotes an area that was not clear enough for measurement purposes.

incisors, upper left lateral incisor, and lower left lateral incisor). Bitemark orientation is clearly an issue as Fig. 10*b* suggests a more typical wider maxillary arch form. Figure 10*c* exhibits a north-south maxillary to mandibular bitemark orientation and the corresponding biter's overlay. The maxillary dentition is larger in intercanine width, and the mandibular smaller, than the biter's dental overlay. The question of correct orientation of the bite is further illustrated in Fig. 11*a–d*.

Enlarging or reducing a bitemark photograph made certain parameters "fit" better, but others worse. This is illustrated by bitemarks created adjacent to each other on the same limb. All three sets of bitemarks followed the pattern of inconsistent increases and decreases. Tables 2–4 describe the deviation in angles of rotation between teeth, mesial to distal length, and intercanine diameter between the bitemark and the biter's dentition. Note that the discrepancies are not only significant between bites, but also within a single bite.

Figure 12 illustrates three bites on the top of the thigh. The bite on the left depicts a maxillary intercanine width *decrease* of 0.7 mm with a mandibular *increase* of 0.8 mm compared with the dentition that created it.

Conversely, the maxillary dentition intercanine width in the middle bite in Fig. 12 *increased* by 0.5 mm while the mandibular also *increased*, but by much more, 2.5 mm. Comparing the two bitemarks, the angulation for #10 (upper left lateral incisor) in the left bite is not nearly as steep as for the middle bitemark. For the middle bite, the angulation for #23 is drastically different than the

dentition that created the bite. It now appears to be perpendicular to tooth #22 (lower left canine). The third bite in this series, shown in Fig. 13, depicts a maxillary intercanine width 3.6 mm *larger* and a mandibular width 2.9 mm *larger* than the dentition that created them. All three of these bites show a difference in overall arch shape from the dentition that caused them.

In Fig. 14, the maxillary overlay is slightly larger for the intercanine width but the mandibular overlay is smaller. In Fig. 15, the original bitemark photograph was enlarged by 10% while the dental overlay remained unchanged. While maxillary correlation improved when comparing bitemark to overlay, the mandibular correlation worsened. Uniform increase/decrease in resizing bitemark photographs while the dental overlay remained unchanged resulted in disproportionate differences. In this experiment, the causative dentition was known and represented a gold standard. The results refute the use of the technique of uniformly increasing/decreasing the size of only one of the two elements (bitemark photograph or dental overlay).

When comparing a perpetrator dentition to a more severely distorted bitemark, additional problems can arise. In Fig. 16, the biter's dental overlay is immediately above the bitemark. The bitemark pattern produced resulted in a much flatter and straighter appearance than the biter's dentition (this is the same bite from Fig. 4). A second, arbitrarily chosen dental overlay with a relatively straight alignment is placed above the biter's overlay. The straighter

TABLE 3—*Deviation in mesial to distal width for each of three series of bites created.*

| | Deviation in Mesial to Distal Width | | | | | | | | |
| | Cadaver 1 | | | Cadaver 2 | | | Cadaver 3 | | |
Tooth	Bite 1, mm (%)	Bite 2, mm (%)	Bite 3, mm (%)	Bite 1, mm (%)	Bite 2, mm (%)	Bite 3, mm (%)	Bite 1, mm (%)	Bite 2, mm (%)	Bite 3, mm (%)
6	+0.1 (1.6)	+0.3 (4.9)	+0.5 (8.2)	+1.7 (27.9)	−0.4 (6.6)	−0.3 (4.9)	X	−0.5 (8.2)	−0.5 (8.2)
7	−1.2 (18.2)	−1 (15.2)	−0.8 (12.1)	−0.6 (9.1)	−0.6 (9.1)	−1.7 (27.8)	+0.2 (9.1)	−1.2 (18.2)	+0.1 (15.2)
8	−0.5 (6.3)	−0.3 (3.8)	−1.2 (15)	−0.5 (6.3)	−0.1 (1.3)	−0.2 (2.5)	−1.2 (15)	−0.6 (7.5)	−1 (12.5)
9	−0.3 (3.7)	−1.2 (14.6)	−2.4 (29.3)	−0.8 (9.8)	−0.3 (3.7)	−0.8 (9.8)	−1.3 (23.2)	−1.5 (18.3)	−2.2 (26.8)
10	+0.8 (13.6)	0	−1.2 (20.3)	−0.1 (1.7)	+0.4 (6.8)	0	+0.4 (6.8)	+0.4 (6.8)	−0.9 (14.5)
11	−0.3 (4.8)	0	−1.1 (17.7)	+0.3 (4.8)	−0.7 (11.3)	+0.2 (3.2)	−0.3 (4.8)	X	X
22	−0.7 (10.6)	−0.8 (14.3)	−0.6 (10.7)	−0.5 (8.9)	−0.4 (7.1)	−0.5 (8.9)	+0.3 (5.4)	+0.5 (8.9)	−0.2 (3.6)
23	+0.9 (17.3)	+0.7 (13.5)	+0.3 (5.8)	−1 (18.5)	+0.8 (15.4)	+1.6 (30.1)	+0.5 (9.6)	+0.3 (5.8)	−0.8 (15.4)
24	−1.3 (16.3)	−0.7 (13)	−0.7 (13)	−0.9 (16.7)	0	0	−0.5 (9.3)	−0.7 (13)	−1.6 (29.6)
25	0	−0.3 (5.8)	−0.6 (11.5)	−0.2 (3.8)	−0.2 (3.8)	−0.3 (5.8)	−0.4 (7.7)	+0.1 (19.2)	+0.5 (9.6)
26	+0.8 (16.3)	+1 (20.4)	+0.3 (6.1)	+0.6 (12.2)	+1 (20.4)	+1.3 (26.5)	+0.6 (12.2)	+0.5 (10.2)	+0.5 (10.2)
27	−0.9 (17)	+0.1 (18.9)	−1.4 (26.4)	−0.5 (9.4)	−0.3 (5.7)	0	+1.4 (26.4)	+0.8 (15.1)	+0.2 (3.8)

Note discrepancies not only between bites, but also between teeth within the same bite. X denotes an area that was not clear enough for measurement purposes.

TABLE 4—*Intercanine width deviation for each of three series of bites created.*

| | Intercanine Deviation | | | | | | | | |
| Teeth | Cadaver 1 | | | Cadaver 2 | | | Cadaver 3 | | |
Arch	Bite 1, mm (%)	Bite 2, mm (%)	Bite 3, mm (%)	Bite 1, mm (%)	Bite 2, mm (%)	Bite 3, mm (%)	Bite 1, mm (%)	Bite 2, mm (%)	Bite 3, mm (%)
Maxillary	+1.8 (5.4)	+1.5 (4.5)	−2.8 (8.4)	−0.7 (2.1)	+0.5 (1.5)	+3.6 (10.8)	X	X	X
Mandibular	+2.4 (10)	+2.1 (2.1)	−1 (4.2)	+0.8 (3.3)	+2.5 (10.5)	+2.9 (12.1)	+2.5 (10.5)	+3.5 (14.6)	+2 (8.4)

Note discrepancies not only between bites, but also between teeth within the same bite. X denotes an area that was not clear enough for measurement purposes.
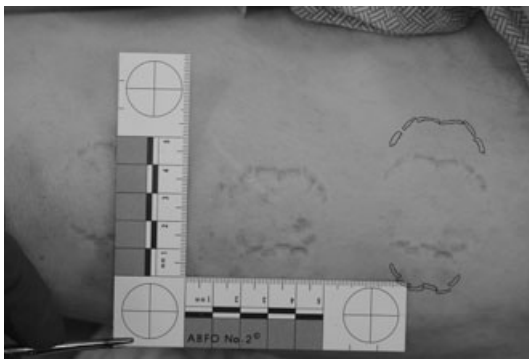
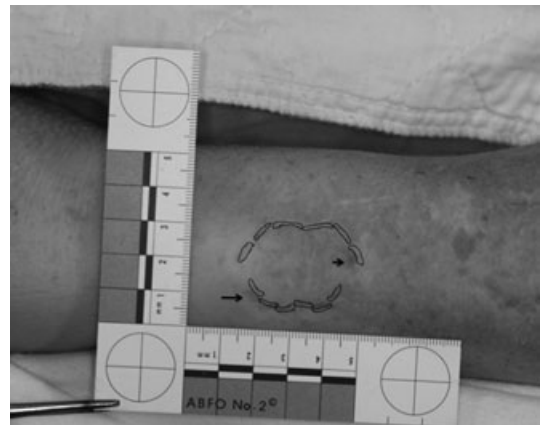FIG. 13—*The overlay of the biter is placed around the third bite in the series.*

FIG. 14—*Notice discrepancies between the bitemark and the overlay of the biter (arrows).*

overlay does not, however, correlate well with the bitemark (Fig. 17). Increasing the bitemark photograph by 10% greatly improves the correlation with the straighter exemplar dentition (Fig. 18). The obvious concern with this technique is that this dentition did not create the bitemark. Use of this method can lead to exclusion of the perpetrator, or worse, erroneous biter identification.

Discussion

Skin pre-tension does not have a uniform distribution in a human body. Tension not only varies from person to person but also varies at a single site on the same individual. Tension is always greater parallel to tension lines and more relaxed perpendicular to them, resulting in anisotropy in skin. Therefore, the degree of distortion will not be uniform throughout a bitemark. There may be intra-arch as well as interarch distortion. The magnitude of these distortional changes can also vary considerably both within and between each arch. To the authors' knowledge, there have been no scientific studies that give direction as to whether an image of an overlay or bitemark should be increased or decreased and if so, to what extent.
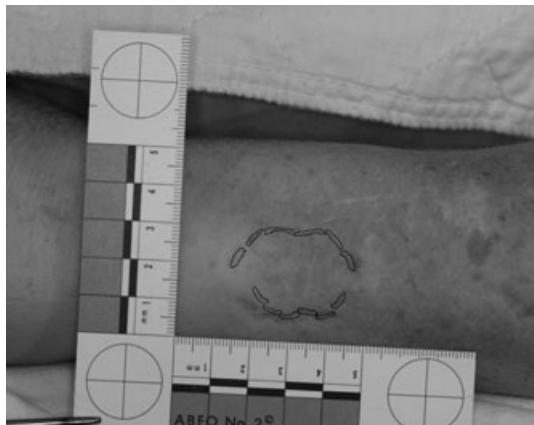
FIG. 15—*The photograph has been enlarged by 10%, while the overlay remained the same size. Note the better correlation with the maxillary dentition but the worsening of the lower.*
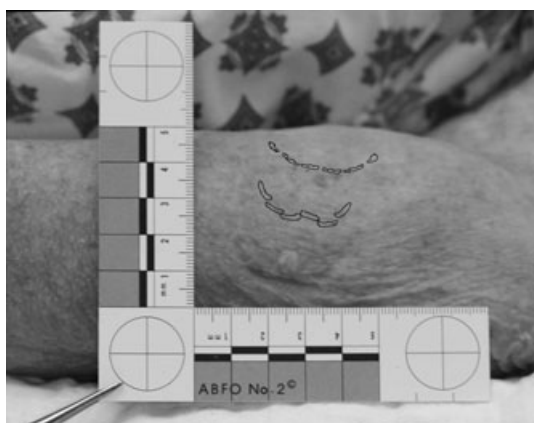

FIG. 17—*Notice poor correlation of the straight overlay (arrows).*

FIG. 16—*Same bitemark as seen in Fig. 4. The overlay of the biter is immediately above the bitemark. An arbitrarily chosen overlay of a relatively straight dentition is placed above the biter's overlay.*

FIG. 18—*This photograph has been enlarged by 10% but the overlays are still the same size. Note improvement of the correlation of the straight overlay.*

The variables associated with the skin are too complex to allow prediction of a single distortion factor.

This study indicated that arbitrary distortion of a bitemark photograph to "match" a dental overlay in an attempt to compensate for tissue distortion is not an appropriate technique. The anisotropic nature of human skin cannot at this time be precisely anticipated to arrive at a percentage enlargement or reduction of an image in any given direction. Results showed distortional ranges were nonuniform both between bites, as well as within each bite. Thus, enlarging/decreasing the photograph uniformly would not correct the distortion that resulted.

The results of this study indicate that unless a closed scenario exists, it may not be prudent to profile a biter from a bitemark. Marked deviation from the biters' dentition occurred in 38% of the postmortem bites on adult cadaver skin because of distortion. There are two potential perils here for the forensic practitioner. First, distortion effects could lead to dental profiling an innocent person. Second, arbitrary interpretation of distortion could be used to explain discrepancies in a bitemark in order to include a presupposed suspect. This potential bias could steer an investigation into ways that might exclude entire populations or, worse, could lead to the arrest and conviction of an innocent person.

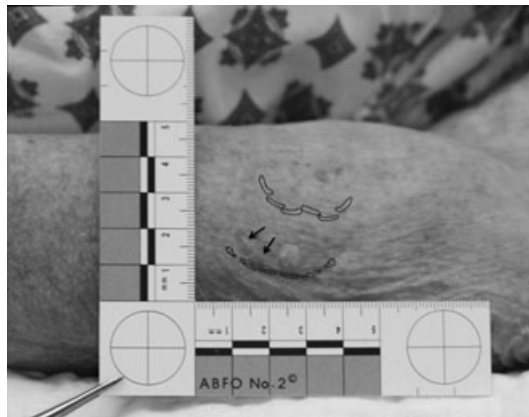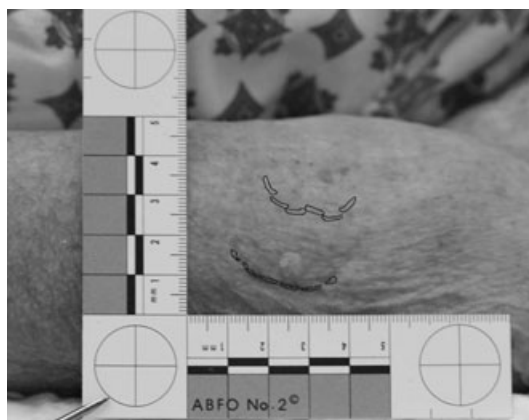It is important to note that the bitemarks shown in this current project did not involve any postural distortion. Including the effects of postural change may worsen distortion, creating a further deviation from the causative dentition (7).

It is acknowledged that experimentation occurred on cadaver skin and that results may differ on live tissue. However, cadaver models have been used in many fields to test biomechanical properties of the skin, and are accepted techniques. The use of cadavers was seen as an advantage as only clear indentations were studied as opposed to a diffuse bruise. In a complex field of analysis such as this, it is important to be able to control as many variables as possible. The use of cadavers can be regarded as providing the basic groundwork for scientific understanding of bitemark distortion. Restrictions imposed by human subject review boards may mean that there are aspects of bitemark analysis that may never be studied in living human skin, such as dynamic interactions and bruising.

Experimentally created bites allowed for a gold standard situation, as the biter was known. This permits investigation into skin properties, how a bite can be distorted, and comparison to other potential biting dentitions.

As in the pattern recognition process of fingerprint examination, it cannot be predetermined which features might be present either in the impressing object or the imprint. Thus, one cannot say *a priori* that features may be compared between the dentition and bitemark (24). It is during the comparison of bitemark and suspect's dentition that a prudent examiner identifies which characteristics are common and are clear enough to be recognized. A feature that

was noted during examination of the dentition may not be accurately represented in skin or vice versa. Stipulating in advance what may or may not be present can lead to possible bias, inaccurate scenario account, or dental perpetrator misdirection.

As stated in an earlier study, every occasion in which a dentition comes in contact with skin can be considered a unique event (7). The data derived showed no correlation and was not reproducible, that is, the same dentition could not create a measurable impression that was consistent in all of the parameters in any of the test circumstances.

There may be compelling evidence associated with a bitemark, including the presence of DNA, crime scene context, corroboration of victim accounts, timing of injury/death, exclusion, perpetrator identification and other factors, which will continue to make bitemark evidence important in court. However, the authors of this article urge caution in presumptive unilateral alteration of photographic bitemark evidence to "fit" a suspect dentition as well as definitive dental profiling based upon the bitemark.

The goal of this research is to continue to establish the basis of scientific and objective bitemark analysis, thereby minimizing to the greatest extent possible the likelihood that forensic science will lead to an innocent person's wrongful conviction, and worse yet, execution based on erroneous bitemark interpretation.

References

1. Dorion RBJ, editor. Bitemark evidence. New York, NY: Marcel Dekker (CRC Press), 2005.
2. State v. Frimpong, 1218308 (California 2008) (en banc).
3. Aws G. Analysis of symmetry in the anterior human dentition and its application in the evaluation and correction of postural distortion in the photographic recording of human bite marks [dissertation]. Edinburgh, UK: Univ of Edinburgh. College of Medicine & Veterinary Medicine, 2006.
4. Johnson LT, Wirtz TS, Radmer TW. The verdict is in: can dental characteristics be quantified? Proceedings of the 60th Annual Meeting of the American Academy of Forensic Sciences; 2008 Feb 17–23; Washington, DC. Colorado Springs, CO: American Academy of Forensic Sciences, 2008.
5. Johnson LT. Quantification of the individual characteristics of the human dentition, Final Report MFRC07-F-06 Feb 1, 2006–Oct 30, 2007, http://www.mfrc.ameslab.gov/R-D_Reports/Quantification_of_Characteristics_of_Human_Dentition_Final.pdf (accessed April 2009).
6. Spitz W, Spitz DJ, editors. Medicolegal investigation of death: guidelines for the application of pathology to crime investigation. Springfield, IL: Charles Thomas, 2006.
7. Bush MA, Miller RG, Bush PJ, Dorion RBJ. Biomechanical factors in human dermal bitemarks in a cadaver model. J Forensic Sci 2009;54(1): 167–76.
8. Miller RG, Bush PJ, Dorion RBJ, Bush MA. Uniqueness of the dentition as impressed in human skin: a cadaver model. J Forensic Sci 2009;54(4):909–14.
9. Bush MA, Thorsrud K, Miller RG, Dorion RBJ, Bush PJ. The response of skin to applied stress: investigation of bitemark distortion in a cadaver model. J Forensic Sci 2010;55(1):71–6.
10. Millington PF, Wilkinson R. Skin. Cambridge: Cambridge University Press, 1983.
11. Daly CH. Biomechanical properties of dermis. J Invest Dermatol 1982;79(Suppl 1):17s–20s.
12. Edwards C, Marks R. Evaluation of biomechanical properties of human skin. Clin Dermatol 1995;13(4):375–80.
13. Meijer R, Douven LF, Oomens CW. Characterization of anisotropic and non-linear behaviour of human skin *in vivo*. Comput Methods Biomech Biomed Engin 1999;2(1):13–27.
14. Reihsner R, Menzel EJ. On the orthogonal anisotropy of human skin as a function of anatomical region. Connect Tissue Res 1996;34(2):145–60.
15. Gibson T, Stark H, Evans JH. Directional variation in extensibility of human skin *in vivo*. J Biomech 1969;2(2):201–4.
16. Langer K. On the anatomy and physiology of the skin. I. The cleavability of the cutis. (Translated from Langer, K. 1861). Zur Anatomie und Physiologie der Haut. I. Uber die Spaltbarkeit der Cutis. Sitzungsbericht der Mathematisch-naturwissenschaftlichen Classe der Kaiserlichen Academie der Wissenschaften, 44, 19. Br J Plast Surg 1978;31(1):3–8.
17. Langer K. On the anatomy and physiology of the skin: II. Skin tension (1862). Gibson T, translator. Br J Plast Surg 1978;31(1):8–13.
18. Langer K. On the anatomy and physiology of the skin III (1862). Gibson T, translator. Br J Plast Surg 1978;31(3):185–99.
19. Borges AF. Relaxed skin tension lines. Dermatol Clin 1989;7(1):169–77.
20. Wilkes GL, Brown IA, Wildnauer RH. The biomechanical properties of skin. CRC Crit Rev Bioeng 1973;1(4):453–95.
21. Byard RW, Gehl A, Tsokos M. Skin tension and cleavage lines (Langer's lines) causing distortion of ante- and postmortem wound morphology. Int J Legal Med 2005;119(4):226–30.
22. Paphangkorakit J, Osborn JW. Effects on human maximum bite force of biting on a softer or harder object. Arch Oral Biol 1998;43(11):833–9.
23. Paphangkorakit J, Osborn JW. The effect of pressure on a maximum incisal bite force in man. Arch Oral Biol 1997;42(1):11–7.
24. Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council. Strengthening forensic science in the United States: a path forward. Washington, DC: National Academies Press, 2009.

Additional information and reprint requests:
Mary A. Bush, D.D.S.
Department of Restorative Dentistry
Laboratory for Forensic Odontology Research
School of Dental Medicine, SUNY at Buffalo
B1 Squire Hall, S. Campus
Buffalo, NY 14214
E-mail: bushma@buffalo.edu

PAPER

ODONTOLOGY

Raymond Richmond,[1] *B.Sc., M. Phil., Ph.D., D.C.P., F.H.E.A. and Iain A. Pretty,*[2]
B.D.S. (Hons), M.Sc., Ph.D., M.F.D.S. R.C.S. (Ed.)

Identification of the Edentulous Individual: An Investigation into the Accuracy of Radiographic Identifications*

ABSTRACT: The dental identification of edentulous individuals can be challenging based on the lack of antemortem materials and unique features visible on radiographs. The constant resorption of the alveolar ridges further complicates the process. The purpose of this study was to quantify the error rate and reliability of dental identifications based on a comparison of synthesized antemortem and postmortem radiographs of edentulous individuals. Ten observers examined ten cases on two occasions and reported dichotomous and conclusion level decisions. These were analyzed using Kappa and Receiver Operator Characteristics. The mean area under the curve was 0.75 and the mean sensitivity was 0.57 and specificity was 0.83. These results suggest that dental identifications of edentulous individuals using radiographs alone have a high error rate and should be dual reported. These data add further weight to the argument that all dental prostheses should be labeled.

KEYWORDS: forensic science, radiography, denture marking, postmortem, identification

Dental identification is one of the commonest means used in identifying deceased individuals who cannot be recognized visually, and despite the introduction of DNA analysis, forensic odontologists continue to be called upon by investigative agencies in cases involving severe head and neck trauma, gross decomposition, fire, and other perimortem assaults (1). Postmortem identification requires the investigation of biometric features that are unique to the individual and are capable of withstanding severe perimortem conditions. To this end, dental features still remain one of the most effective modalities for postmortem identification because number, individuality of shape, and restorative condition of a person's teeth may be the only means of determining identity when other features have long since disappeared (2,3). Variations in dental characteristics such as tooth angulation, morphology, and/or degree of restoration commonly provide a satisfactory number of distinguishing characteristics with which to compare with dental records and antemortem radiographs (4).

However, what does the forensic dentist do when faced with a deceased edentulous individual who has lost all or most of the aforementioned characteristics that have proven so valuable? The bone morphology of such individuals provides a potentially reliable source of information in such cases, but time intervals between antemortem radiographs and postmortem examination can be

considerable, particularly given the reduced need for diagnostic views of edentulous patients.

In countries such as Sweden and in 21 out of 50 states of America, this problem has been acknowledged, and in these locations at least, legislation has been introduced for the mandatory marking of complete dentures (5–7). In the rest of the world, however, such practice appears to be undertaken purely on a voluntary basis.

The importance of dental records in the task of identifying found human remains has been discussed extensively, and over many years in the dental literature (8,9). However, the task of identifying found human remains based on dental comparisons of postmortem and antemortem radiographs is "labor-intensive, subjective, and has several drawbacks, including the following: inherently poor image quality, difficulty matching the viewing angles in postmortem radiographs to those taken antemortem, and the fact that the state of the dental remains may entirely preclude the possibility of obtaining certain types of radiographs postmortem" (10,11). Nevertheless, Borrman and Grondahl (4) and (12) reported that identification of edentulous victims may be achieved via the study of bone trabeculations recorded within antemortem and postmortem radiographs. In their study they attempted to evaluate the degree of accuracy in establishing the identity of edentulous individuals via the use of occlusal antemortem and postmortem radiographs of the maxilla. The authors concluded that examiners who are well trained in oral radiology are capable of determining positive identification from just this information alone, although over half of the examiners in the study were unable to correctly establish identity in all 20 cases (4).

The provision of implant-retained complete lower dentures is becoming increasingly common and therefore provides another potential source of evidence for human identification. A study by

Korkchi (13) appears to support this statement, when they investigated the accuracy of seven forensic dentists and seven police officers in determining identity via the use of comparable intraoral radiographs obtained from a series of edentulous implant patients. The authors stated that "the investigation was based on 34 patients, from 26 of whom a matched pair of radiographs was constructed in such a way that one, taken after insertion of an implant supported prosthesis in the anterior part of the mandible, was regarded as the antemortem radiograph, while another picture from a later follow-up examination served as the postmortem x-ray."

A similar set of radiographs to the aforementioned were created from the remaining eight cases in the study; these eight radiographs were also paired but did not match. Surprisingly, Korkchi and his colleagues found that the total number of errors made by the forensic dentists were higher (26) than those made by the police officers (18) and one police officer managed to match all 26 radiographs correctly. Nevertheless, all observers managed to match up at least 12 of the 26 cases correctly. It should be noted that in all cases the observers stated that it was the design of the prosthesis attached to the implants that led to the identification, rather than either the bone's morphological appearance or the implant itself.

The purpose of the current study is to determine, via the use of dental radiographs, the rate of error among a population of forensic odontologists during their attempt to establish the identity of a series of edentulous individuals.

Materials and Methods

Radiographs

Ten sets of radiographs of edentulous subjects were collected from the dental records of the School of Dentistry. Records were selected where two radiographs of the same dental structures were taken with at least a 24-month interval. These two sets of radiographs became the antemortem and postmortem views. The radiographs were then digitized using a high-resolution scanner and loaded onto a compact disc. A total of 10 cases were compiled comprising five negative identifications and five positive identifications. Negative identifications were produced by randomizing the antemortem and postmortem views of five radiographic sets. A copy of the CD was then sent by mail to 10 examiners of varying levels of forensic odontological experience. For this study, the examiners were classified as either experienced or novice.

The participants were asked to view the radiographs and provide two decisions for each case. The first was a dichotomous decision —simply either positive or negative—and the second decision enabled the examiners to provide a level of certainty using the conclusion levels shown in Table 1. The responses from each of the 10 participants were received via email. After a minimum interval

of 3 weeks they were sent a further compact disc containing the same cases; however, in this instance the order of presentation had been altered. On this occasion the examiners were asked simply to record their dichotomous decision, positive or negative identification. The data from each of the 10 participants were again received via email, and the three sets of data from the study were entered into SPSS (SPSS Inc., Chicago, IL) for statistical analysis from which the following variables were assessed:

- overall accuracy,
- intraexaminer agreement,
- interexaminer agreement,
- sensitivity and specificity.

Results

A total of 10 participants of various degrees of experience took part in this study. For simplicity of analysis the participants were split into the following groups: (i) experts and (ii) nonexperts. The mean number of identifications undertaken by those in the experts group was 189.2 (since commencing forensic work) and in the novice group 36.3. The odontologists' responses were anonymized and any identifying comments removed from their responses.

Accuracy

The term "accuracy" for this study may be defined as percentage of correct matches achieved by an examiner. This definition may appear self-evident; however, it fails to cater for possibility of correct responses that may have been recorded purely by chance. Hence, to eliminate this effect, the statistical test known as Kappa was used, which produces a scale that has been interpreted by Landis and Koch (Table 2). The accuracy results together with their Kappa scores for both attempts made by all 10 examiners are presented in Table 3. Table 4 contains the average values for sensitivity and specificity for each of the examiners' two attempts. Sensitivity is a measure of the ability of the examiners to correctly identify positive identifications. Specificity is the proportion of negative identifications correctly labeled. Table 5 contains the data from the Receiver Operator Characteristics (ROC) analysis. These data represent those decisions where a level of certainty was provided. The closer the value is to 1.0, the more accurate the examiner.

Agreement

An analysis of variance was applied to the grouped data in Table 3 (i.e., experts vs. novices), and a statistically significant difference was found between the two groups in terms of accuracy ($p = 0.023$) in that the experts' scores were found to be more accurate than that of the nonexperts. However, according to Table 2 both groups scored only fair to moderate in terms of Kappa scores, thus indicating that many of the identifications appear to have been

TABLE 1—*Conclusion levels available to the participants in the study.*

Level	Description
5. Inconclusive	Insufficient evidence to make any statement
4. Exclusion	Negative identification
3. Possible	Could be, may or may not be
2. Probable	More likely than not
1. Positive	No reasonable or practical possibility that it is someone other than that represented in the antemortem records

TABLE 2—*Guide to the interpretation of Kappa scores. From Landis and Koch, 1977.*

Kappa	Strength of Agreement
0.00–0.01	Poor
0.01–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

TABLE 3—*Number of correct matches between antemortem and postmortem radiographs.*

		Attempt 1		Attempt 2	
Examiner	Experience	No. of Correct Matches	Kappa	No. of Correct Matches	Kappa
1	Novice	5	0	4	−0.2
2	Novice	7	0.4	6	0.4
3	Expert	9	0.8	10	1.0
4	Novice	7	0.4	5	0
5	Expert	7	0.4	7	0.4
6	Expert	5	0	5	0
7	Novice	4	−0.2	5	0
8	Expert	8	0.6	8	0.6
9	Novice	10	1.0	9	0.8
10	Expert	7	0.4	8	0.6
Mean for all groups (SD)		6.9 (±1.85)	0.38 (±0.37)	6.7 (±2.00)	0.36 (±0.39)
Mean for expert group (SD)		7.2 (±1.41)	0.44 (±0.28)	7.4 (±1.81)	0.48 (±0.36)
Mean for novice group (SD)		6.6 (±2.30)	0.32 (±0.46)	5.8 (±1.92)	0.20 (±0.40)

TABLE 4—*Average values for sensitivity and specificity for both attempts made by each examiner.*

		Attempt 1		Attempt 2	
Examiner	Experience	Sensitivity	Specificity	Sensitivity	Specificity
1	Novice	0.2	0.8	0	0.8
2	Novice	0.8	0.8	0	0.8
3	Expert	0.8	1	1	1
4	Novice	0.6	1	0.6	0.6
5	Expert	0.4	1	0.4	1
6	Expert	0.2	0.8	0.2	0.8
7	Novice	0.4	0.4	0.6	0.6
8	Expert	1	0.8	1	0.8
9	Novice	1	1	1	0.8
10	Expert	0.4	1	0.6	1
Average		0.57	0.83	0.57	0.78

TABLE 5—*Average area under the curves for each of the 10 examiners following the Receiver Operator Characteristics analysis.*

				Asymptotic 95% Confidence Interval	
Examiner	Experience Level	Area Under the Curve	Std. Error	Lower Bound	Upper Bound
1	Novice	0.64	0.18	0.28	0.99
2	Novice	0.80	0.16	0.47	1.12
3	Expert	0.88	0.12	0.63	1.12
4	Novice	0.78	0.16	0.45	1.10
5	Expert	0.74	0.16	0.41	1.06
6	Expert	1.00	0.00	1.00	1.00
7	Novice	0.52	0.19	0.13	0.90
8	Expert	0.86	0.13	0.60	1.11
9	Novice	1.00	0.00	1.00	1.00
10	Expert	0.28	0.17	−0.05	0.61

made by chance. The mean intraexaminer agreement (between examination 1 and 2), measured by unweighted Kappa, was 0.74 (±0.13) and the mean intraexaminer agreement (between the examination 2 for each examiner) was 0.68 (±0.22).

Discussion

For each case in this study, a comparison between an antemortem radiograph and a postmortem radiograph was conducted. This procedure is common practice in forensic casework where dental methods are used for confirmatory identification based upon solid presumptive evidence (1,14). However, it is important to note that in such cases involving multiple victims e.g., such as those involved in mass disaster incidents, it is usually necessary to compare a single postmortem radiograph to several antemortem records. In fact, odontologists will use several other sources of evidence at their disposal when completing a dental identification procedure.

Such evidence may include:

- written case history notes,
- working casts,
- photographs,
- laboratory prescriptions,
- the victim's dentures, etc.

It is therefore important to state that the results of this study should be viewed in context i.e., that in no way do they represent the degree of accuracy achieved in a typical dental identification procedure. However, they do appear to indicate the likely degree of accuracy for identifications involving edentulous individuals. It should also be noted that in this case many of the radiographs provided were panoramic views, as these are most commonly taken on edentulous individuals (7). As these radiographic sets were synthesized, the "postmortem" views were also panoramic, making cross-comparison simpler. In forensic practice it can be very difficult to acquire postmortem radiographs and frequently intraoral radiographs are used instead (15–18). This is rarely an issue for dentate patients, but in edentulous cases, where the overall morphology of the alveolar ridge is of interest, this may make identification more complex. Therefore, while accepting that the examiners in this case were not provided with all the materials that one would normally expect to utilize in a dental identification, they were provided with radiographs of potentially higher quality than those usually obtained postmortem.

It is helpful to compare these results to similar studies studying the reliability of dentate identifications. One such study, using a similar methodology to the current work, found that the mean accuracy was 85.5%, with the most experienced examiners achieving a success rate of 91% (19). These data compare favorably with the current study where the average success rate was 69% and that for the expert group similar at 70%. A similar pattern is seen in the ROC data between the studies. This is certainly a reflection of the complexity of these radiographic comparisons compared to those in the previous study (19). The pattern of sensitivity and specificity is also similar, with poorer sensitivity than specificity suggesting that the examiners are more often correct when excluding individuals than when positively identifying them.

The changing pattern of the dental (alveolar) ridge caused by resorption can make identifying characteristics complex to identify (20–24). The absence of restorations, teeth, or other familiar structures further complicates this process. Although the sample was too small to assess statistically, there was a trend for those cases with larger time intervals between radiographic examinations to be more often scored incorrectly.

The complexity of these identifications was also manifest when the cases correctly identified between first and second dichotomous assessments were analyzed. Many of the cases "flipped" between examinations, with correct responses changed to incorrect and *vice versa*. This pattern is often seen in dichotomous decision making where the examiners are uncertain about their judgments and are being "forced" into choosing between a potential false positive or false negative.

The findings of a better accuracy in the expert group versus the novice group reflect findings from an earlier study with similar methodology (25). These data support the view that experience in complex dental identifications is crucial. Given the relatively high error rate among the number of examiners that have taken part in this study, a reasonable recommendation would seem to be that edentulous identifications should always receive a second opinion before an identification is confirmed.

This study also adds further support for calls to make denture labeling mandatory as such labels, if designed to withstand appropriate postmortem assaults, can facilitate dental identifications.

References

1. Pretty IA, Sweet D. A look at forensic dentistry—Part 1: the role of teeth in the determination of human identity. Br Dent J 2001;190(7):359–66.
2. Deadman WJ. The identification of human remains. Can Med Assoc J 1964;91:808–11.
3. Whelton RL. The identification of the deceased. Mil Med 1965;130:665–8.
4. Borrman H, Grondahl HG. Accuracy in establishing identity in edentulous individuals by means of intraoral radiographs. J Forensic Odontostomatol 1992;10(1):1–6.
5. Borrman HI, Rene N. Denture marking—a questionnaire for patients and dentists. J Forensic Odontostomatol 1997;8:3–10.
6. Richmond R, Pretty IA. Contemporary methods of labelling dental prostheses—a review of the literature. J Forensic Sci 2006;51(5):1120–6.
7. Richmond R, Pretty IA. Antemortem records of forensic significance among edentulous individuals. J Forensic Sci 2007;52(2):423–7.
8. Pretty IA. Forensic dentistry: 1. Identification of human remains. Dent Update 2007;34(10):621–2, 624–6, 629–30 passim.
9. Rothwell BR. Principles of dental identification. Dent Clin North Am 2001;45(2):253–70.
10. Tohnak S, Mehnert AJ, Mahoney M, Crozier S. Synthesizing dental radiographs for human identification. J Dent Res 2007;86(11):1057–62.
11. Tohnak S, Mehnert A, Crozier S, Mahoney M. Synthesizing panoramic radiographs by unwrapping dental CT data. Conf Proc IEEE Eng Med Biol Soc 2006;1:3329–32.
12. Bengtsson A, Olsson T, Rene N, Carlsson GE, Dahlbom U, Borrman H. Frequency of edentulism and identification marking of removable dentures in long-term care units. J Oral Rehabil 1996;23(8):520–3.
13. Korkchi M, Lekholm U, Dahlbom U, Borrman H. Accuracy in identification of implant treated patients by use of intraoral radiographs. J Forensic Odontostomatol 1995;13(1):4–8.
14. Sweet D. Why a dentist for identification? Dent Clin North Am 2001;45(2):237–51.
15. Schwartz S, Woolridge ED. The use of panoramic radiographs for comparisons in cases of identification. J Forensic Sci 1977;22(1):145–6.
16. Serman NJ, Nortje CJ. A retrospective pantomographic study of the edentulous patient—the forensic implications. J Dent Assoc S Afr 1982;37(4):237–41.
17. Happonen RP, Laaksonen H, Wallin A, Tammisalo T, Stimson PG. Use of orthopantomographs in forensic identification. Am J Forensic Med Pathol 1991;12(1):59–63.
18. Du Chesne A, Benthaus S, Teige K, Brinkmann B. Post-mortem orthopantomography—an aid in screening for identification purposes. Int J Legal Med 2000;113(2):63–9.
19. 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(6):1325–30.
20. Felton DA. Edentulism and comorbid factors. J Prosthodont 2009;18(2):88–96.
21. Demarosi F, Leghissa GC, Sardella A, Lodi G, Carrassi A. Localised maxillary ridge expansion with simultaneous implant placement: a case series. Br J Oral Maxillofac Surg 2009;47(7):535–40.
22. Cooper LF. The current and future treatment of edentulism. J Prosthodont 2009;18(2):116–22.
23. Blahout RM, Hienz S, Solar P, Matejka MH, Ulm CW. Quantification of bone resorption in the interforaminal region of the atrophic mandible. Int J Oral Maxillofac Implants 2007;22(4):609–15.
24. Stellingsma C, Vissink A, Meijer HJ, Kuiper C, Raghoebar GM. Implantology and the severely resorbed edentulous mandible. Crit Rev Oral Biol Med 2004;15(4):240–8.
25. Soomer H, Lincoln MJ, Ranta H, Penttila A, Leibur E. Dentists' qualifications affect the accuracy of radiographic identification. J Forensic Sci 2003;48(5):1121–6.

Additional information and reprint requests:
Iain A. Pretty, B.D.S., M.Sc., Ph.D., M.F.D.S. R.C.S. (Ed.)
Dental Health Unit
3A Skelton House, Lloyd Street North
Manchester Science Park
Manchester M15 6SH
U.K.
E-mail: iain.pretty@manchester.ac.uk

PAPER

PATHOLOGY AND BIOLOGY

Mark P. V. Begieneman,[1,2,3] *B.Sc.; Frank R. W. van de Goot,*[1,2,3] *M.D.; Jan Fritz,*[1];
Rence Rozendaal,[1] *M.D., Ph.D.; Paul A. J. Krijnen,*[1,3] *M.Sc.; and Hans W. M. Niessen,*[1,3,4] *M.D., Ph.D.*

Validation of Ultrastructural Analysis of Mitochondrial Deposits in Cardiomyocytes as a Method of Detecting Early Acute Myocardial Infarction in Humans*

ABSTRACT: In the present study, ultrastructural analysis of mitochondrial deposits (black dots within mitochondria) as a method for the detection of early acute myocardial infarction (AMI) was evaluated. In 24 patients with AMI and six controls, analysis was performed in the heart of infarcted patients and noninfarcted controls. In the infarction area in lactate dehydrogenase (LDH)-diagnosed AMI, the percentage of positive mitochondria was significantly higher compared to corresponding heart tissue in control patients and compared to noninfarcted areas within these patients. Also in patients with a clinically diagnosed AMI but no LDH decoloration, a significant higher percentage of positive mitochondria was found in the left ventricle compared to controls and noninfarcted areas. In patients with AMI, an increase in mitochondria with deposits was found in the infarction area compared to controls and noninfarcted tissue within the same patient, suggesting that electron microscopical changes in mitochondria can be used for the diagnosis of AMI less than 3 h old.

KEYWORDS: forensic science, forensic pathology, acute myocardial infarction, electron microscopy, mitochondria, autopsy

In routine clinical pathological examination during autopsy, but also in forensic pathology, it is important to identify or rule out acute myocardial infarction (AMI) as a cause of death. This is because AMI remains a leading cause of mortality in the Western world (1–3).

At a macroscopic level, AMI can be identified using a nitro blue tetrazolium staining method, and by doing so, a decreased staining intensity identifies lactate dehydrogenase (LDH) leakage and thus infarction areas (4). However, LDH decoloration is only possible from 3 h after onset of AMI onward (5). Routine histochemical analysis can only identify infarctions beyond those 3 h of AMI duration (5).

Ultrastructural analysis of the heart has been used as a method to identify early infarctions. In the ischemic canine heart ultrastructural changes in mitochondria, including swelling of mitochondria, disorganized cristae, and formation of small osmiophilic amorphous densities (6,7), which are composed of lipids and possibly proteins (8,9), were detected at 2 h after onset of AMI (10). In rats, these ultrastructural changes in the mitochondria were observed as early as 1 h after onset of AMI (11). Also in humans, it was shown that AMI induced damage to mitochondria (12,13). The ability to use

electron microscopic changes for the early detection of AMI in ischemic canine hearts has been questioned, as the same ultrastructural changes have been associated with autolysis (10). The same was found in rats (14). These autolytic effects were also shown in human heart tissue, where mitochondrial deposits could be detected as early as 30 min postmortem (15).

In the present study, we analyzed whether quantitative differences exist in the amount of mitochondrial deposits in cardiomyocytes between infarction and noninfarction areas in the heart after AMI and non-AMI hearts in autopsy material and if so whether these can be used to define early infarcts (earlier than 3 h after onset of AMI) in human autopsy.

Materials and Methods

Human Heart Tissue

Human hearts were obtained at autopsy (*n* = 30) as soon as possible but at least within 48 h after death. All patients were seen by medical personnel prior to death. In addition, a clinical diagnosis of AMI was made prior to autopsy. LDH staining (it has to be noticed that tetrazolium used for LDH is harmful) was performed to indicate AMI, decoloration indicates affected myocardium. In patients without LDH decoloration indicating affected myocardium, AMI was clinically defined by ECG. In those patients, left ventricular heart tissue was sampled from suspect areas related to corresponding atherosclerotic changed coronary arteries at risk and/or signs of older infarctions (replacement fibrosis).

In patients with AMI, tissue from the infarction area (left ventricle [LV]) was analyzed via electron microscopy, while in controls

corresponding areas of the heart were analyzed. In addition, tissue from the right lateral ventricular wall of the heart from all subjects was analyzed. In patients with AMI, this noninfarcted tissue served as an internal control.

This study has been approved by and performed according to the guidelines of the ethics committee of the VU Medical Centre, Amsterdam. Use of leftover material after the pathological examination has been completed, is part of the patient contract in our hospital.

Electron Microscopy

Heart tissue of all patients was analyzed using electron microscopy. Heart tissue was fixed in 4% formalin and refixated in 2% (v/v) glutaraldehyde for 30 min and 1.5% (w/v) osmium tetraoxide for 10 min. The tissue was then dehydrated with acetone and embedded in Epon 812. Ultra thin sections were collected on 300-mesh Formavar-coated Nickel grids. The sections were contracted with uranyl acetate and lead citrate and were examined in a Jeol-1200 EX electron microscope. Ten electron microscopy pictures were analyzed per patient, five of the LV and five of the right ventricle (RV) (magnification 7500×). When comparing the mean amount of mitochondria in the left and right ventricle of control patients, a significant higher mean amount of mitochondria was found in the RV (LV: 131 ± 13 per 30 electron microscopy [EM] pictures magnification 7500× vs. RV: 216 ± 25 per 30 EM pictures magnification 7500×, $p = 0.020$). At an ultrastructural level, mitochondrial deposits appeared as black dots within mitochondria (Fig. 1). Such mitochondria were defined as positive and indicative of irreversible cell damage (5). It has to be noticed that those deposits do not differ ultrastructurally from deposits formed by autolysis in control patients (Fig. 1d). Mitochondria with and without deposits were counted separately in each picture. The percentage of positive mitochondria (mitochondria with deposits) in the left or right ventricle was then calculated as the total score of each specimen.

Statistical Analysis

Statistical analysis was performed using the SPSS statistics program (windows version 14.0, SPSS Inc., Chicago, IL). Data were analyzed using a repeated measure ANOVA with post hoc Bonferroni tests and paired T-tests. Levene's test was used for homogeneity of variances. p-Values at the 0.05 level were considered significant. The factor shows the increase of the percentage of positive mitochondria in the LV in comparison to the RV of the same group or LV of another group. The positive predictive value (PPV) was calculated using the internal control (score of RV). We calculated the PPV by scoring the true positive cases (positive for AMI and 1.28 times more deposits in the LV compared to RV) and true negative cases (negative for AMI and 1.28 times more deposits in the LV compared to RV). After that the total amount of true positive cases was divided by the total amount of true positive and true negative cases, times 100%.

Results

Included were 24 patients who died of AMI and six control patients who died from a cause not related to cardiac disease
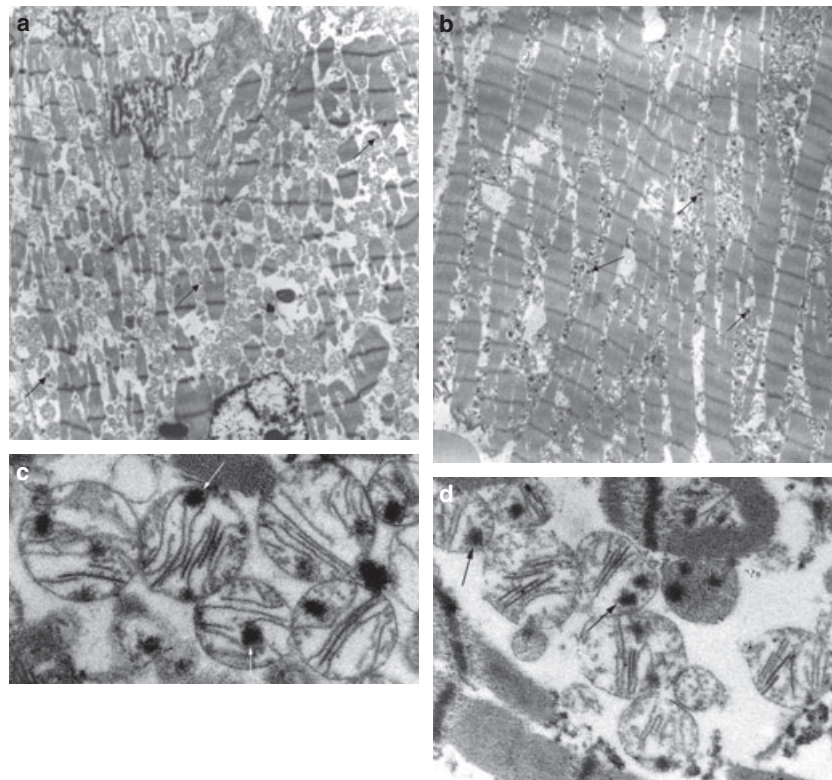

FIG. 1—(a) Electron microscopy picture of mitochondria with deposits (arrows) in the left ventricle of a control patient. Magnification 7500×. (b) Electron microscopy picture of mitochondria with deposits (arrows) in the left ventricle of a patient with acute myocardial infarction (LDH diagnosed). Magnification 7500×. (c) Higher magnification of an electron microscopy picture of deposits (arrows) in the left ventricle of a patient with acute myocardial infarction. Magnification 20 000×. (d) Higher magnification of an electron microscopy picture of deposits (arrows) in the left ventricle of a control patient. Magnification 20 000×. LDH: lactate dehydrogenase.

TABLE 1—*Clinical data of patients included in the study.*

	Controls (n = 6)	Acute Myocardial Infarction (LDH Decolorization) (n = 13)	Acute Myocardial Infarction (Clinically Diagnosed, no LDH Decolorization) (n = 11)
Male	5	10	7
Age mean,	47	58	61
range (in years)	13–77	30–87	34–83
Female	1	3	4
Age mean,	92	65	74
range (in years)	–	49–79	71–79
Cause of death	RI (n = 2) Epileptic insult (n = 1) APE (n = 2) Viral infection lungs (n = 1)	AMI (n = 13)	AMI (n = 11)

LDH: lactate dehydrogenase, RI: respiratory insufficiency, AMI: acute myocardial infarction; APE: acute pulmonary embolism.

(Table 1). In the AMI group, patients were included when no extravascular neutrophilic granulocytes could be detected in the heart indicative for AMI of less then 12 h old (5). In 13 patients with AMI, LDH staining showed LDH decoloration of the affected myocardium indicative of an infarction of 3 h old or older. These patients with AMI had infarction of either the left ventricular anterior wall (n = 7), both lateral and posterior wall (n = 2), posterior wall (n = 1), lateral wall (n = 1), or both lateral and anterior wall (n = 2). In these patients, the infarct age was histologically determined to be between 3–6 h (no neutrophilic granulocytes in blood vessels) (n = 12) and between 6–12 h old (neutrophilic granulocytes in blood vessels but without extravasation) (n = 1) (5). In the remaining 11 patients with AMI, AMI was clinically diagnosed; however, no LDH decoloration was observed. In these patients, the infarct age was therefore determined to be <3 h.

In control patients, the mean percentage of positive mitochondria (Fig. 1a) in the LV of the heart was 48 ± 2%, varying from 26%
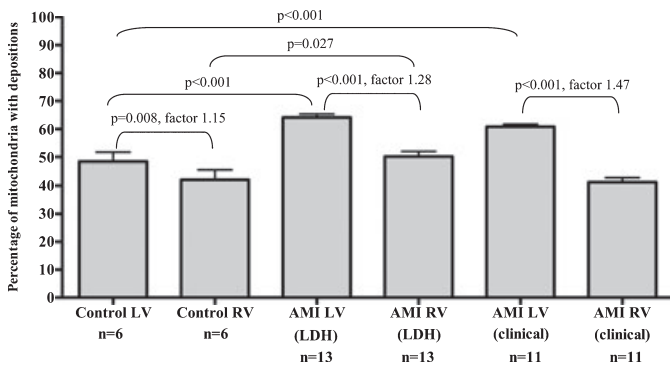

FIG. 2—*Comparison between the mean percentages of mitochondria with deposits in the left and right ventricle of patients with acute myocardial infarction (LDH decoloration and clinically diagnosed, no LDH decoloration) and control patients. Error bars represent the standard error of the mean, factor represents % of mitochondria with deposits of the LV divided by % of mitochondria with deposits of the RV. LV: left ventricle, RV: right ventricle, AMI (LDH): acute myocardial infarction, with LDH decolorization, AMI (clinical): acute myocardial infarction, clinically diagnosed, no LDH decolorization, AMI: acute myocardial infarction, LDH: lactate dehydrogenase.*

up to 68% underlining that indeed because of postmortal autolysis (15), deposits are formed within the mitochondria (Fig. 2). In patients with a LDH-diagnosed AMI (varying from 3 h up to 6–12 h), the mean percentage of positive mitochondria in the infarction area (LV) (Fig. 1b, c), however, was 64 ± 2%, varying from 52% up to 82% (Fig. 2). The mean percentage of positive mitochondria in the infarction area of patients with a LDH-diagnosed AMI was significantly higher than the mean percentage of positive mitochondria in the LV of control patients ($p < 0.001$, factor 1.33). In patients with a clinically diagnosed AMI without LDH decolorization at autopsy (<3 h old), the mean percentage of positive mitochondria in the infarcted area (LV) was 61 ± 2%, varying from 53% up to 67% (Figs. 1 and 2). The mean percentage of positive mitochondria in the infarcted area of these patients with a clinically diagnosed AMI was significantly higher than the mean percentage of positive mitochondria in the LV of control patients ($p < 0.001$). To compare the percentage of positive mitochondria in patients with AMI (both LDH and clinically diagnosed only) between infarcted tissue and noninfarcted tissue within the same patient, the percentage of positive mitochondria was determined in the infarction area from the LV and noninfarcted tissue from the RV. In both groups of patients with AMI, the mean percentage of positive mitochondria in the infarction area was significantly higher (respectively, factor 1.28 and 1.47) than the mean percentage of positive mitochondria in the RV (64 ± 2% vs. 50 ± 2% for LDH diagnosed AMI and 61 ± 2% vs. 41 ± 2% for clinically diagnosed AMI, $p < 0.001$) (Fig. 2). In control patients, the mean percentage of positive mitochondria was also significantly higher in the LV than RV (48 ± 2% vs. 42 ± 3%, $p = 0.008$). However, this was only a factor 1.15 higher compared to 1.28 and 1.47 in patients with AMI. Also the mean percentage of positive mitochondria in the RV of patients with a LDH-diagnosed AMI was significantly higher compared to the RV of control patients (50 ± 2% vs. 42 ± 3%, $p = 0.027$). However, no significant difference was found between the percentage of positive mitochondria in the RV of control patients and the RV of clinically diagnosed patients with AMI.

We also analyzed the number of patients that had a significant increase in the amount of positive mitochondria in the LV compared to the RV in patients with AMI. In patients with a LDH-diagnosed AMI, 10 of 13 (85%) patients had a significantly higher mean percentage of positive mitochondria in the LV than in the RV, while in patients with a clinically diagnosed AMI, 8 of 11 (73%) patients had a significantly higher mean percentage of positive mitochondria in the LV than RV. We also calculated a PPV using the internal control (scores of the RV), and this was found to be 89.47%, meaning that in 89.47% of the cases the finding absolutely indicates AMI.

We next analyzed whether increased time between death and autopsy (Δt [hours]) correlated with increased mitochondrial deposits because of autolysis in the LV and RV of the heart. Hence, the patients with AMI (LDH diagnosed) were divided into two groups: patients with a Δt of ≤12 h and 13 up to 48 h. The mean percentages of positive mitochondria in these two groups in the LV were 62 ± 1% (≤12 h) and 66 ± 2% (13 up to 48 h) respectively, and were not significantly different. However, in the RV, the mean percentages of positive mitochondria in these two groups were 46 ± 1% (≤12 h) and 55 ± 3% (13 up to 48 h), and were significantly different ($p = 0.013$). This shows that patients with AMI (LDH diagnosed) with a higher Δt do not have a higher amount of mitochondrial deposits in the LV; however, the RV patients with a higher Δt does have a higher amount of mitochondrial deposits. The data of the RV indicate that mitochondrial deposits are formed because of autolysis and do increase significantly in accordance

with time elapsed between death and autopsy, at least within the Δt investigated here (up to 48 h). In the LV, no significant increase in mitochondrial deposits was found in accordance with time elapsed between death and autopsy.

Discussion

In the present study, quantitative ultrastructural analysis of mitochondrial amorphous densities or mitochondrial deposits (positive mitochondria) indicative for irreversible cell damage as a method for the determination of early AMI was evaluated.

We have found that in the infarction area of patients with a LDH diagnosed AMI, the percentage of positive mitochondria was significantly higher (factor 1.33) compared to corresponding heart tissue in control patients and compared to noninfarcted areas (RV) within patients with AMI. This was found in 85% of the patients. Also in patients with a clinically diagnosed AMI only (no LDH decoloration at autopsy) a significantly higher amount of positive mitochondria was found in the LV compared to controls and noninfarcted areas (RV) (factor 1.47). This was found in 73% of the patients.

These results thus indicate that ultrastructural analysis of mitochondrial deposits in the heart can be used to define early AMI of less than 3 h old.

It is known that postmortem autolysis can also cause formation of deposits in mitochondria in the canine, rat, and human heart (10,14,15). In the present study, an effect of autolysis was also found in control and patients with AMI; in both patient groups we have found formation of deposits in mitochondria in the left and right ventricle of the heart. However, in patients with AMI (LDH diagnosed), we found an increase in positive mitochondria in the infarction area (LV) compared to noninfarcted area (RV) within the same patient (factor 1.28) and the LV of the heart in control patients (factor 1.33). This thus suggests that besides autolysis, there is an additional effect of ischemia in patients with AMI (LDH diagnosed). In contrast, Ludatscher et al. (15) found that after 2–18 h postmortem, deposits are found in all mitochondria because of autolysis; however, they did not quantify the amount of mitochondrial deposits and in their concomitant figure of myocardial autopsy material taken 18 h after exitus not all mitochondria have deposits. Also they describe that these patients died because of various diseases which was not further specified, and they did not describe any effects of these various diseases on the formation of mitochondria in the heart. Also in our control patients, we did not find deposits in all mitochondria (48 ± 2%) and compared to patients with AMI there was a significant difference in the amount of deposits found.

Furthermore, we also showed that these same patients with AMI with the longest Δt (13 up to 48 h) between time of death and time of autopsy did not have a significant higher percentage of positive mitochondria in the LV compared to patients with a lower Δt (≤12 h). However, in the RV a significant increase was found, indicative that there is an autolytic effect which does increase in time. As is also shown in the control patients, there is an autolytic effect; however, in the LV of patients with AMI, the percentage of mitochondria with deposits because of the AMI may be so high that an additional autolytic effect is too small in comparison to alter the percentage of positive mitochondria.

In conclusion, our results indicate that ultrastructural analysis of mitochondrial deposits is a valid method to detect early myocardial infarction of less than 3 h old when the amount of mitochondrial deposits is a factor 1.28 larger in the area of the heart suspected of putative AMI, compared with distant, noninfarcted part of the heart
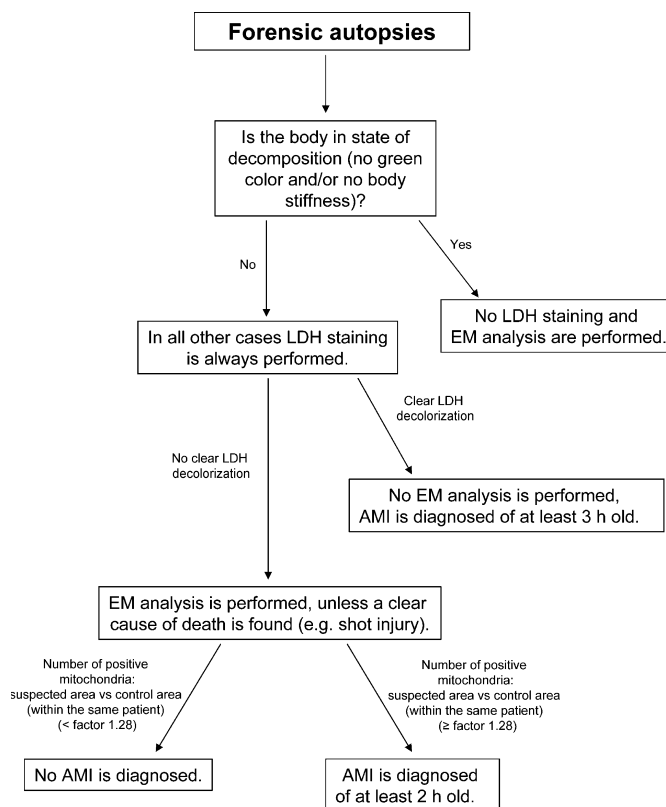

FIG. 3—*Diagram when to use LDH staining and EM analysis for diagnosing AMI in forensic autopsies. LDH: lactate dehydrogenase, EM: electron microscopy, AMI: acute myocardial infarction, h: hours.*

of the RV of the same patient, or a factor 1.33 when compared to the noninfarcted LV of a control patient.

According to literature, the detection of deposits in mitochondria can be seen at earliest at 2 h post AMI (10). This ultrastructural analysis is a particularly useful method in cases of sudden or unexplained death in case LDH staining is not conclusive at autopsy. Electron microscopy then can contribute significantly in diagnosing an AMI in between 2 and 3 h old, thus before LDH decolorization. A diagram when to use LDH staining and EM analysis for diagnosing AMI is depicted in Fig. 3.

When using this method, heart samples should be collected during autopsy from suspected areas of the heart related to corresponding atherosclerotic changed coronary arteries at risk and/or signs of older infarctions (replacement fibrosis). It is also recommended to take samples from noninfarcted tissue from the same patient to indicate the "background" signal caused by autolysis. Suspected infarcted and noninfarcted tissue should then be compared to determine a possible AMI. The cost of electron microscopy analysis in the Netherlands is approximately 250 euro (330 US dollars).

However, our results also show that this method is not absolute and that in our study, in approximately 25% of the patients with an infarction of less than 3 h old we could not detect evidence for myocardial infarction at an ultrastructural level, although a sampling error never can be excluded in those cases.

References

1. de la Grandmaison GL. Is there progress in the autopsy diagnosis of sudden unexpected death in adults? Forensic Sci Int 2006;3:138–44.
2. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: global burden of disease study. Lancet 1997;349(9064):1498–504.

3. Murray CJ, Lopez AD. Mortality by cause for eight regions of the world: global burden of disease study. Lancet 1997;349(9061):1269–76.

4. Nachlas MM, Shnitka TK. Macroscopic identification of early myocardial infarcts by alterations in dehydrogenase activity. Am J Pathol 1963;42:379–405.

5. Robbins SL, Cotran RS. Pathologic basis of disease. In: Kumar V, Abbas AK, Fausto N, editors. Disease of organ systems: the heart. Philadelphia, PA: Elsevier Saunders, 2005;577–81.

6. Jennings RB, Baum JH, Herdson PB. Fine structural changes in myocardial ischemic injury. Arch Pathol 1965;79:135–43.

7. Jennings RB, Ganote CE. Structural changes in myocardium during acute ischemia. Circ Res 1974;35(Suppl 3):156–72.

8. Buja LM, Dees JH, Harling DF, Willerson JT. Analytical electron microscopic study of mitochondrial inclusions in canine myocardial infarcts. J Histochem Cytochem 1976;24(3):508–16.

9. Jennings RB, Shen AC, Hill ML, Ganote CE, Herdson PB. Mitochondrial matrix densities in myocardial ischemia and autolysis. Exp Mol Pathol 1978;29(1):55–65.

10. United States, Canadian Academy of Pathology I. Cardiovascular pathology, clinicopathologic correlations and pathogenetic mechanisms, 37th edn. Philadelphia, PA: Williams & Wilkins, 1995.

11. Yanagiya N. Ultrastructural and histochemical changes of mitochondria in global ischemic cardiac muscle of rat. Cell Mol Biol (Noisy-le-grand) 1994;40(8):1151–64.

12. Hein S, Scheffold T, Schaper J. Ischemia induces early changes to cytoskeletal and contractile proteins in diseased human myocardium. J Thorac Cardiovasc Surg 1995;110(1):89–98.

13. Schaper J. Effects of multiple ischaemic events on human myocardium—an ultrastructural study. Eur Heart J 1988;9(Suppl A):141–9.

14. Tomita Y, Nihira M, Ohno Y, Sato S. Ultrastructural changes during in situ early postmortem autolysis in kidney, pancreas, liver, heart and skeletal muscle of rats. Leg Med (Tokyo) 2004;6(1):25–31.

15. Ludatscher RM, Hashmonai M, Peleg H. The irreversible ischaemic lesion of human myocardium. Comparison with the experimental animal model. Acta Anat (Basel) 1984;118(2):91–5.

Additional information and reprint requests:
Mark P. V. Begieneman, B.Sc.
VU Medical Center
Department of Pathology
Room no. 0E46
de Boelelaan 1117
1007 HV Amsterdam
the Netherlands
E-mail: mpv.begieneman@vumc.nl

PAPER

PATHOLOGY AND BIOLOGY

Timothy G. Baumer,[1] *B.S.; Nicholas V. Passalacqua,*[2] *M.S.; Brian J. Powell,*[1] *B.S.;*
William N. Newberry,[3] *M.S.; Todd W. Fenton,*[2] *Ph.D.; and Roger C. Haut,*[1] *Ph.D.*

Age-Dependent Fracture Characteristics of Rigid and Compliant Surface Impacts on the Infant Skull—A Porcine Model*,†

ABSTRACT: This study documents skull fracture characteristics on infant porcine specimens under known impact conditions with respect to age and interface. A single impact causing fracture was conducted on the skull of porcine specimens aged 2–28 days ($n = 76$). Paired rigid and compliant impacts at the same energy were conducted at each specimen age. Impact force, impact duration, and fracture length were recorded. Energy required to initiate skull fracture increased with specimen age. For a given energy, impact of the skull with a compliant interface caused more fracture damage than with a rigid interface for specimens aged under 17 days, but less damage for specimens aged 24–28 days. The documentation of energy required to cause fracture and resulting fracture propagation with respect to impact interface and age may be of critical importance in forensic investigations of infant skull trauma.

KEYWORDS: forensic science, skull fracture, infant, porcine, interface dependency, age dependency

Owing to the paucity of skull fracture tolerance data for infants and young children, pediatric trauma involving single-event head injuries with related cranial fractures represents one of the greatest challenges to forensic pathologists and anthropologists. Head injuries account for 80% of fatal child abuse in younger children (1); however, the most common injuries from accidental falls of 2 m or more also occur to the head (2). Furthermore, in a study of 64 falls from 3 m or more of subjects ranging from 0–18 years, 48% of the subjects with head injuries were under 5 years of age and 50% of these head injuries involved skull fracture (3). Distinguishing between cases of inflicted trauma (abuse) and accidental trauma (falls) can be difficult, as both may produce similar types of injuries (4). The most commonly fractured cranial bone in both accidental injury and abuse cases is the parietal (5–7), and linear, complex, and depressed fractures can be seen in both types of cases (8,9). The risk of injury is also dependent on the contacting surface and fall height (10,11). Other variables, such as the area struck, thickness of the skull, thickness of scalp and hair, and impact direction can also affect the fracture pattern (12,13). To develop reliable models for skull fracture prediction, the effects of these aforementioned variables must be better understood in scenarios relating to both inflicted and accidental trauma.

While several studies have documented important information regarding the response of the infant skull and brain to blunt force impact, these data are limited to one or two age groups. And because of ethical considerations, experimentation on human pediatric specimens is very limited. While scaling of the adult skull has met with some success in predicting impact response of the pediatric skull (14), the head of an infant is smaller and geometrically unlike that of an adult (15) and the validity of predicting skull fracture patterns in infants from adult data has not been investigated. Using adult data to predict skull fracture patterns in the pediatric skull may also be problematic because of the different structural (ease of deformation and decreased threshold to fracture) and mechanical properties of the infant skull (16–18). However, one animal model of the human infant has begun to emerge regularly in the current literature. A study by Dickerson and Dobbing (19) shows a similarity in development of the central nervous system when correlating months of the human to weeks of the pig. Other studies have used porcine models to predict fracture loads for the infant human femur (20) and strain in the braincase and sutures (21). More recently, age-related rates of change in the bending rigidity of the infant human and infant porcine skulls has been found to be similar for a correlation based on months in humans to days in pigs (22).

The hypotheses of the current study were twofold. First, impact energy required to initiate a skull fracture would depend on specimen age. Second, the extent of fracture damage for a given energy impact would depend on impact interface. These data using a developing porcine model may provide a baseline for the characterization of fracture patterns resulting from blunt force impacts onto

different contact interfaces for the developing human pediatric skull, the results of which may have ultimate utility in discriminating between fatal cases of accidental injury versus inflicted abuse.

Materials and Methods

Porcine specimens aged from 2–28 days ($n = 76$) were obtained from a local supplier and stored at −20°C. All animals died of natural causes. The specimens were frozen within 12 h of death. None of the specimens had outward signs of head injury, which was later confirmed during specimen preparation.

Each specimen was thawed at room temperature for 24 h prior to testing. All soft tissue was removed from the left side of the skull, leaving the periosteum to help prevent drying of the bones and sutures. This side of each skull was placed in a bed of air-hardening epoxy (Fiber Strand; Martin Senour Corp., Cleveland, OH). Capturing the skull in this manner allowed deformation of the skull bones and sutures, but eliminated gross motions (translations and rotations) of the skull as a whole. After the potting material had hardened, the skull was bathed in phosphate-buffered saline solution in regular intervals until the experimental impact. Each skull, with the scalp in place, was impacted in the center of the right parietal bone. The impact location was controlled by a custom-made, 4 degree of freedom, lockable fixture that was built to hold the potted skull (Fig. 1). The skull was oriented so that the impact was normal to the bone surface.

Each specimen was subjected to a single impact with a gravity-accelerated mass (GAM) (Fig. 2). A load transducer (4.45 kN capacity, model AL311CV; Sensotec, Columbus, OH) was mounted on the GAM behind the impact interface to record forces. A single impact was assured by using an operational amplifier comparator circuit to monitor the impact forces and energize an electromagnetic solenoid to catch the impactor immediately after the load returned to zero. Two interchangeable impact interfaces were created for the GAM to represent rigid and compliant impact interfaces. The rigid interface was a solid aluminum cylinder with approximately 16 cm^2 of impact surface. The compliant interface was a deformable aluminum material (1.10 MPa crush strength Hexcel; Hexcel Corp., Stamford, CT), approximately 3 cm thick with a 16 cm^2 impact surface, fastened to the rigid interface. Both interfaces had greater surface area than the resulting contact area on the skull.

Input energy was adjusted by changing the drop height of the 1.67-kg GAM. A second mass of 1.92 kg was necessary for specimens 21 days and older, as the drop height was limited in the laboratory. The mass of the impact interface for each impact was included in the total mass of the GAM, and the force data were
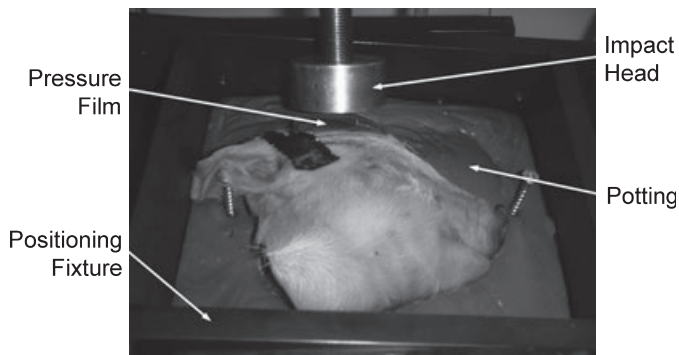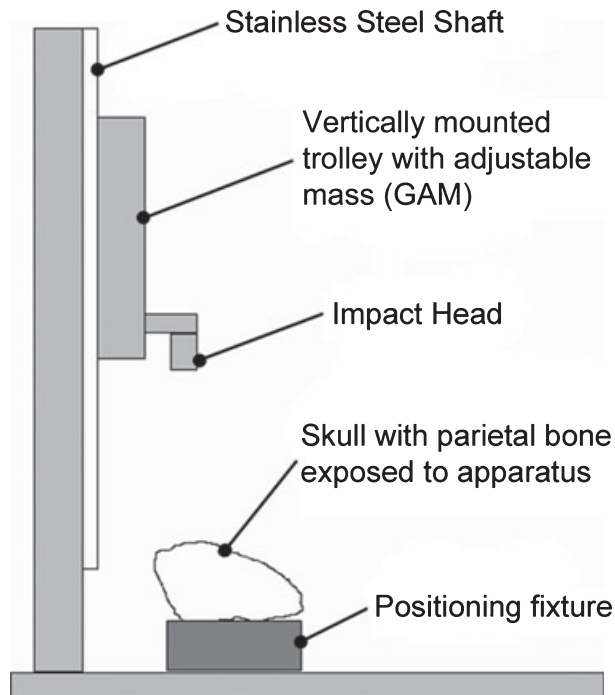

FIG. 2—Drop test fixture set-up.

checked for inertial effects. Preliminary data collected at several ages were used to estimate the input energy needed to initiate a fracture using the rigid interface for each specimen age. In a few cases ($n = 7$), the first impact with a rigid interface did not cause a fracture, so a second, higher energy impact was performed. Impacts were always conducted in pairs. All skulls impacted with the compliant interface were impacted at the same energy that caused fracture in the paired rigid impact on another skull of the same age. While the input energy was the same for rigid and compliant impacts at each age, the energy of impact required to cause fracture initiation had to be increased with specimen age. The forces and impact durations were recorded at 10,000 Hz.

Pressure-sensitive film packets (Prescale; Fuji Film Ltd., Tokyo, Japan) were attached with tape to the skull at the impact site to measure contact area and pressures generated during impact. Low- (0–10 MPa) and medium- (10–50 MPa) range pressure films were stacked together and sealed between two sheets of polyethylene (0.04 mm thick) to prevent exposure to fluids during the impact (23). Using the previously documented procedure, the film was scanned and converted to pressure using a dynamically calibrated density-to-pressure scale. Only the low-pressure film was analyzed in these experiments as the pressures were not high enough to record on the medium-pressure film.

Computed tomography scans were conducted for each skull after impact (625-μm-thick coronal slices). The width of the skull vault and average thickness of the parietal bone were obtained from these scans using reconstruction software (Osirix version 3.0, Open Source General Public License). The width of the vault was measured at its widest point from parietal bone to parietal bone. Average parietal bone thickness was determined from a slice near the center of the parietal bone.

Visual inspection was conducted on each skull after the soft tissue and periosteum were removed from the impacted side. All suture damage and any visible bone fractures were recorded. Each specimen was then cleaned via standard anthropological procedures,

FIG. 1—Skull positioning and set-up showing rigid impact head.

removing all remaining soft tissue. Complete fracture diagrams and measurements were made for each specimen. Total length of all skull fractures was measured to the nearest millimeter with a flexible measuring tape, which conformed to the curvature of the skull. In cases of multiple locations or complex (nonlinear) fractures, each fracture was measured and summed for a combined total fracture length.

The impact and pressure data were analyzed for age effects using linear regression analyses. Comparisons between the interfaces were performed using a two-factor (age, interface) ANOVA. Statistically significant effects were reported for $p < 0.05$.

Results

Linear regression analyses showed a significant increase (0.0516 mm/day) in average parietal bone thickness with age ($p < 0.001$), but no change in skull vault width with age.

The input energy required to cause fracture increased from 1.5 J at 2 days of age to 11.5 J at 28 days of age. A characteristic rapid drop in force appeared in the force–time plot for all rigid interface impacts that were associated with fracture (Fig. 3). In rigid interface impacts that did not cause fracture, a more pronounced peak was observed. Rapid drops in force during compliant interface impacts could not always be associated with fractures because of the deformation of the interface.

For both impact interfaces, linear regression analyses showed a significant increase in impact force with age that was required to produce a fracture ($p < 0.001$) (Fig. 4). The slopes and intercepts of impact forces versus age were not different based on interface.

The area of contact for the compliant interface significantly increased with age ($p < 0.001$), but no significant change with age was noted for the rigid interface (Fig. 5).

Bone fractures for all ages, and for both interfaces, occurred at sites away from the point of impact (Fig. 6). These bone fractures initiated at a suture–bone interface. Frequently, bone fractures were accompanied by diastatic fractures that were generally located near the fracture tip at the bone–suture interface. Diastatic fractures caused by a rigid interface impact were not present until specimens were 10 days of age, while diastatic fractures caused by a compliant interface were present as early as 4 days of age (Fig. 7). Diastatic fractures were common for both interfaces between 10 and 17 days. Neither interface showed suture damage after 17 days.

A two-factor (age, interface) ANOVA performed on the total fracture length (bone and diastatic combined) revealed significantly
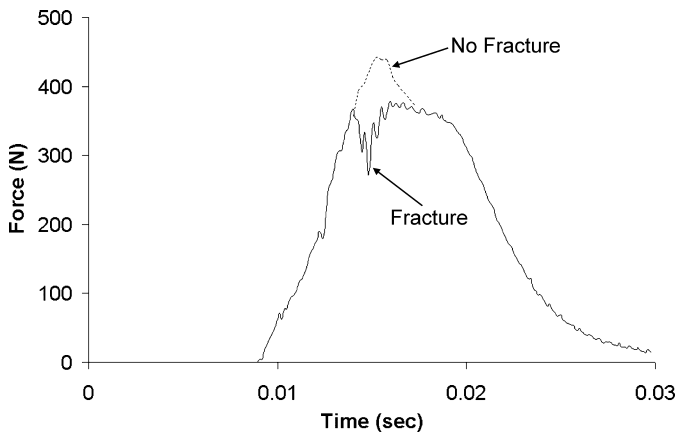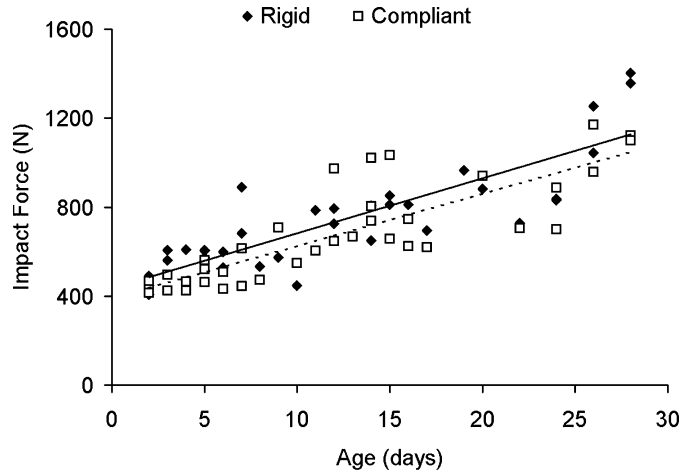

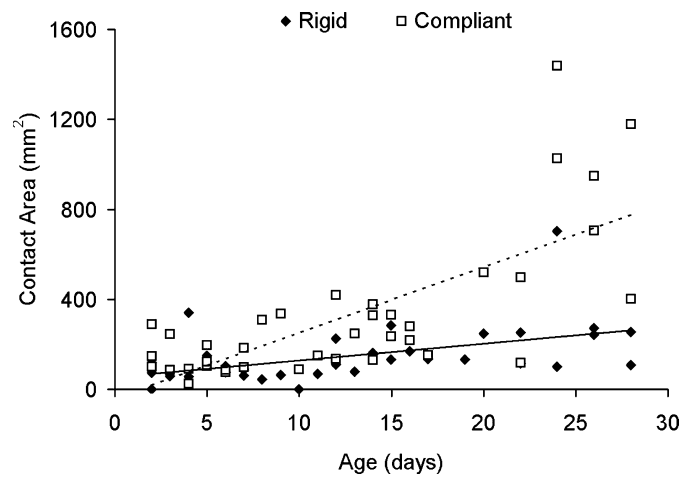FIG. 4—*Peak force during impact for rigid and compliant interfaces.*

FIG. 5—*Contact area recorded on pressure sensitive film during impacts for rigid and compliant interfaces.*

more damage caused by the compliant interface than by the rigid interface on pig skulls <17 days of age, similar amounts of damage from both interfaces between 17 and 22 days of age, and significantly more damage caused by the rigid interface than by the compliant interface for specimens aged 24–28 days ($p < 0.05$) (Fig. 8).

Discussion

In an effort to better understand fracture of the developing skull under known conditions, the current study examined characteristics of impact-induced fracture to the infant porcine skull with respect to age and impact interface. As would be expected, the peak impact force that produced fractures increased with age. While an increase in the bending modulus of the porcine skull with age would account for this increase in force with age, a previous study shows that the bending modulus of infant porcine bone and bone–suture–bone specimens under 4-point bending does not change significantly with specimen age (22). Measurements of the average parietal bone thickness, however, showed a significant increase with age, which seemed to correlate with changes in the impact force. It is expected that the increasing strength of the skull with age is because of changes in geometry (thickness) and not because of an increase in bone quality with age.
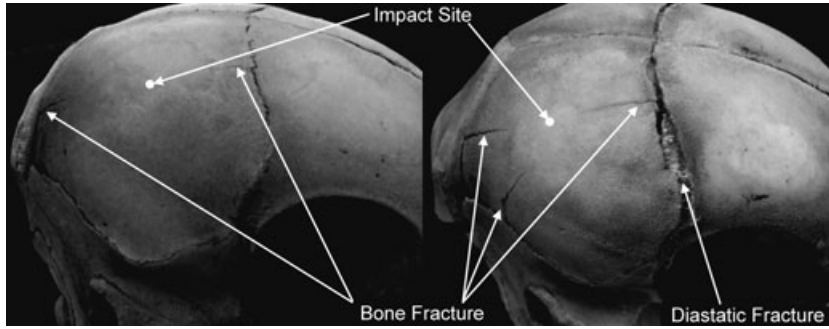

FIG. 3—*Overlay of force–time plots showing typical characteristics of an impact that caused fracture and one that did not.*

FIG. 6—*Representative fracture sites for (a) rigid and (b) compliant interface impacts (five-day-old specimens shown). Fracture sites initiated at bone–suture boundaries. Compliant interface impact showed increased fracture lengths at similar initiation sites as the rigid interface impacts as well as diastatic fracture of coronal suture.*
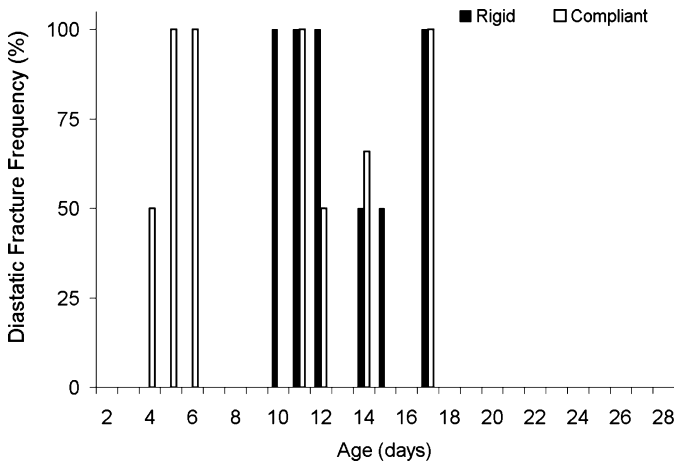

FIG. 7—*Frequency of diastatic fractures caused by both rigid and compliant interfaces versus age.*
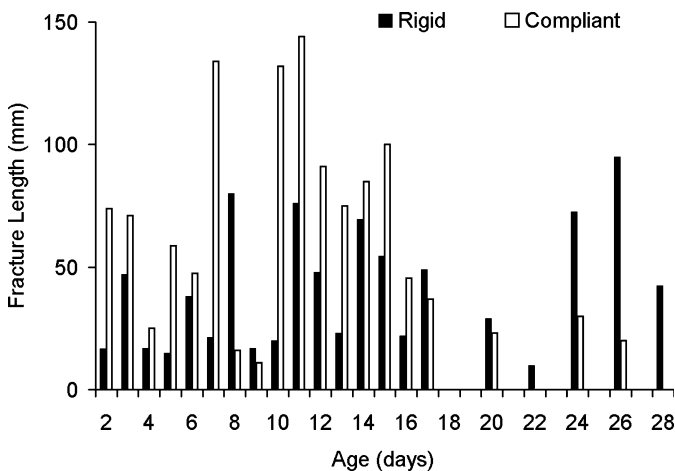

FIG. 8—*Average fracture length generated by rigid and compliant impacts at the same impact energy.*

There has been debate in the forensic community as to the mechanism of fracture initiation on adult human skulls, specifically, whether fractures initiate at the point of impact (24) or at remote sites of "outbending" (25). In the current study, 70 of 76 infant porcine specimens displayed fracture initiation at a suture–bone boundary and not at the center of the parietal bone (the impact location). The other six specimens involved comminuted skull fractures that obscured the site of initiation. The locations of fracture initiation on infant porcine specimens in the current study agreed with the findings of Gurdjian et al. (25) for adult human skulls; however, it is currently unclear whether the fracture initiation sites observed in this study are because of the immature nature of the cranial bone, differences in porcine cranial geometry, or a combination of these factors. Despite potential differences limiting direct comparisons with adult human fracture information, the infant porcine model developed in the current study showed that fracture initiation occurred in relatively predictable locations at bone–suture interfaces for impacts on the parietal bone against a flat surface. This information may be expanded in the future to other interface conditions and should be correlated with fracture patterns in the human infant skull from case files. These findings suggest that skull fractures on the developing skull from impacts with rigid or compliant flat surfaces may be located remote from the site of the impacting force and nearer to the cranial sutures.

The current study produced a greater amount of fracture on younger infant porcine skulls using a compliant than rigid interface. The concentrated force of an impact with a rigid interface was expected to produce a larger amount of bone fracture than an impact with a compliant interface where the same force intensity would be distributed over a larger contact area and contact stresses would be reduced. This study, however, showed that the relative amount of fracture between these interfaces was dependent on age. For infant pigs <17 days of age, a compliant interface caused significantly more damage than a rigid interface for the same impact energy, while for specimens 24 days of age and older, the rigid interface generated more damage than the compliant interface. Specimens younger than 17 days of age also showed frequent damage to the sutures, while specimens older than 17 days of age had bone fractures alone. In a previous study on beam specimens cut from infant porcine skulls and loaded in 4-point bending, it has been shown that there is a significant difference in the relationship between bone and bone–suture–bone material properties for specimens aged 14 and 18 days (22). These changes resulted in a significant increase in the relative strength of the coronal suture at 18 days to values more similar to that of bone. This effect correlates well with the disappearance of suture damage and a change in interface-dependent fracture differences documented in the current study. The age-dependent response of the developing skull to impact with rigid and compliant surfaces is important in forensic cases for determining potential differences between testimony describing the events leading to injury and the observed fracture pattern. Specifically, the infant porcine model suggests that very young skulls may actually be more susceptible to extensive skull fracturing from an impact against a compliant surface than from an

impact against a more rigid surface. Further investigations will be needed to create a more comprehensive set of patterns for various impact interfaces and energies.

In an effort to explain the interface- and age-dependent results of the current study, it was noted that the contact area for the compliant interface impacts was generally greater than that for the rigid interface impacts. Yet, the generated impact force was not different between interfaces. It was hypothesized that the increased contact area for the compliant interface impacts resulted in a higher state of stress close to the boundaries of the parietal bone (sutures). It may then be suspected that the mechanical properties of the sutures were more critical during impacts with a compliant than with a rigid interface. For specimens under 17 days of age, when the sutures are significantly more compliant than bone, large deformations at the sutures during impact for both interfaces may have generated high levels of stress near the boundaries of the parietal bone at the sutures. Therefore, the risk of damage at the suture may have increased, resulting in the high frequency of diastatic fractures. Furthermore, the higher state of peripheral stresses caused by the compliant interface may have promoted fracture propagation, resulting in the increased fracture length documented for the compliant interface versus the rigid interface. The increase in suture strength that occurs between 14 and 18 days of age may have allowed a better distribution of impact stresses across the boundaries of the parietal bone, resulting in lower overall levels of stress in the parietal bone and across suture lines. This effect may help explain why in the relatively older infants the compliant interface actually generated less fracturing to the skull than for the rigid interface, because the rigid interface would be expected to generate more localized contact stresses nearer the site of impact.

In summary, this study documented an increase in impact energy required to initiate skull fractures as specimen age increased. Furthermore, there was relatively more fracture damage generated for a compliant than a rigid interface with equal input energy at the younger infant ages. While a direct correlation with infant human fracture patterns is not currently available from case files, the experimental observations from the developing porcine model may aid in understanding some of the limited data that is available today. The experimental documentation of fracture initiation energy and the extent of skull fracture with respect to interface as a function of specimen age may be of critical importance in future investigations of infant cranial bone fracture patterns as well as hold significant forensic implications for the diagnosis of fatal child abuse.

References

1. Case ME, Graham MA, Handy TC, Jentzen JM, Monteleone JA. Position paper on fatal abusive head injuries in infants and young children. Am J Forensic Med Pathol 2001;22:112–22.
2. Yagmur Y, Guloglu C, Aldemir M, Orak M. Falls from flat-roofed houses: a surgical experience of 1643 patients. Injury 2004;35(4):425–8.
3. Lallier M, Bouchard S, St-Vil D, Dupont J, Tucci M. Falls from heights among children: a retrospective review. J Pediatr Surg 1999;34(7):1060–3.
4. Billmire ME, Myers PA. Serious head injury in infants: accident or abuse? Pediatrics 1985;75:340–2.
5. Hobb CJ. Skull fracture and the diagnosis of abuse. Arch Dis Child 1984;59:246–52.
6. Meservy CJ, Towbin R, McLaurin RL, Myers PA, Ball W. Radiographic characteristics of skull fractures resulting from child abuse. Am J Roentgenol 1987;149:173–5.
7. Leventhal JM, Thomas SA, Rosenfield NS, Markowitz RI. Fractures in young children: distinguishing child abuse from unintentional injuries. AMA Am J Dis Child 1993;147:87–92.
8. Reece R, Sege R. Childhood head injuries: accidental or inflicted. Arch Pediatr Adolesc Med 2000;154:11–5.
9. Wheeler DS, Shope TR. Depressed skull fracture in a 7-month old who fell from bed. Pediatrics 1997;100:1033–4.
10. Bertocci GE, Pierce MC, Deemer E, Aguel F, Janosky JE, Vogeley E. Using test dummy experiments to investigate pediatric injury risk in simulated short-distance falls. Arch Pediatr Adolesc Med 2003;157(5):480–6.
11. Chalmers DJ, Marshall SW, Langley JD, Evans MJ, Brunton CR, Kelly AM, et al. Height and surfacing as risk factors for injury in falls from playground equipment: a case-control study. Inj Prev 1996;2:98–104.
12. Knight B. Forensic pathology. London: Edward Arnold, 1991.
13. Cooperman DR, Merten DF. Skeletal manifestation of child abuse. In: Reece M, Ludwig S, editors. Child abuse: medical diagnosis and management. Philadelphia, PA: Lippincott, Williams, and Wilkins, 2001;123–56.
14. Prange MT, Luck JF, Dibb A, Van Ee CA, Nightingale RW, Myers BS. Mechanical properties and anthropometry of the human infant head. Stapp Car Crash Journal 2004;48:279–99.
15. Schneider LW, Lehman RJ, Pflug MA, Owings CL. Size and shape of the head and neck from birth to four years. Washington, DC: The Consumer Product Safety Commission, 1986. Report No.: UMTRI-86-2.
16. Thibault KL, Margulies SS. Age-dependent material properties of the porcine cerebrum: effect on pediatric head injury criteria. J Biomech 1998;31(12):1119–26.
17. Prange MT, Meaney DF, Margulies SS. Defining brain mechanical properties: effects of region, direction, and species. Stapp Car Crash Journal 2000;44:205–13.
18. Prange MT, Margulies SS. Regional, directional, and age-dependent properties of the brain undergoing large deformations. J Biomech Eng 2002;124(2):244–52.
19. Dickerson J, Dobbing J. Prenatal and postnatal growth and development of the central nervous system of the pig. Proc R Soc Lond 1967;166(1005):384–95.
20. Pierce M, Valdevit A, Anderson L, Inoue N, Hauser D. Biomechanical evaluation of dual-energy x-ray absorptiometry for predicting fracture loads of the infant femur for injury investigation: an in vitro porcine model. J Orthop Trauma 2000;14(8):571–6.
21. Herring S, Teng S. Strain in the braincase and its sutures during function. Am J Phys Anthropol 2000;112:575–93.
22. Baumer TG, Passalacqua NV, Powell BJ, Fenton TW, Haut RC. Age dependent mechanical properties of the infant porcine parietal bone and a correlation to the human. J Biomech Eng 2009;131:111006-1/6.
23. Atkinson PJ, Newberry WN, Atkinson TS, Haut RC. A method to increase the sensitive range of pressure sensitive film. J Biomech 1998;31(9):855–9.
24. Kroman AM. Experimental study of fracture propagation in the human skull: a re-testing of popular theories. Proceedings of the 56th Annual Meeting of the American Academy of Forensic Sciences; 2004 Feb 16–21; Dallas, TX. Colorado Springs, CO: American Academy of Forensic Sciences, 2004;CH79:314.
25. Gurdjian ES, Webster JE, Lissner HR. The mechanism of skull fracture. Radiology 1950;54(3):313–39.

Additional information and reprint requests:
Roger C. Haut, Ph.D.
Orthopaedic Biomechanics Laboratories
A407 East Fee Hall
Michigan State University
East Lansing, Michigan 48824
E-mail: haut@msu.edu

PAPER

PATHOLOGY AND BIOLOGY; TOXICOLOGY

D. Kimberley Molina,[1] *M.D.*

A Characterization of Sources of Isopropanol Detected on Postmortem Toxicologic Analysis

ABSTRACT: Isopropanol is an important chemical to forensic pathologists in that intoxication can result in death yet presence does not necessarily indicate intoxication. Several reports have been published, which indicate that isopropanol can be created endogenously in certain situations including diabetes mellitus, starvation, dehydration, and chronic ethanol use; however, a large-scale analysis addressing all of the possible causes of postmortem isopropanol detection has not been performed. A retrospective review of all cases examined at the Bexar County Medical Examiner's Office between 1993 and 2008 in which isopropanol was detected in routine alcohol screening was undertaken. The cases were categorized by the source of the isopropanol, and the concentrations of isopropanol and acetone were analyzed. Analysis revealed isopropanol concentrations to be low (<100 mg/dL) in cases of antemortem and postmortem creation and in postmortem contamination and high (>100 mg/dL) in cases of antemortem exposure. These results are consistent with other published reports.

KEYWORDS: forensic science, isopropanol, forensic toxicology, acetone, postmortem, diabetes, infection, decomposition

Isopropanol, also known as rubbing alcohol, is a commonly used antiseptic and disinfectant and is often found as a component of antifreeze and window cleaners. If ingested, isopropanol is rapidly absorbed from the gastrointestinal tract, although it can also be passively absorbed through the skin. It is less toxic than methanol or ethylene glycol but is a more potent central nervous system (CNS) depressant than ethanol (1). In addition to CNS depression, isopropanol intoxication can result in vasodilation, hypotension, pulmonary edema, and gastritis with secondary gastrointestinal bleeding (1). The toxicity of isopropanol is based on the parent compound itself, and not its metabolite, unlike methanol and ethylene glycol, which are metabolized to toxic aldehydes.

Isopropanol is metabolized by alcohol dehydrogenase (ADH) to acetone. It is excreted by the kidneys and lungs and has an approximate half-life of 2.5–8 h (1), although concurrent ethanol administration can prolong the half-life because of competition for the ADH enzyme. Metabolically, isopropanol results in ketonemia without acidemia and produces an osmolal gap. Signs and symptoms of toxicity usually develop within 2–3 h after ingestion and include lethargy, hypotension, abdominal pain, nausea/vomiting, headache, respiratory depression, shock, and/or coma. Individuals exposed to isopropanol may exude a fruity smell, secondary to the ketones, similar to those experiencing diabetic ketoacidosis. Ingestion of 250 mL of isopropanol has been reported to cause death, although survival has been reported after ingestion of 500 mL (1). Treatment of isopropanol intoxication is mostly supportive, though hemodialysis has been used successfully.

Isopropanol has been reported in the blood in cases of diabetes, starvation, dehydration, infection, and in the setting of chronic ethanol use. Davis et al. examined the metabolism of acetone to explain how isopropanol is created in these situations (2). Acetone

is created when two molecules of acetyl coenzyme A (acetyl CoA) combine to form acetoacetate, which is then decarboxylated. Acetyl CoA is a component in the tricarboxylic acid cycle, but when it is present in abundance, acetone is created. Acetyl CoA is created as a byproduct of fatty acid metabolism and is increased in disease states which utilize a great deal of fatty acids, such as diabetes mellitus, starvation, and chronic ethanol intake. These disease states can result in increased concentrations of acetyl CoA which is then converted to acetone. Isopropanol is converted to acetone by ADH; however, this directionality can be reversed under certain conditions. ADH can convert acetone to isopropanol in the presence of H_2O and NAD^+ and especially in the presence of not only increased acetone levels, but increased NAD^+ as well, which is also found in increased amounts in the disease states discussed previously. Lewis et al. (3) took this hypothesis one step further and found that isopropanol could routinely be detected in the blood of rats after exposure to acetone.

Isopropanol is an important chemical to forensic pathologists in that intoxication can result in death yet presence does not necessarily indicate intoxication. A comprehensive review of all deaths occurring at the Bexar County Medical Examiner's Office (BCMEO) in which isopropanol was detected on routine toxicologic analysis was undertaken in an attempt to clarify the differing issues affecting the interpretation of isopropanol concentrations.

Materials and Methods

A database search was performed for all cases examined at the BCMEO between 1993 and 2008 (inclusive) in which isopropanol was detected in routine alcohol screening. During the time periods studied, alcohol detection and quantitation were performed by direct injection method analyzed by gas chromatography paired with a flame ionization detector. The cases were then evaluated for the source of the isopropanol, and the concentrations of isopropanol, ethanol, and acetone were assessed, when present.

[1]Bexar County Medical Examiner's Office, San Antonio, TX.

A concentration of 0 mg/dL denotes that the chemical was not detected on routine screening. The limit of detection of isopropanol and acetone had some variability during the time period of the study, but averaged 1 mg/dL for both isopropanol and acetone.

Results

A total of 152 decedents were identified in which isopropanol was detected on routine alcohol analysis. Nine sources of the isopropanol were identified which fell into one of four categories: antemortem exposure to isopropanol (intoxication); creation of isopropanol either antemortem or postmortem (diabetes mellitus, infection, dehydration/starvation, chronic ethanol use, decomposition); postmortem contamination (tissue and/or cornea procurement, embalming); and unknown. The data is summarized in Tables 1 and 2.

Exposure

During the period queried, seven individuals were found that had an antemortem exposure to isopropanol, all of which died as a result of this exposure (isopropanol intoxication). The median concentration of isopropanol in the blood found in exposure cases was 175 mg/dL (range 148–3700 mg/dL) with a median acetone concentration of 150 mg/dL (range 7–200 mg/dL). The median concentration of isopropanol in the vitreous was 240 mg/dL (range 130–244 mg/dL) with a median acetone concentration of 159 mg/dL (range 7–190 mg/dL). The average isopropanol-to-acetone ratio was 5.5 (range 0.8–25) in the blood and 5.6 (range 1–18.6) in the vitreous. In all but one case, the blood isopropanol concentration was greater then the blood acetone concentration. In that case, the decedent was witnessed drinking a bottle of isopropanol prior to becoming unresponsive. The blood isopropanol concentration was 160 mg/dL with an acetone concentration of 200 mg/dL. Unfortunately, vitreous testing was not performed. The explanation for this discrepancy is unknown.

In only one case was the vitreous isopropanol concentration less than the acetone concentration (vitreous isopropanol 181 mg/dL, acetone 185 mg/dL). In that case, the blood isopropanol level was greater than that of acetone (blood isopropanol 148 mg/dL, acetone 134 mg/dL). In all exposure cases, the concentration of isopropanol was >100 mg/dL in the blood and vitreous.

Postmortem Contamination

A total of 34 cases were identified where the presence of isopropanol was thought to be an artifact due either to embalming (15 cases) or postmortem tissue or cornea procurement (19 cases). The median isopropanol concentration in the vitreous for the embalmed cases (no blood was obtained in these cases) was 20 mg/dL (range 0–36 mg/dL). Acetone was not detected in a single embalmed case although methanol was detected in all but one. For postmortem procurement cases, the median blood isopropanol concentration was 25 mg/dL (range 0–201 mg/dL) and in vitreous was 18 mg/dL (range 0–92 mg/dL). Acetone was not present in any of the cases, either in the blood or vitreous fluid. In 89.5% of the procurement cases ($n = 17$), the isopropanol concentrations (blood or vitreous) were <100 mg/dL.

Creation

Creation of isopropanol within the body by either microorganisms, as in the case of infection or decomposition, or as a byproduct of increased acetone concentrations, as in diabetes mellitus, starvation, dehydration, and chronic ethanol use, accounted for the majority of cases identified (105 cases). Of these, 88 were thought to represent antemortem creation, either from diabetes mellitus, starvation/dehydration, chronic ethanol use, or infection, and 17 postmortem creation because of the decomposition process. Overall, the median concentration of isopropanol was 14 mg/dL (range 0–910 mg/dL) in the blood with a median acetone concentration of 25 mg/dL (range 0–195 mg/dL). In the vitreous, the median concentrations of isopropanol and acetone were 12 mg/dL (range 0–90 mg/dL) and 30 mg/dL (range 0–90 mg/dL). The average blood and vitreous isopropanol-to-acetone ratios were 0.69 and 0.56, respectively.

Antemortem Creation

Dehydration and/or Malnutrition—Nine cases were identified where the isopropanol present was believed to be because of a malnourished or dehydrated state. In those cases, the median concentration in blood of isopropanol was 8 mg/dL (range 2–22 mg/dL) with a median acetone concentration of 12 mg/dL (range 6–32 mg/dL). Isopropanol was detected in only one case in the vitreous (concentration 7 mg/dL). Acetone was present in the vitreous in all dehydration/malnutrition cases and had a median concentration of 18 mg/dL (range 7–210 mg/dL). The average ratio of isopropanol to acetone in the blood was 0.84 (range 0.25–2.75). Only one case had a ratio >1, all other cases had ratios <0.9. In the vitreous, the average ratio could not be calculated because of the absence of isopropanol in all but one case. For that one case, the isopropanol-to-acetone ratio was 0.39.

Diabetes Mellitus—Diabetics represented the single largest group of individuals with isopropanol present in this study, with a

TABLE 1—*Summary of blood isopropanol and acetone concentrations by category.*

Category	N	Isopropanol Conc (mg/dL)			Acetone Conc (mg/dL)		
		Median	Min	Max	Median	Min	Max
Creation	105	14	0	910	25	0	195
Decomposition	17	15.5	0	870	24.5	0	46
Dehydration/malnutrition	9	8	2	220	12	6	32
Diabetes mellitus	39	11.5	0	50	46	0	195
Chronic ethanol use	29	15	0	71	19	0	195
Infection	11	70	0	910	58	0	79
Contamination	34	25	0	201	0	0	0
Embalming	15	NA	NA	NA	NA	NA	NA
Tissue procurement	19	25	0	201	0	0	0
Exposure	8	175	148	3700	150	40	200
Unknown	5	14	0	52	0	0	0

TABLE 2—*Summary of vitreous isopropanol and acetone concentrations by category.*

Category	N	Isopropanol Conc (mg/dL)			Acetone Conc (mg/dL)		
		Median	Min	Max	Median	Min	Max
Creation	105	12	0	90	30	0	90
Decomposition	17	23	0	90	28	0	41
Dehydration/malnutrition	9	7	0	7	18	7	210
Diabetes mellitus	39	8	0	50	55	0	112
Chronic ethanol use	29	12	0	81	18	0	23
Infection	11	10	0	30	57	0	89
Contamination	34	19.5	0	92	0	0	0
Embalming	15	20	<1	36	0	0	0
Tissue procurement	19	18	0	92	0	0	0
Exposure	8	240	130	244	159	43	190
Unknown	5	50	0	90	0	0	0

total of 39 cases. The median concentration of isopropanol in the postmortem blood was 11.5 mg/dL (range 0–50 mg/dL) with a median acetone concentration of 46 mg/dL (range 0–93 mg/dL). In the vitreous, the median concentrations of isopropanol and acetone were 8 mg/dL (range 0–50 mg/dL) and 55 mg/dL (range 0–112 mg/dL), respectively. The isopropanol-to-acetone ratio averaged 0.39 (range 0.04–1) in the blood and 0.08 (range 0.02–1.1) in the vitreous.

Chronic Ethanol Use—Chronic ethanol users were the second largest subset of individuals with isopropanol present in the postmortem blood/vitreous, with a total of 29 cases. The median concentration of isopropanol in the postmortem blood was 15 mg/dL (range 0–71 mg/dL) with a median acetone concentration of 19 mg/dL (range 0–195 mg/dL). In the vitreous, the median concentrations of isopropanol and acetone were 12 mg/dL (range 0–81 mg/dL) and 18 mg/dL (range 0–231 mg/dL), respectively. The isopropanol-to-acetone ratio averaged 1.1 (range 0.2–3.3) in the blood and 0.77 (range 0.26–2.75) in the vitreous. Ethanol was present in the postmortem blood in 82% of cases and was present in the vitreous in 82% of cases. The median ethanol concentrations in the blood and vitreous were 180 mg/dL (range 0–401 mg/dL) and 191 mg/dL (range 0–415 mg/dL), respectively.

Infection—An infectious process was thought to be the source of isopropanol in 11 cases. The median blood isopropanol and acetone concentrations were 70 mg/dL (range 0–910 mg/dL) and 58 mg/L (range 0–79 mg/L), respectively. Seventy percent of the cases had isopropanol concentrations of <100 mg/dL in the blood. In the vitreous, the median concentrations of isopropanol and acetone were 10 mg/dL (range 0–30 mg/dL) and 57 mg/dL (0–89 mg/dL), respectively. The isopropanol-to-acetone ratio could only be calculated in two cases each for blood and vitreous as only two cases had acetone present (the same two cases had acetone present in the blood and vitreous). The calculated ratios were 0.14 and 0.81 for the blood and 0.09 for the vitreous.

Postmortem Creation

Isopropanol was thought to be the byproduct of decomposition in 17 of the study cases. The median isopropanol concentration in decomposed bodies was 15.5 mg/dL (range 0–87 mg/L) in the blood and 23 mg/dL (range 0–90 mg/dL) in the vitreous. The median concentrations in blood and vitreous for acetone were 24.5 mg/dL (0–46 mg/dL) and 28 mg/dL (0–41 mg/dL), respectively. For decomposed bodies, ethanol was also present in the blood in 88% of cases and in the vitreous in 91.7% of cases with a median concentration of 64 mg/dL (16–167 mg/dL) and

28 mg/dL (5–200 mg/dL) in the blood and vitreous, respectively. The isopropanol-to-acetone ratio averaged 0.93 in the blood (range 0.4–1.9) and 2.5 in the vitreous (range 0.3–13.6).

Analysis could not be made regarding the time since death for the postmortem creation cases, as it could not be determined for most of the decedents. However, in all of these cases, there were significant decompositional changes present, including skin slippage, bloating, and skin/tissue discoloration.

Unknown—Five cases with an unknown source of isopropanol were found and are summarized in Table 3. Three of the cases were in children, two of which had causes of death which could not be identified and one of which died as a result of trauma. The two additional cases were elderly persons who died of asphyxia by choking and a pulmonary embolus, respectively, but both of whom had cancer of unidentified origins. In three of the five cases, isopropanol was present only in the postmortem blood; in one it was present in both the blood and vitreous; and in the last case isopropanol was only present in the vitreous. Acetone was not present in the blood or vitreous of any of the cases. The median concentration of isopropanol in the postmortem blood was 14 mg/dL (range 0–52 mg/dL), and the two vitreous concentrations were 10 mg/dL and 90 mg/dL (median 50 mg/dL).

Discussion

Isopropanol Creation

In cases where isopropanol was thought to result from antemortem creation, either because of diabetes mellitus, chronic ethanol use, infection, dehydration, and/or malnutrition, the isopropanol concentration tended to be low, usually <100 mg/dL, with the

TABLE 3—*Summary of the cases where the source of isopropanol is unknown.*

Age	Cause of Death	Manner of Death	Blood Isopropanol Conc* (mg/dL)	Vitreous Isopropanol Conc* (mg/dL)
8 weeks	SIDS	Natural	13	0
14 days	Undetermined	Undetermined	52	0
2 years	Subdural hemorrhage	Accident	10	10
85 years	Pulmonary embolus[†]	Natural	15	0
75 years	Choking[†]	Accident	0	90

*Acetone was not present in the blood or vitreous in any of the cases.
[†]Underlying cancer was also present.

exception of infection, where the concentrations tended to be higher. This data is consistent with previous studies. Davis et al. (2) found concentrations ranging from 1 to 29 mg/dL for isopropanol and 7–62 mg/dL for acetone in individuals not exposed to isopropanol but with diabetes mellitus, gastrointestinal disorders, and sepsis. Jenkins et al. (4) examined 162 cases of isopropanol presence because of acetone creation in individuals either in fasting states or with diabetes mellitus and found blood isopropanol concentrations to range from 2–39 mg/dL with acetone levels ranging from 2–87 mg/dL. Jenkins also reported isopropanol and acetone concentrations to range from 2–38 mg/dL (mean 5 mg/dL) and 5–110 mg/dL (mean 40 mg/dL) in the vitreous and 2–38 mg/dL (8 mg/dL) and 9–113 mg/dL (mean 44 mg/dL) in the urine, respectively (4). Lewis et al. (3), in their study of 27 deaths where isopropanol was present because of either diabetes mellitus or liver disease found concentrations of 10–44 mg/dL for isopropanol and 10–56 mg/dL for acetone. Bailey examined five cases of acetonemic patients all of whom had diabetes mellitus and found isopropanol concentrations ranging from 2 to 7.3 mg/dL (one case with 297 to be discussed below) and acetone concentrations of 5.8–30.1 mg/dL (5). In a study by Alexander et al. (6), which looked at individuals with isopropanol present in the postmortem blood, 26 individuals were found who died with isopropanol present but not from isopropanol intoxication. In these cases the isopropanol concentration ranged from trace to 160 mg/dL and the acetone concentration from 20 to 320 mg/dL (6).

In the current study, dehydrated and/or malnourished decedents tended to have the lowest concentrations of isopropanol present, followed closely by diabetics and chronic ethanol users (all averaging around 15 mg/dL). The cases of isopropanol production because of infection tended to be significantly higher (median 174 mg/dL). It could be hypothesized that this may be because of a double mechanism of creation. Individuals with severe infections may have elements of metabolic disturbances, including dehydration, malnutrition, and stress reaction, leading to increased fatty acid metabolism and the creation of acetone, and thus of isopropanol. However, both George et al. and Kugelberg et al. argue that isopropanol can be created by microorganisms (7,8). Specifically, George et al. (7) showed that certain *Clostridium* species produce isopropanol. In one of the infection cases analyzed in this report, *Clostridium novyi* was actually grown on postmortem blood cultures (blood isopropanol 15 mg/dL; no acetone detected). It can be hypothesized that isopropanol can itself be created directly by microorganisms in the cases of an infectious process. This is further supported by the lack of acetone present in the infection cases; acetone was present in only 18.2% of cases.

The isopropanol-to-acetone ratio is helpful in cases where the isopropanol is thought to have been created from acetone. In these cases, the ratio tends to be <1, even in infection cases where the isopropanol concentrations tended to be higher. DiMaio and DiMaio (9), in their *Forensic Pathology* textbook, argue that in disease states creating acetone and isopropanol, there should be high levels of acetone and low levels of isopropanol. Lewis et al. (3) found that when rats were given acetone, isopropanol could be detected in the blood shortly after the exposure and that the isopropanol concentration was consistently less than that of acetone. Jenkins et al. (4) showed the isopropanol-to-acetone ratio in femoral blood to be <1 in 97% of their cases of fasting and diabetic patients. In Bailey's analysis of five cases of people with diabetes mellitus, the isopropanol-to-acetone ratio was found to be <0.5 in all cases except one, where the isopropanol concentration was 297 mg/dL and the ratio was 5.1 (5), raising the possibility of exposure in addition to the diabetes.

In the current study, diabetics had the lowest average isopropanol-to-acetone ratio (average 0.39), followed by infectious causes (average 0.48) and dehydration/malnutrition (average 0.84), although in all cases the average ratio was <1. The average ratio for chronic ethanol users was the highest at just above 1 (average 1.1), although the presence of elevated concentrations of ethanol in these cases may assist in ascertaining the source of the isopropanol.

Decomposition—No previous studies looking at isopropanol in decomposed bodies were found, although numerous studies have examined the creation of ethanol in the decomposition process (8,10,11). In general, isopropanol concentrations in decomposed bodies tended to be low, being slightly greater than the concentrations found in diabetes mellitus, malnutrition/dehydration, and chronic ethanol use, but lower than those found in infectious cases. All decomposition cases examined had both isopropanol and acetone concentrations <100 mg/dL in the blood and <50 mg/dL in the vitreous fluid. The isopropanol-to-acetone ratio was not helpful in decomposed bodies as the ratio had quite a large range for both blood and vitreous (0.3–13.6). However, low concentrations of ethanol were present in approximately 90% of cases, indicating postmortem creation may be possible. As discussed previously, both George et al. and Kugelberg et al. argue that isopropanol can be created by certain microorganisms (7,8), particularly *Clostridium* species (7).

Postmortem Contamination

The issue of isopropanol presence because of postmortem contamination has not been previously discussed in the literature. We found 34 cases in which postmortem contamination was thought to be the cause of the presence of isopropanol. In these cases, the body had either been embalmed or postmortem tissue and/or cornea procurement had occurred. In the embalmed bodies, acetone was not present while methanol usually was, indicating embalming fluid contamination. Embalming fluid usually contains methanol, but often contains isopropanol, formaldehyde, and/or ethanol.

The issue of contamination because of postmortem procurement is not as straightforward. In San Antonio, the tissue procurement agencies will often wash the entire body with isopropanol prior to procurement as well as rinsing off the surface of the eyes. It is hypothesized that the contamination could either be from the puncture of the skin/scleral surface prior to complete drying of the isopropanol, leading to direct contamination of the sample or, alternatively, to the passive absorption into the blood and vitreous through the respective tissues. Further research would need to be conducted to delineate the exact mechanism, although it can be established that contamination results in relatively low concentrations of isopropanol with an absence of acetone, indicating a lack of metabolic processing of the isopropanol.

Exposure

Several articles have been published addressing isopropanol intoxication deaths (4,6,12–14). Alexander et al. looked at 31 deaths due to isopropanol intoxication and found blood isopropanol concentrations ranging from 10 to 250 mg/dL (mean 140 mg/dL) with acetone concentrations ranging from 40 to 300 mg/dL (mean 170 mg/dL) (6). They report that concentrations >150 mg/dL can result in coma, while concentrations >200 mg/dL result in death. The results of the current report are consistent with this study, with a median concentration of isopropanol in the blood of 175 mg/dL and of acetone of 150 mg/dL. Our results are also consistent with Jenkins et al. (4) that, in exposure cases, the blood isopropanol

concentration should be greater than that of acetone, yielding an isopropanol-to-acetone ratio of >1. These results are significantly different from isopropanol results where the isopropanol is either created or artifact.

Unknown

If one applies the general principles discussed previously to the unknown cases described, a hypothesis explaining the isopropanol concentrations can be devised for each case. As none of the cases had acetone present, this would be most consistent with either an infectious process or postmortem artifact. None of the cases were embalmed, and while none of them had tissue harvesting performed prior to autopsy, all of the cases came through a hospital prior to being brought to the BCMEO, creating the possibility of medical contamination. The other possibility would be infection, as 82% of cases of isopropanol production because of an infectious process did not have acetone present. The infectious hypothesis would fit best for the two children dying of undetermined causes, as perhaps the cause of death, and of the isopropanol, was an infectious agent which was not identified. The two elderly cases may represent an infectious process as well, or may represent antemortem creation secondary to a malnurished state because of the underlying cancer. For all of these cases, the concentrations of isopropanol were low and, in addition to the absence of acetone, the possibility of ante-mortem exposure can confidently be excluded.

Conclusion

Determining the cause of isopropanol present in postmortem blood requires a thorough death investigation, including examination of the scene, examination of the body, and review of the pertinent medical history. In addition to a thorough investigation, some general rules can be applied to assist in determining the source of the isopropanol. Overall, cases of isopropanol intoxications tend to have high isopropanol concentrations (100% were >100 mg/dL) while cases of isopropanol creation have low concentrations (96.4% were <100 mg/dL). In addition, specific results can also assist in deciding from where the isopropanol originated. Cases where there is actual isopropanol exposure tended to have isopropanol-to-acetone ratios >1 and cases of antemortem isopropanol production tended to have isopropanol-to-acetone ratios <1. In cases where isopropanol is present because of postmortem contamination (artifactual) either because of embalming or tissue procurement, acetone was absent. In decomposition, blood and vitreous isopropanol-to-acetone ratios were variable, but low concentrations of ethanol and a decomposed state of the body were present. In applying these general principles to cases where isopropanol is detected on postmortem analysis and where the source of the isopropanol is not readily apparent, reasonable explanations of a possible cause may be ascertained.

References

1. Silvilotti MLA. Ethanol, isopropanol, and methanol. In: Dart RC, editor. Medical toxicology, 3rd edn. Philadelphia, PA: Lippincott Williams & Wilkins, 2004;1215–6.
2. Davis PL, Dal Cortivo LA, Maturo J. Endogenous isopropanol: forensic and biochemical implications. J Anal Toxicol 1984;8:209–12.
3. Lewis GD, Laufman AK, McAnalley BH, Garriott JC. Metabolism of acetone to isopropyl alcohol in rats and humans. J Forensic Sci 1984;29(2):541–9.
4. Jenkins AJ, Merrick TC, Oblock JM. Evaluation of isopropanol concentrations in the presence of acetone in postmortem biological fluids. J Anal Toxicol 2008;32:719–20.
5. Bailey DN. Detection of isopropanol in acetonemic patients not exposed to isopropanol. J Toxicol Clin Toxicol 1990;28(4):459–66.
6. Alexander CB, McBay AJ, Hudson RP. Isopropanol and isopropanol deaths—ten years' experience. J Forensic Sci 1982;27(3):541–8.
7. George HA, Johnson JL, Moore WEC, Holdeman LV, Chen JS. Acetone, isopropanol, and butanol production by *Clostridium beijerinckii* (syn. *Clostridium butylicum*) and *Clostridium aurantibutyricum*. Appl Environ Microbiol 1983;45(3):1160–3.
8. Kugelberg FC, Jones AW. Interpreting results of ethanol analysis in postmortem specimens: a review of the literature. Forensic Sci Int 2007;165:10–29.
9. DiMaio VJ, DiMaio D. Interpretive toxicology: drug abuse and drug deaths. Chapter 23. Forensic pathology, 2nd edn. New York, NY: CRC Press, 2001;520.
10. Grellner W, Iffland R. Assessment of postmortem blood alcohol concentrations by ethanol levels measured in fluids from putrefactive blisters. Forensic Sci Int 1997;90:57–63.
11. Honey D, Caylor C, Luthi R, Kerrigan S. Comparative alcohol concentrations in blood and vitreous fluid with illustrative case studies. J Anal Toxicol 2005;29(5):365–9.
12. Parker KM, Lera TA. Acute isopropanol ingestion: pharmacokinetic parameters in the infant. Am J Emerg Med 1992;10:542–4.
13. Jerrard D, Verdile V, Yealy D, Krenzelok E, Menegazzi J. Serum determinations in toxic isopropanol ingestion. Am J Emerg Med 1992;10: 200–2.
14. Adelson L. Fatal intoxication with isopropyl alcohol (rubbing alcohol). Am J Clin Pathol 1962;38:144–51.

Additional information and reprint requests:
Kimberley Molina, M.D.
Bexar County Medical Examiner's Office
7337 Louis Pasteur Drive
San Antonio
TX 78229
E-mail: kmolina@bexar.org

PAPER

PATHOLOGY AND BIOLOGY

Burcu Sabanoğlu,[1] *M.Sc. and Osman Sert,*[1] *Ph.D.*

Determination of Calliphoridae (Diptera) Fauna and Seasonal Distribution on Carrion in Ankara Province*

ABSTRACT: This study was conducted from March 2006 to 2007. The aim of the study was to determine the forensically significant Calliphoridae (Diptera) species and their seasonal distribution in Ankara province at Beytepe Campus. Pig carcasses were killed by a veterinary using pentobarbital sodium with intervals from 2 to 3 weeks to a month. Samples were collected from the carcass everyday in the morning, in the afternoon, and at sunset. Temperature and parameters of weather, such as cloudy, sunny, rainy, and snowy, were recorded at each visit. *Phaenicia sericata* (Meigen), *Chrysomya albiceps* (Wiedemann), *Calliphora vomitoria* (Linnaeus), and *Calliphora vicina* (Robineau-Desvoidy) species, which belong to Calliphoridae family, were collected. It was shown that the seasonal distribution of the collected species was different from each other. The species were determined on carcass between the following times; *P. sericata*, from April to November; *C. albiceps*, from May to November; *Ca. vomitoria*, from February to June and from September to December; and *Ca. vicina*, from February to May and in June and September. Calliphoridae fauna and its annual seasonal distribution in Turkey on 12 pig carcasses were reported for the first time in this study.

KEYWORDS: forensic sciences, forensic entomology, Calliphoridae, Diptera, postmortem interval, Turkey

Forensic entomology is an extensive discipline that uses entomology in the judicial system (1). Forensic entomology, with a focus on justice trial, is separated into three main areas; urban entomology, stored products entomology, and medico-criminal entomology (1).

One of the most crucial questions that comes to mind is the time of death when a corpse is found. Measured body temperature, livor, and rigor mortis analyses can be used to determine postmortem interval (PMI) (2–4) for a few days after death. On the other hand, data obtained from these analyses will no longer be reliable after this few days period. Entomological data, however, can be used as evidence for determining PMI from early stages of decomposition to advance stages (4–7). Kashyap and Pilay (8) made a comparative analysis of various applications that are widely used (autopsy report, entomological, and second degree evidences) in 16 cases. It was recognized that entomological method is statistically more reliable when compared with pathological method that is especially based on differences observed in 72 h after death (livor mortis, rigor mortis, algor mortis, and decomposition) according to the results of this study.

According to Payne (1965), *c.* 400 insect species are found at different decomposition stages on pig carcasses (9). Determination of the insect fauna in a specific area, identification of different immature stages of the insect species (egg, first, second, third instar larvae, and pupa) that appear on carrion, and the knowledge of the development time of these insect species at different temperatures can provide reliable evidence for the determination of PMI (4,6,10). Questions such as time of death, season of death, whether the corpse is moved away from the crime scene to another place after death, geographic location of death, location of specific sites of trauma that occurred on the body, sexual molestation, and use of drugs can be answered following the determination and examination of entomological evidences (11).

The aim of this study was to determine Calliphoridae (Diptera) species on pig carcass and identify seasonal distribution of these species in Ankara province at Beytepe Campus.

Materials and Methods

Pig (*Sus scrofa* L.) carcasses were provided from Baskent University Generation and Research Center in Ankara. Each pig was killed by a veterinary employed at the center using pentobarbital sodium with intervals from 2 to 3 weeks to a month. This drug is known to have impacts on the respiration system of pigs (12). Pigs were washed and transferred to the study area following their death.

Two forestation areas at Beytepe Campus, Hacettepe University, were chosen as the study areas. Eight and four pigs were placed, respectively, in the first area and second area. The flora of the study site consisted mainly of dark pine (*Pinus nigra* L.). Distance between the carcasses in dark pine forest was far enough to prevent them affecting each other.

The study was separated into four experimental periods as spring, summer, autumn, and winter to facilitate the examination. The placing dates of pig carcasses in these experimental periods and decomposition stages of pig carcasses are given in Table 1 and Fig. 1, respectively.

A pig carcass is an attractive food source for many big vertebrate scavengers (4). Therefore, carcasses were placed inside $1.0 \times 1.0 \times 1.5$ m sized wire cages.

A "Forensic Entomology Daily Data Sheet" was prepared to provide a standardized study. Temperature and parameters of weather, such as cloudy, sunny, rainy, and snowy, were recorded on this data sheet at each visit. Additionally, Calliphoridae species seen on carcass, number of organisms belonging to these species, and their locations on the carcass as well as decomposition stage of the carcass and its physical changes were recorded in the sheet. Ambient temperature was measured with an electronic thermometer and hygrometer; soil temperature was measured with mini thermometer with 15-cm probe; and the temperature 1.5 m over carcass was measured with alarm thermometer with external probe.

Samples were collected from the carcass every day in the morning, in the afternoon, and at sunset. Attention was paid to collect samples in the afternoon when fly activity was high. The adult samples on carrion were collected using net, whereas the egg, larvae, and pupae samples were collected using forceps. A few of the adult

samples were killed using ethyl acetate to keep them for collection. Egg, larvae samples, pupae, and adults were preserved in 70% ethanol. Each sample box was labeled accordingly. The Calliphoridae samples brought to the laboratory were identified using adult and third larval instar identification keys in Greenberg and Kunich's (13) book, *Entomology and Law: Flies as Forensic Indicator*.

Results

In this study, four species from three subfamilies of Calliphoridae family were identified: *Calliphora vicina* and *Calliphora vomitoria* belonging to Calliphorinae subfamily, *Phaenicia sericata* belonging to Lucilinae subfamily, and *Chrysomya albiceps* belonging to Chrysominae subfamily were identified in Ankara province at Beytepe Campus between the time period of March 2006 and 2007. The species, identified during the experiment period, were determined to be coming to the carcasses in 2–3 min following the placement in the area, depending on the ambient temperature. Maximum and minimum air temperatures, in which the species were seen, are given in Table 2. The Calliphoridae species laid eggs

TABLE 1—*The placing dates of pig carcasses in each experimental period.*

Season	Pig Number	Date
Spring	1	21 March
	2	21 April
	3	25 May
Summer	4	22 June
	5	12 July
	6	4 August
Autumn	7	29 August
	8	4 September
	9	10 November
Winter	10	6 December
	11	30 January
	12	15 February

TABLE 2—*Maximum and minimum temperatures in which the species were seen.*

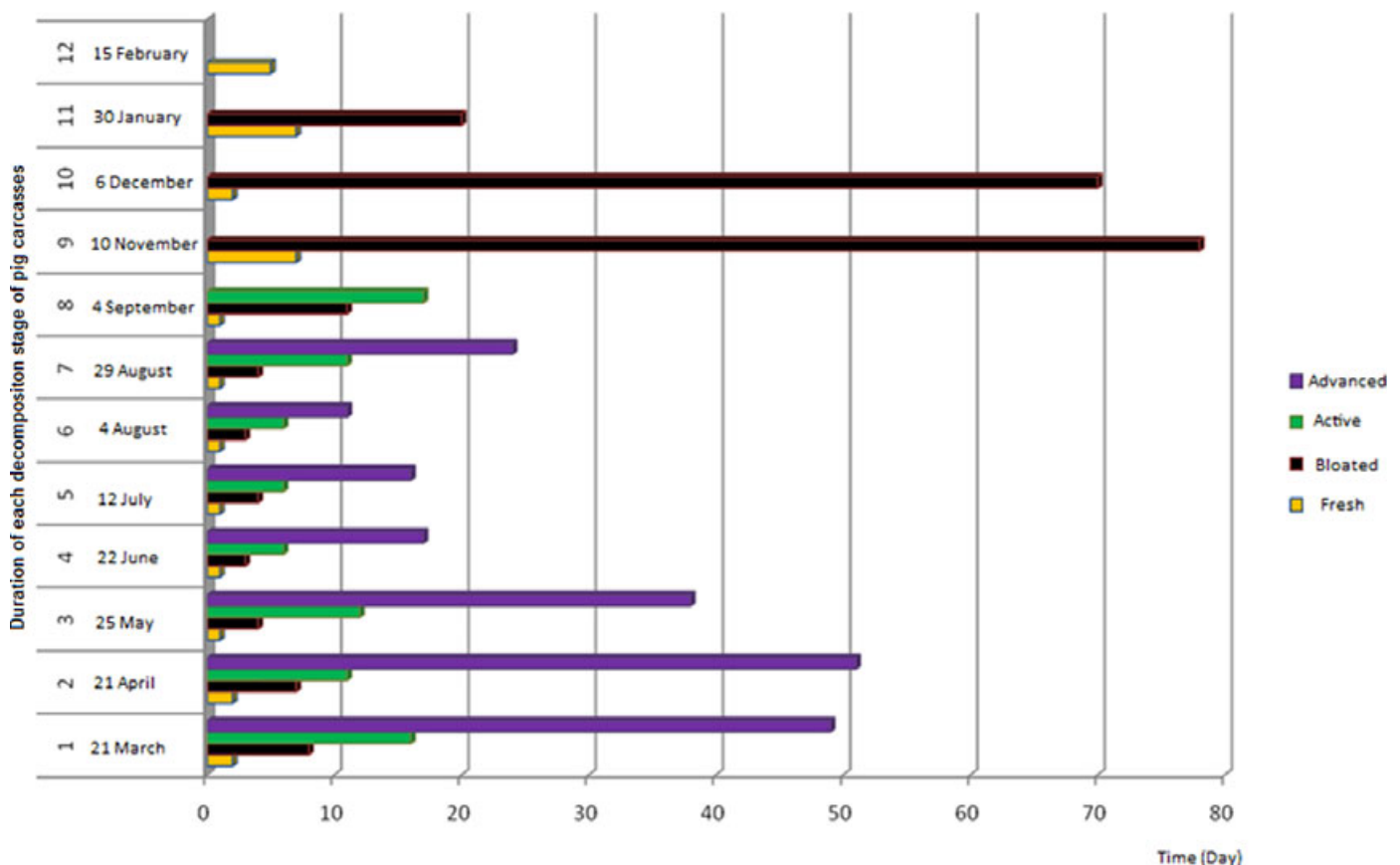| Species	First Seen	Last Seen	Min. Temperature (°C)	Max. Temperature (°C)
Calliphora vomitoria	21 March	March 2008	6.5	35
Calliphora vicina	21 March	March 2008	5	35
Phaenicia sericata	1 April	9 November	8	41
Chrysomya albiceps	26 May	9 November	11	41

FIG. 1—*Duration of each decomposition stage of pig carcasses.*

TABLE 3—*Monthly mean moisture, air, over carcass, on carcass, ground, soil temperature from 5, 10, and 15 cm.*

Months	Temperatures							
	Monthly Mean Moisture (%)	Monthly Mean Air Temp. (°C)	Monthly Mean Over Carcass Temp. (°C)	Monthly Mean On Carcass Temp. (°C)	Monthly Mean Ground Temp. (°C)	Monthly Mean Soil Temp. 5 cm (°C)	Monthly Mean Soil Temp. 10 cm (°C)	Monthly Mean Soil Temp. 15 cm (°C)
March	33.4	18.9	18.2	38.2	19.6	14.4	11.7	10.1
April	36.5	20.5	21.7	36	24.2	18.8	16.1	14.6
May	32.3	26.3	27.8	32.3	31.3	22.8	20.1	18.1
June	31.4	30	31.8	36.2	36.6	28.4	25.6	23.5
July	29.6	30.2	31.8	33.8	34.5	27.8	25.3	23.7
August	19.2	33.2	35.4	37	39.3	30	26.5	25
September	32.8	24.3	24.6	29	27.5	20.6	18.5	17.8
October	41.5	19.6	18.8	21.2	22	14.8	13.1	12.5
November	46.2	10.7	8.3	7.2	9.2	6.6	5.5	4.7
December	43.8	6.3	3.6	2.4	2.6	1.2	1.8	2.5
January	49.4	4.2	2.3	0.4	−0.6	−0.7	0	0.5
February	42.1	7.9	6.4	4.8	4.2	2.9	2.3	1.7

45 min following the placement of the carcass in the study area during summer months, when the ambient temperature was the highest during the experiment (Table 3). In spring and autumn, when the ambient temperature was lower (Table 3), this interval was observed to be 1.5 h. It was observed that *Ca. vomitoria* laid eggs 7 days after the pig carcass was brought to the study area at the end of November; when the ambient temperature fell to 4–5°C.

Rainfall condition and cloudiness influenced the behavior of Calliphoridae species during the experiment. They did not lay eggs in rainy days, and this result was parallel to the findings of Introna et al. (10). Another factor that prevented the Calliphoridae species from laying eggs was the presence of *Vespula germanica* (Fabricius, 1793) from the order Vespidae (Hymenoptera). The number of individuals from these species increased in July and August, thus preventing Calliphoridae species from laying eggs. These species had caused a decline in Calliphoridae eggs by disengaging their larvae and adults and had also effected on the localities where flies laid eggs on the carcass. The Calliphoridae species laid eggs under the carcass instead of the natural body apertures, which they generally prefer, to ensure their egg's safety. Additionally, Vespidae (*V. germanica* Fabricius, 1793), individuals belonging to Silphidae (*Thanatophilus ferrugatus* Solsky, 1874) and Staphylinidae (*Creophilus maxillosus* L., 1758) families (Coleoptera) also consumed Calliphoridae larvae on the carcasses.

The spring experiments were carried out between the dates of March 21 and June 21, 2006; and *Ca. vicina*, *Ca. vomitoria*, and *P. sericata* species were identified (Figs. 2 and 3).

The summer experiments were carried out between the dates of June 22 and August 29, 2006. The time intervals between the placements of the pig carcasses in the study area were shortened because of the increase in ambient temperature during summer. Therefore, the decomposition rate of pig carcasses increased, whereas affinity of Calliphoridae species to the pig carcasses decreased rapidly. During the summer experiments, *Ca. vicina*, *Ca. vomitoria*, *P. sericata*, and *C. albiceps* species were seen on pig carcasses (Figs. 2 and 3). On the other hand, *Ca. vomitoria* and *Ca. vicina* species (the ones which prefer colder temperatures) were not seen on pig carcasses, following the last week of June, in the entire summer period.

The autumn experiments started on August 29, 2006 and continued until November 10, 2006. Four species from Calliphoridae family were identified on the pig carcasses; *Ca. vomitoria*, *Ca. vicina*, *P. sericata*, and *C. albiceps* (Figs. 2 and 3) within this period. *Phaenicia sericata* and *C. albiceps* species were not observed on the pig carcasses after the second week of November.

The winter experiments started on December 6, 2006 and continued until March 2007. No Calliphoridae samples and even *Ca. vicina* and *Ca. vomitoria* species, which prefer low temperatures, were observed on carcasses from the beginning of the first week of December. *Calliphora vomitoria* and *Ca. vicina* adults were observed on the carcasses again following the beginning of the second week of February, when the ambient temperature was suitable for insect development (Figs. 2 and 3).

Discussion

Phaenicia sericata was the species that had the longest duration on the pig carcasses in this study. This species was present on the pig carcasses and laid eggs from April 1, 2006 until November 10, 2006. Maximum individual number of *P. sericata* (95 individuals)
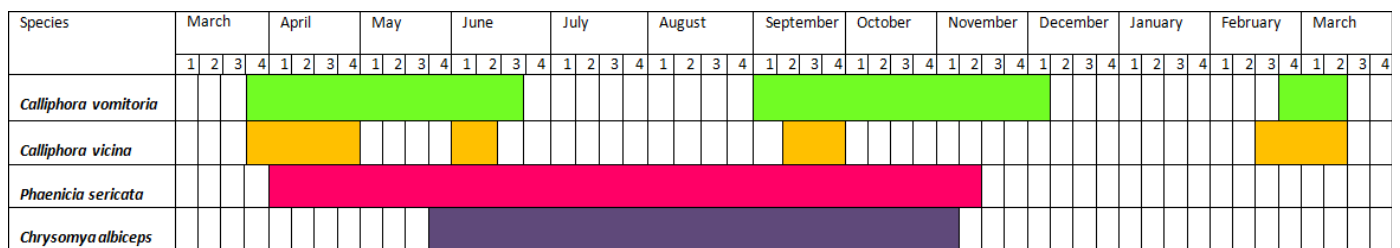

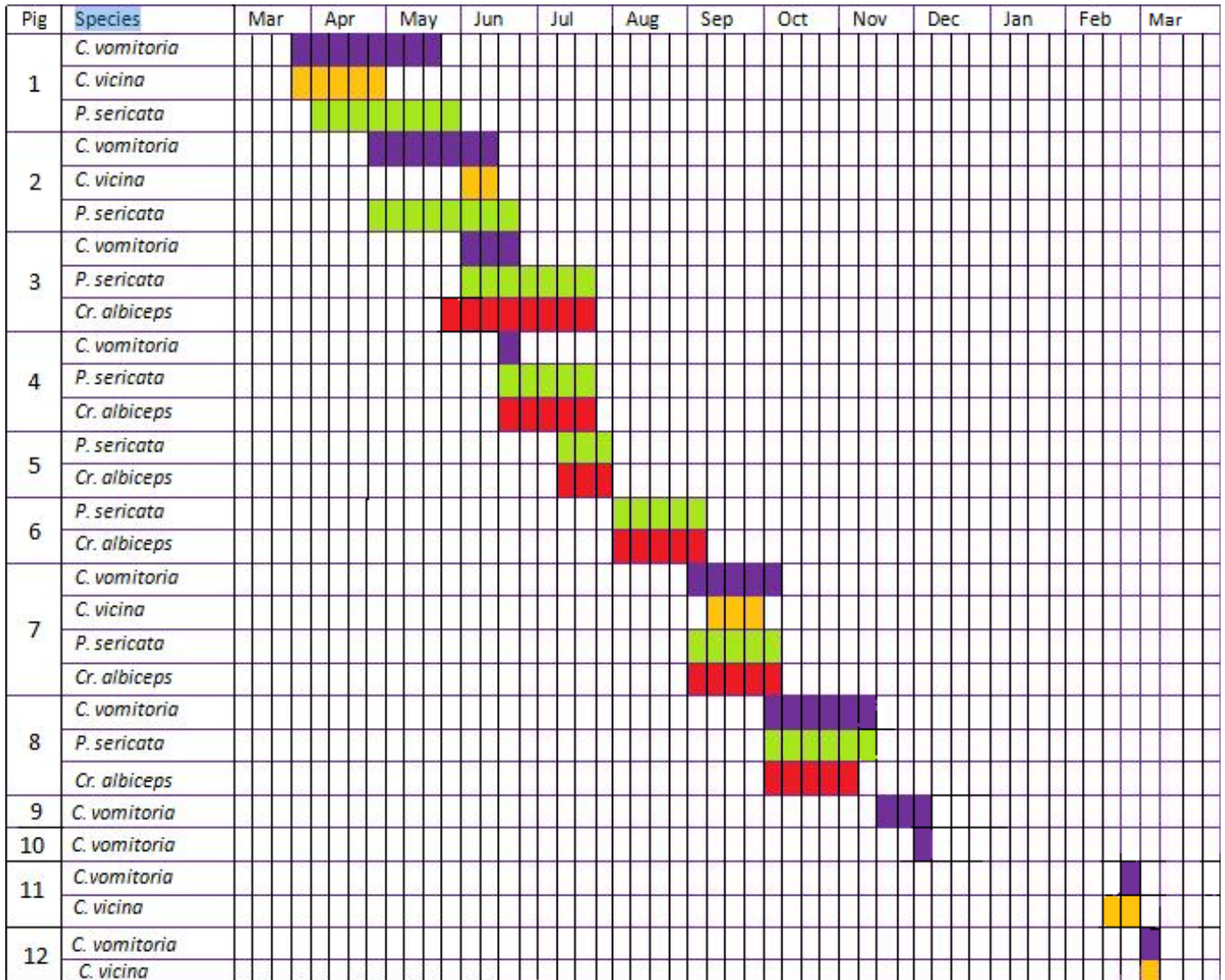FIG. 2—*Distribution of identified species for 1 year period.*

Pig	Species	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
1	C. vomitoria													
	C. vicina													
	P. sericata													
2	C. vomitoria													
	C. vicina													
	P. sericata													
3	C. vomitoria													
	P. sericata													
	Cr. albiceps													
4	C. vomitoria													
	P. sericata													
	Cr. albiceps													
5	P. sericata													
	Cr. albiceps													
6	P. sericata													
	Cr. albiceps													
7	C. vomitoria													
	C. vicina													
	P. sericata													
	Cr. albiceps													
8	C. vomitoria													
	P. sericata													
	Cr. albiceps													
9	C. vomitoria													
10	C. vomitoria													
11	C.vomitoria													
	C. vicina													
12	C. vomitoria													
	C. vicina													

FIG. 3—*Distribution of species for each pig carcass.*

was observed when adults on pig carcasses were counted at the end of April 2006.

Calliphora vomitoria was another species that was seen for a longer period on the pig carcasses. Grassberger and Frank (14) determined that *Ca. vomitoria*, along with *Phaenicia* spp., was one of the dominant species in a study conducted in Vienna during springtime. Haskell et al. (11) stated that *Calliphora* spp. (*Ca. vomitoria*, *Ca. vicina*, and *Calliphora livida*) were densely observed on carcasses in spring. According to their study, their intensity declined because of the increase in ambient temperature and did not appear on pig carcasses in summer until autumn when the ambient temperature decreased again. Likewise, *Ca. vomitoria* was observed from the first day of the study (March 21) until June 23, 2006 on pig carcasses in Ankara province at Beytepe Campus in this study. The species was not observed from June 23 to August 29, and then it was seen again on August 30 with the decrease in ambient temperature and continued to be seen on carcasses until December 10. *Calliphora vomitoria* appeared again after February 22 with the increase in the ambient temperature and the melt down of snow on carcass. Based on the adult count on the pig carcasses, *Ca. vomitoria* was the dominant species in March, April, October,

and November and the only species to be seen on pig carcass on December. *Calliphora vomitoria* is found in rural areas abundantly (15). The campus is considered to be representing both an urban area in this study because of the settlements (dormitories, cafeterias, and departments) and the rapid conversion of natural areas into settlements, and a rural area because of its rich woodlands.

Greenberg and Povolny (16) reported that *Ca. vicina* existed in subtropics during winter and in temperate zone during spring and autumn. Introna et al. (10) identified *Ca. vicina* samples especially during autumn time in their study conducted in Maryland. Tantawi et al. (17) declared that *Ca. vicina* was observed only during winter on carcass in their Alexandria study. Within our study, however, *Ca. vicina* was observed in March, April, September, and February but not observed in May, June, October, and December; when *Ca. vomitoria* was observed. It was also identified that the number of individuals of *Ca. vicina* was fewer than other samples collected on carcasses.

Chrysomya albiceps was the most frequently seen species during summer on the carcasses. The species was first seen on May 26 and existed until November 6 when the fourth pig carcass was placed in the study area. *Chrysomya albiceps* was seen along with

P. sericata and *Ca. vomitoria* on pig carcasses, which were placed in the area in May, June, September, and October; whereas it was seen with only *P. sericata* in July and August. Similar to this study, Tantawi et al. (17) reported that *C. albiceps* was seen in spring, summer, and autumn, while it was not observed in winter on carcasses in their study that was held in Alexandria. Smit (18) also found that *C. albiceps* females did not lay egg in winter. It was observed that only *C. albiceps* larvae pupate under and near the pig carcasses when compared with postfeeding larvae of other Calliphoridae species identified within the study. This finding is parallel to the findings of the study that was carried out by Tantawi et al. (17) in Alexandria.

Conclusion

An accurate PMI identifying process is the most basic criteria in forensic entomology. It may be possible to find out the criminal or to decrease the number of the suspects in a death case (19) only with a true estimation. In forensic entomology, a definite PMI estimation is provided with a precise analysis and assessment of entomological data (20).

In conclusion, four Calliphoridae species, which are used for PMI estimation in any death case, were identified in this study. These four species were seen together in May, June, September, and October. It was determined that adults and larvae of *Ca. vicina* and *Ca. vomitoria* can be used in any case in spring and autumn when ambient temperature is low; and *C. albiceps* can be used to estimate PMI in a death case, which took place in summer. Another identified species was *P. sericata* within the study. The species was seen for the first time in April, and it existed until mid November in this study. It is therefore also reliable to use this species in a probable death case that occurs in spring, summer, or autumn.

References

1. Hall RD. Introduction: perceptions and status of forensic entomology. Forensic entomology: the utility of arthropods in legal investigations. Boca Raton, FL: CRC Press, 2001;1–15.
2. Bass WM. Preface. In: Byrd JH, Castner JL, editors. Forensic entomology: the utility of arthropods in legal investigations. Boca Raton, FL: CRC Press, 2001;ix–x.
3. Amendt JR, Krettek R, Zehner R. Forensic entomology. Naturwissenschaften: Springer, 2004;91(22):51–65.
4. Smith KGV. A manual of forensic entomology. London: Trustees of the British Museum (Natural History), 1986.
5. Goff ML, Omori AI, Gunatilake K. Estimation of postmortem interval by arthropod succession. Am J Forensic Med Pathol 1988;9(3):220–5.
6. Greenberg B. Flies as forensic indicators. J Med Entomol 1991;28(5):565–77.
7. Nuorteva P. Sarcosaphrophagous insects as forensic indicators. In: Tedeschi CG, Eckert WG, Tedeshi LG, editors. Forensic medicine: a study in trauma and environmental hazards. Toronto: W. B. Saunders and Company, 1977;3:1072–95.
8. Kashyap VK, Pilay VV. Efficacy of entomological method in estimation of postmortem interval: a comparative analysis. Forensic Sci Int 1989;40:245–50.
9. Payne JA. A summer carrion study of the baby pig *Sus scrofa* Linnaeus. Ecology 1965;46(5):92–602.
10. Introna F, Suman TW, Smialek JE. Sarcosaprophagous fly activity in Maryland (USA). J Forensic Sci 1991;36(1):238–43.
11. Haskell NH, Hall RD, Cervenka VJ, Clark MA. On the body: insect's life stage presence and their postmortem artifacts. In: Haglund WD, Sorg MH, editors. Forensic taphonomy. Boca Raton, FL: CRC Press, 1997;415–48.
12. http://www.elephantcare.org/Drugs/pentobar.htm.
13. Greenberg B, Kunich JC. Entomology and the law: flies as forensic indicators. Cambridge: Cambridge University Press, 2002.
14. Grassberger M, Frank C. Initial study of arthropod succession on pig carrion in a central European urban habitat. J Med Entomol 2004;41(3):511–23.
15. Anderson GS. Forensic entomology in British Columbia: a brief history. J Entomol. Soc. BC 2001;98:127–35.
16. Greenberg B, Povolny D. Flies and disease, Vol. I. Princeton: Princeton University Press, 1971;57.
17. Tantawi TI, El-Kady EM, Greenberg B, El-Ghaffar HA. Arthropod succession on exposed rabbit carrion in Alexandria, Egypt. J Med Entomol 1996;33:566–80.
18. Smit B. A study of the sheep blowflies of South Africa. Onderstepoort: Report of the Director of Veterinary Service, 1931;17:299.
19. Catts EP, Haskell NH. Entomology & death—a procedural guide. Clemson, SC: Joyce's Print Shop, Inc., 1990.
20. Schoenly K, Reid W. Dynamics of heterotrophic succession in carrion arthropod assemblages: discrete series or a continuum of change. Oecologia 1987;73:192–202.

Additional information and reprint requests:
Burcu Sabanoğlu, M.Sc.
Research Assistant, Faculty of Science
Department of Biology
Applied Biology Section
Hacettepe University
Ankara, 06532
Turkey
E-mail: burcus82@hacettepe.edu.tr

PAPER

PSYCHIATRY & BEHAVIORAL SCIENCES

*Katja Björklund,[1,2] Ph.D.; Helinä Häkkänen-Nyholm,[1,3] Ph.D.; Lorraine Sheridan,[4] Ph.D.;
Karl Roberts,[5] Ph.D.; and Asko Tolvanen,[6] Ph.D.*

Latent Profile Approach to Duration of Stalking

ABSTRACT: Stalking behavior and victim–stalker relationship are often the principal known factors in a stalking case. Thus, they are of great importance when trying to identify factors contributing to stalking duration. The present study aims to identify distinct subgroups of stalking victims based on measures of behavioral stalking dimensions. These victim subgroups, stalking dimensions, and victim–stalker relationship are examined in relation to stalking duration. Using a sample of 137 university students, latent profile analysis (LPA) revealed five distinct victim subgroups based on stalker behavior dimensions: surveillance, low-profile, social lurker, wide scope, and baseline stalkers. The subgroups were significantly related to stalking duration and explained a considerable amount of the variance along with the stalking dimensions and victim–stalker relationship. Connections to stalking literature and utility of person-orientated methods in stalking research are discussed.

KEYWORDS: forensic science, forensic psychology, stalking, duration of stalking, latent profile analysis, university students

Stalking has become the label for a long-term pattern of unwanted behaviors involving the persistent pursuit and repeated intrusions directed by one person toward another (1–5). It has a considerable negative impact on the victim's social, psychological, and physical well-being (6–15). According to previous research, most stalking cases persist for a year or less (16,17). However, the average stalking duration is almost 2 years (15,17,18), while the mean duration of stalking among college students is *c.* 150 days, i.e., *c.* 5 months (19,20). Recent research suggests that the longer stalking continues, the greater the potential damage to the victim (14,21). It appears that the persistence of stalking along with the unpredictable nature of stalking could be the main source of harm to a considerable proportion of victims (21). Thus, being able to analyze stalker behaviors and estimate the duration of these behavioral patterns are of value to both health care and law enforcement professionals (21,22). Currently, prior victim–stalker relationship is the best predictor of stalking duration (21,23). However, there is a paucity of information available regarding predictors of stalking duration, and as such better predictive models are needed (21). The present study aims to address some of these issues in a sample of university students by analyzing behavioral stalking dimensions, victim subgroups, and victim–stalker relationship in relation to stalking duration.

Several studies have analyzed stalking in university students by thematically clustering various stalking behaviors (20,24–29). However, only one of these studies (29) analyzed stalking behaviors in

relation to duration of stalking. For example, Cupach and Spitzberg (24) identified four factors: pursuit, violation, threat, and hyper-intimacy, while Spitzberg et al. (29) found two factors: pursuit and aggression, which were both significantly related to stalking duration. Haugaard and Seri (26) also reported a two-factor solution, one factor characterized by stalking behaviors involving personal contact and the other factor involving noncontact stalking behavior. Dennison and Stewart (25) found four categories of stalking behavior: direct communication, covert pursuit, self-harm, and other harm. Also, Sinclair and Frieze (27,28) found four subcategories based on factor analysis: approach (normal courtship approach behavior), surveillance, intimidation, and verbal and physical aggression.

The duration of stalking has been shown to vary according to the nature of the prior relationship, with ex-intimate partners being the most persistent stalkers and strangers stalking for the shortest period (2,30). Furthermore, studies have indicated that violence is not uncommon in stalking cases, especially when the stalker is an ex-intimate partner of the victim (31–34), although rates vary considerably across studies. Thus, one of the major causes of concern over stalking is the fear that relatively mild forms of harassment might escalate into potentially dangerous behavior (22,23,33,35). In an extensive review, the average rate for physical violence was 32% and for sexual violence 12% (18). Relatively lower levels of threats and physical violence have, however, been reported among student populations (6,19). In Cupach and Spitzberg (24), 20–30% of the students across two samples reported being subjected to the more threatening stalking behaviors (e.g., threatening notes and property damage). Roberts (4) on the other hand found much higher rates of threats (63%) and stalking violence (36%). Logan et al. (36) found that stalking victimization was significantly associated with physical and psychological abuse for women and with psychological abuse for men. Haugaard and Seri (26) found that 22% of female and 7% of male student victims feared for their safety, and this was related to the duration of stalking.

[1]Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland.
[2]Finnish Student Health Service, Helsinki, Finland.
[3]Forensic Laboratory, National Bureau of Investigation, Vantaa, Finland.
[4]School of Life Sciences, Heriot Watt University, Edinburgh, UK.
[5]Australian Graduate School of Policing, Charles Sturt University, Manly, Australia.
[6]Department of Psychology, University of Jyväskylä, Jyväskylä, Finland.

Considering the wealth of research on stalking that now exists, it is surprising that the duration of stalking has been relatively unexplored both theoretically as a concept and empirically (37,38). As Purcell et al. (14,38) have earlier noted, there have been no attempts to empirically establish whether there exists a discrete disjunction between brief, self-limiting episodes of stalking and the type of persistent stalking that puts the victim at risk of psychological and physical damage. They used their data (i.e., median split) and ROC analysis to distinguish brief outbursts of intrusiveness (under 2 weeks) from damaging and persistent episodes (over 2 weeks) of stalking. A significant proportion of the victims had experienced a brief burst of harassment by a stranger stalker. The results, however, indicated that the continuation of stalking beyond a threshold of 2 weeks is associated with a more intrusive, threatening and psychologically damaging course of stalking (14,38). In contrast to these findings of Purcell and colleagues, James and Farnham (31) recently found serious violence to be associated with a shorter duration, leading to an observation that different kinds of violence have different kinds of associations and that risk factors may vary accordingly.

The present study has three main goals: First, it was hypothesized that in line with previous studies (26,29), the stalking behaviors could be clustered into at least two somewhat separate dimensions with a focus either on pursuit or aggression. The second aim was to identify distinct subgroups of stalking victims based on measures of the stalking dimensions. The third aim was to also investigate how the duration of stalking is associated with the stalking dimensions, the victim subgroups, and the nature of the prior victim–stalker relationship.

Materials and Methods

Materials

A questionnaire presented in Sheridan et al. (39) was used, with minor changes. This instrument has been tried and tested in several studies (39–42). As a whole, the questionnaire measures participants' perceptions of what constitutes stalking and their personal experiences of stalking. A working definition of stalking was provided on the questionnaire as "Persistent unwanted behavior consisting of several attempts to approach, contact, or communicate that the recipient didn't want and did not encourage." The present study relates to the section of the questionnaire where the respondents were asked to choose, from a list of 47 behaviors, those that they had personally been subjected to. In addition, respondents were asked to provide information concerning the characteristics of the stalker, the nature of their prior relationship with the stalker, and the duration of the stalking episode. For data analysis, stalkers were categorized into strangers, ex-intimates, and acquaintances, in line with previous research (7,10,43,44).

Sample

The sample comprised university students from five large universities across the country studying either within the Faculty of Arts or departments of psychology. Students from the Faculty of Arts were chosen because this faculty represents one of the largest and most heterogeneous faculties. Psychology students were chosen as the stalking questionnaire used in the present study is planned to be tested among psychology students across 22 countries. Participation was voluntary for all respondents. The electronic questionnaire and the cover letter were distributed by e-mail-lists provided by student associations. In Finland, university students automatically belong to an association of their own and distributing information (concerning activities, etc.) through the e-mail-list is a usual practice. The study was approved by the ethical board of the department of psychology and the information technology department of the University of Helsinki.

Altogether 615 students responded to the electronic questionnaire. To analyze individual stalking behaviors, only those students who had been subjected to one stalking episode during their lifetime were included ($n = 137$, 22.3% of the whole sample) in the present study for further analysis. Of these, 51.5% were psychology students. Of the 137 respondents, 91.2% were female, which is a higher proportion than the general proportion of female students at the Faculty of Arts (76.5%) and department of psychology (80.1%). The average age of the respondents was 24 (SD = 3.24, range 18–37 years), which is slightly lower than the average age of all Finnish university students ($M = 29.2$ years including postgraduate students).

Statistical Analysis

The statistical analysis was conducted using SPSS for Windows, version 15, and Mplus, version 4.2 (45). *In the first stage of the analysis,* the goal was to find the optimal number of stalking behavior dimensions (the first research question). All stalking behavior variables were entered to exploratory factor analysis using weighted least squares mean and variance adjusted (WLSMV) estimator to uncover the underlying structure (dimensions). The exploratory process using WLSMV estimation was continued by setting the remaining variables to load on individual factors. The final factor solution (consisting of the majority of the original variables) was estimated using the maximum likelihood (ML) method to obtain the most reliable factor score estimates. These factor score estimates were saved as variables and used in the second stage of the analysis.

In the second stage of the analysis, the aim was to find latent subgroups where the profiles differ from each other regarding stalking dimensions (the second research question). This was done via mixture analysis implemented in the Mplus program (45). This type of mixture model is referred as "latent profile analysis" (LPA) in the literature. This modeling is based on the idea that the observed data can represent subpopulations, i.e., latent classes, and that these classes can be identified and their parameters estimated (46,47). The advantage of the mixture model (unlike the traditional cluster analyzes) is that it is based on and permits the use of statistical criteria for deciding the number of latent classes. These statistical criteria are Akaike's information criteria (AIC) (48), Baye's information criteria (BIC) (49), adjusted Baye's information criteria (aBIC) (50), the adjusted Vuong, Lo, Mendell and Rubin (LMR) (51) test, and the parametric bootstrapped likelihood ratio test (BLRT) (45,52). When comparing fit statistics for latent models, lower values of AIC, BIC, and aBIC indicate a better model fit. The LMR and BLRT tests are based on hypotheses of successive numbers of latent classes (e.q. k or $k-1$ number of classes).

The subgroups discovered in the second stage were used *in the third stage of the analysis.* In line with the third research question, the duration of stalking (measured in days) was predicted by entering the subgroup variable, group-centred stalking dimension scores, victim–stalker relationship variable, along with some other background variables into a general linear model (GLM) in SPSS. The model also examined the interaction effect of the group and stalking dimension scores and victim–stalker relationship. Statistically nonsignificant effects were excluded from the final model. The

factor scores used in the GLM were group mean centered, so that the relation between duration and stalking scores would be independent from the differences between the latent class means. Because the duration of stalking variable was very skewed, the natural logarithmic transformation was used that successfully normalized the distribution.

Results

Sample Characteristics

Most of the stalkers were male (91.0%). The average age of the stalkers as reported by the respondents was 25.5 (SD = 8.48, range 10–56 years). Of the stalkers, 26.3% were 18 years or younger at the time of the stalking, 59.1% were 19–30 years old, and 13.9% were over 30 years. In total, 13.1% of the stalkers were strangers, 22.6% ex-partners, and 64.2% acquaintances (e.g., fellow students, coworkers, clients).

Duration of Stalking

The mean duration of stalking was 296.1 days (SD = 465.46, median = 90, range 1–2555 days). For 25.8% of the respondents, the stalking had lasted 30 days or less, for 26.5% 31–90 days, for 17.4% 91–180 days, and for 30.3% 181 days or more. The duration of stalking was not significantly associated with either stalker or victim gender. However, the duration had a significant negative correlation with stalker age (Pearson $r = -0.18$, $p < 0.05$). Of the stalkers aged 18 or younger, 40% stalked for more than a half a year, while the corresponding figures for the 19–30-year-old stalkers and older than 30-year stalkers were 35.4% and 10.0%, respectively. Univariate ANOVA indicated victim–stalker group differences for the duration of stalking ($F_{2,129} = 9.81$, $p < 0.001$). The average duration was significantly shorter among strangers

(43.6 days), than among acquaintances (326.5 days) and ex-partners (362.5 days). For 55.6% of the stranger stalkers, the duration was less than 30 days, while the corresponding figure for acquaintances was 23.8% and ex-partners 13.3%. None of the stalking by strangers persisted for more than half a year (c. 35% of the stalking by acquaintances and ex-partners lasted for more than 180 days).

Defining the Dimensions of Stalking Behavior

All 47 stalking items were entered into the explorative factor model. After the explorative process, the final three-factor model was estimated and each variable was set to load on one of the three factors. The model fitted the data well ($\chi^2(39) = 58.54$, $p = 0.02$, RMSEA = 0.06) and no high modification indexes existed, indicating a clear correspondence between variables and factors. To maximize the reliability of factor scores, the final model was estimated using the ML method. Table 1 shows the factor solution and standardized loadings of stalking behavior variables. This factor solution could be given an interpretation of three distinct dimensions of stalking behavior: violence, surveillance, and contact seeking. The violent dimension includes stalking behaviors comprising intrusive, threatening, and/or violent behaviors. The surveillance dimension includes behaviors that can be characterized as "classic" stalking behavior; including following and loitering near locations where the stalker might see or run into the victim. The third dimension was named contact seeking as it includes stalking behaviors indicative of attempts to learn more about the target of interest (victim), to get to know him/her, and possibly develop a relationship.

In the next stage, the factor scores of the three stalking dimensions for the whole sample were estimated and used as variables in LPA. According to information criteria, AIC, BIC, and aBIC, the five-class solution fits the data best. Also, the p-values of LMR and BLRT tests indicate that the five-class solution (alternative

TABLE 1—*A three-factor solution for stalking with standardized factor loadings for stalking behavior variables.*

Dimension 1 Violence	Dimension 2 Surveillance	Dimension 3 Contact Seeking
Death threats	Driving by home, school, or work	Asking for a date repeatedly
0.908	0.805	0.786
Physical harm	Spying	Asking out "as just friends"
0.871	0.786	0.736
Verbal abuse	Loitering by victim's school or work	Offering to buy a drink in a café or bar
0.871	0.774	0.711
Hurting feelings	Changing school etc. to be near the victim	Engaging in a conversation in a public place
0.826	0.756	0.709
Physical threats	Loitering by victim's home	Phoning after one initial meeting
0.789	0.687	0.663
Property damage	Asking others about the victim	Trying to manipulate/force into dating
0.787	0.675	0.615
Forced sexual contact	Seeing a person at the same time each day	Multiple phone calls
0.733	0.668	0.611
Suicide threats	Visiting places frequented by the victim	Sending or giving gifts
0.698	0.662	0.521
Aggressive behavior when seeing with others	Following	Trespassing on victim's property
0.681	0.640	0.486
Intercepting mail	Trying to get to know victim's friends	
0.660	0.624	
Confining against victim's will	Finding information (phone, address, etc.)	
0.637	0.521	
Secretly taking belongings		
0.614		
Refusing to accept that a relationship is over		
0.559		
Manipulating into dating		
0.482		

Only the correlation between "surveillance" and "contact-seeking" factors was statistically significant ($r = 0.42$).

TABLE 2—*Fit statistics for latent classes (1–7).*

Classes	AIC	BIC	aBIC	LMR	BLRT
1	972.496	998.441	969.974	–	–
2	951.458	988.935	947.815	0.0536	0.0001
3	936.028	985.036	931.265	0.0223	0.0001
4	923.678	987.217	920.793	0.3905	0.0128
5	881.855	953.925	874.850	0.0066	0.0001
6	883.988	967.590	875.862	0.4794	0.1935
7	886.418	981.550	877.171	0.1628	0.3333

AIC, Akaike's information criteria; BIC, Baye's information criteria; aBIC, adjusted Baye's information criteria; LMR, Lo, Mendell and Rubin test; BLRT, bootstrapped likelihood ratio test.
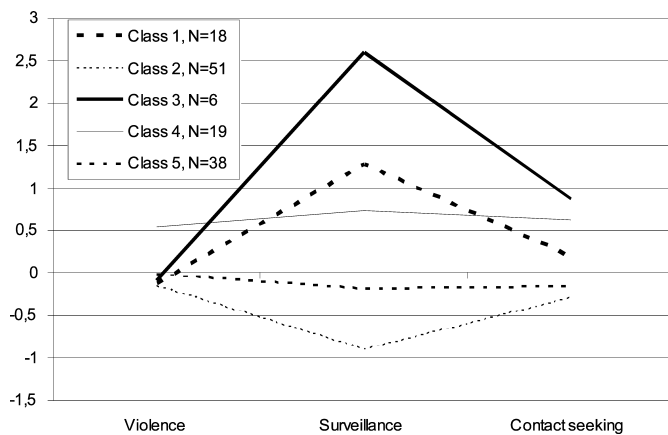
FIG. 1—*Overall means for factor scores was zero and standard deviations between 0.82 and 0.89.*

hypothesis) is a significantly better fit to the data than the four-class solution (null hypothesis) (see Table 2).

Next, each case was assigned to the most probable class (according to highest posterior probability). The average latent class posterior probabilities for the most likely latent class membership was very high, ranging from 0.97 to 1.00, which indicates that the cases could be clearly assigned to their latent classes. The five-class solution (i.e., the latent class mean profiles based on the estimated values of LPA along with the number of cases assigned to each class) in relation to the stalking dimension scores can be seen in Fig. 1. The five identified classes (victim subgroups) consist of victims exposed to stalking behaviors that are homogeneous within classes and heterogeneous across classes.

Four of the five classes (victim subgroups) plot very similarly in terms of the overall mean (e.g., zero mean) within the violence dimension. Only class 4 plots 0.5 over the overall mean. The victims assigned to class 4 also plot on average 0.5 over the overall mean regarding surveillance and contact-seeking behaviors. This class was labeled victims of *wide scope stalkers*. The cases assigned to class 1 plot fairly high regarding surveillance and were labeled victims of *surveillance stalkers*. The second class plots below the overall means, especially in respect of surveillance, and was labeled victims of *low-profile stalkers*. Those assigned to class 3 plot very highly in surveillance and relatively highly in contact seeking. This low on frequency class was labeled victims of *social lurker stalkers*. The victims assigned to class 5 were labeled as victims of *baseline stalkers* as those assigned to this group plot very near the overall means regarding all three stalking dimensions. The distribution of victim–stalker relationship and classes can be seen

TABLE 3—*Distribution of victim–stalker relationship in relation to classes (victim subgroups).*

Class	Stranger (N = 18)	Acquaintance (N = 88)	Ex-Intimate (N = 31)	Total (N)
Victims of...				
I Surveillance stalkers	4 (22.2%)	12 (66.7%)	2 (11.1%)	18
II Low-profile stalkers	8 (15.7%)	31 (60.8%)	12 (23.5%)	51
III Social lurker stalkers	1 (16.7%)	5 (83.3%)	–	6
IV Wide scope stalkers	1 (11.1%)	11 (57.9%)	7 (36.8%)	19
V Baseline stalkers	4 (10.5%)	25 (65.8%)	9 (23.7%)	38
Total	18 (13.6%)	84 (63.6%)	30 (22.7%)	132

TABLE 4—*Final general linear model with duration of stalking as the dependent variable.*

Effect	F	df	p-Value	Partial η^2
Class	4.01	4	0.004	0.122
Surveillance	3.64	1	0.059	0.031
Contact seeking	11.21	1	0.001	0.089
Class × surveillance	10.40	4	0.001	0.266
Class × contact seeking	6.96	4	0.001	0.195
Relationship	7.60	2	0.001	0.117

$R^2 = 0.492$, adjusted $R^2 = 0.421$, error df = 115.

in Table 3. The prior victim–stalker relationship categories were not significantly associated with the classes.

In the last stage of the analysis, a set of variables (stalking dimensions, classes identified in prior analysis, victim–stalker relationship, stalker age) were entered into the GLM to investigate their relation to duration of stalking. The final model for duration of stalking is presented in Table 4. As the stalker age variable and the violence dimension did not reach statistical significance, they were excluded from the final model. Furthermore, as there were no interactions between the victim–stalker relationship category and stalking dimensions, they were also excluded from the final model.

The values of estimated parameters of the final GLM can be seen in the table below. The prediction below is valid for strangers. The prediction is also valid for ex-intimates and acquaintances when adding the constant value 1.507 for ex-intimates and 1.205 for acquaintances, based on the main effect of the relationship variable in Table 4.

The results demonstrate that the expected number of logarithmic stalking days (Constant) was shorter among victims of stranger stalkers than among victims of acquaintances and ex-partners, being highest among victims of ex-partners. Table 5 shows that across victim subgroups, the expected number of stalking days is lowest (3.154) for victims of *low-profile stalkers* (group II), but increases strongly if the factor score of surveillance is higher than the expected mean (−0.900). Furthermore, if the factor score of contact

TABLE 5—*General linear model parameter estimates for stalking days.*

	I Surveillance	II Low Profile	III Social Lurker	IV Wide Scope	V Baseline
Constant	3.547	3.154	4.975	4.073	3.299
Surveillance (β)	−0.382	13.376	0.046	−2.409	1.812
Contact seeking (β)	0.312	−3.086	−1.281	0.325	−0.812

Surveillance group means for each subgroup are 1.271, −0.900, 2.603, 0.734, −0.172, and contact-seeking group means for each subgroup are 0.193, −0.290, 0.883, 0.622, −0.153, respectively.

seeking is higher than the expected mean (−0.290), the number of stalking days decreases clearly. The expected number of stalking days is highest (4.975) for victims of *social lurker stalkers* (group III). Also, the expected mean of contact seeking (0.883) is highest in this group. When the factor score of contact seeking is higher than the expected mean (0.883), the number of stalking days decreases. However, the change in surveillance factor score has no effect on the expected number of days because the regression coefficient is very small (0.046).

Discussion

The purpose of this study was threefold. The first task was to identify at least two stalking behavior dimensions based on prior research. As hypothesized, the factor solution for stalking behaviors could be given an interpretation with three distinct regions: violent, contact seeking, and surveillance stalking. The structure of these three stalking dimensions appears purposeful and consistent in the light of previous research (20,24–29).

The second task was to identify distinct subgroups of stalking victims based on the identification of the above-mentioned stalking dimensions. Results from the LPA revealed five distinct victim subgroups: victims of (I) surveillance, (II) low-profile, (III) social lurker, (IV) wide scope, and (V) baseline stalkers. Victims assigned to the surveillance group (I) were distinguished by their exposure to "classical stalking" patterns, e.g., being followed by their stalker. The second subgroup was the high-frequency group, i.e., most victims were assigned to this low-profile stalker group. Victims of this subgroup were distinguished by reporting less exposure to all types of stalking, especially surveillance. The third, low-frequency group (victims of social lurker stalkers) showed distinctive patterns of exposure regarding both surveillance and contact-seeking stalker behaviors. Victims of wide scope stalkers (group IV) had experienced the highest level of stalking violence in combination with exposure to similar levels (above the overall mean) of surveillance and contact-seeking behaviors. As a whole, a relatively high proportion of the stalkers were engaged in stalking behaviors described as baseline stalking. The distinctive feature of this type of stalking behavior is the absence of any real distinctive features. On the other hand, it is of importance to acknowledge the co-occurrence of the different types of stalking behavior. Ultimately, the most distinctive differences across the victim subgroups emerged concerning the exposure to surveillance.

Previous studies indicate that ex-intimates are the most violent stalkers (31–34). The present results, however, do not confirm previous findings concerning ex-intimate stalkers being the subgroup most prone to violent behavior. Moreover, the victim–stalker relationship categories were not significantly associated with the subgroups, even if some variation across the groups can be seen. Thus, it seems that all victim–stalker relational groups (stranger, acquaintance, and ex-intimate stalkers) engage in a relatively broad range of stalking behaviors, including a variety of co-occurring behaviors from all three stalking dimensions (violence, surveillance, contact seeking).

The third task of the present study was to investigate how the duration of stalking is associated with stalking behavior and the type of prior victim–stalker relationship. As a whole, the mean stalking duration was twice as long in the present study compared to the mean duration reported in studies by Fisher et al. (19) and in Spitzberg and Rhea's (20) work. One possible explanation of this might be that the Finnish study included a higher volume of stalkers known to the victim, and this stalker type is known to stalk for longer periods. Furthermore, in line with prior studies, the duration

of stalking in the present study varied also according to prior relationship, with ex-intimate partners being the most persistent stalkers, and strangers stalking for the shortest period of time. A somewhat surprising finding was that acquaintances stalked for almost as long a period of time as ex-intimate stalkers. In one of every three cases, acquaintance stalkers stalked for more than 6 months. A recent study has reported a similar finding of highly persistent acquaintance stalkers. (21).

Currently, a prior victim–stalker relationship is considered to be the best predictor of the duration of stalking (21,23). It was also a significant predictor of the duration of stalking in the present study. However, the present study shows that the prediction of stalking duration could be considerably improved by stalking behavior dimensions and victim subgroups based on these stalking dimensions. Together, these predictors accounted for a considerable amount of the variance regarding the duration of stalking. There is at least one recent study that also combined other aspects of stalking with prior relationship in a stalker sample, resulting in an improved prediction of stalking persistence (21). Interestingly, the violence dimension had to be dropped from the final model of stalking duration, as it did not show any significant effect on duration except that victims of wide scope stalkers had a higher mean value of violence than any other subgroup. To our knowledge, there are only a few prior studies (31,38) that have examined the association between the duration of stalking and violence, and the results are conflicting. The present findings do not confirm either of these results, as no associations were found. This finding calls for further observation and research. However, the surveillance and contact-seeking dimensions emerged as significant predictors of stalking duration. A more exact prediction for the identified victim subgroups was possible, varying according to the relationship with the stalker and stalking dimension. However, the small group size of one of the subgroups could have limited the power of statistical tests. As a whole, the predictive models might be able to capture some of the multidimensionality of the stalking phenomenon and identify central factors for estimating stalking duration/persistence. At the same time, these same factors, i.e., the co-occurring elements of stalking along with the complexity of the phenomenon, can make the interpretation and applicability very challenging.

The present study has some limitations, which have to be taken into consideration when evaluating the results. First, the results are based on a university population, and because of this age-related bias, one cannot generalize to the public at large. Second, the results are based on retrospective self-reports in which memory distortions may occur. Third, the respondents were provided with a finite list of stalking and harassing behaviors. It is possible that some behaviors that they had in fact experienced were missing from the list. Fourth, because of using an electronic questionnaire, it is not possible to estimate the response rate, as it is not possible to know how many persons in fact read the mail. Also, the present study provides robust information on one specific episode of stalking. However, as most of the previous studies have not differentiated between one vs. multiple stalking episodes, a direct comparison concerning incidence rates should be conducted with caution. Lastly, the definition provided in the stalking questionnaire is rather broad and may have resulted in latent classes that would not emerge in studies using a more strict definition. Also, this might have impacted on the findings regarding the duration of stalking.

Stalker behavior and prior victim–stalker relationship are often the principle known factors in any given stalking case and are therefore of great importance when trying to identify factors contributing to the duration of stalking. The present study has shown

that stalking victim subgroups based on measures of behavioral stalking dimensions and victim–stalker relationship play an important part in explaining the duration of stalking. Both the theoretical and clinical implications concerning latent classes need to be further studied, and their use in predictive models should be tested to achieve greater broader applicability for clinicians and law enforcement. The present findings, however, indicate that to capture the essence of stalking behavior, we should go beyond stalking dimensions.

To conclude, the findings of the present study suggest that it is possible to distinguish victim subgroups that are homogeneous within the group regarding exposure to stalking and are heterogeneous across subgroups. The utility and importance of distinguishing subgroups within various samples have been stressed in several recent studies. There is a growing use of person-centered analysis such as LPA and latent class analysis (LCA) by applied social sciences and recently in forensic sciences (53–56). These studies suggest that person-orientated methods hold both a clinical and forensic utility, e.g., for developing tailored intervention programs for victims and classifying offenders. Thus, it is our hope that these findings would encourage the further application of person-centered analysis within forensic sciences.

References

1. Meloy JR, Gothard S. Demographic and clinical comparison of obsessional followers and offenders with mental disorders. Am J Psychiatry 1995;152:258–63.
2. Mullen PE, Pathé M, Purcell R. Stalkers and their victims. Cambridge: Cambridge University Press, 2000.
3. Pathé M, Mullen PE. The impact of stalkers on their victims. Br J Psychiatry 1997;170:12–7.
4. Roberts KA. Women's experience of violence during stalking by former romantic partners. Violence Against Women 2005;1:89–144.
5. Westrup D, Fremouw WJ. Stalking behavior: a literature review and suggested functional analytic assessment technology. Aggress Violent Behav 1998;3:255–74.
6. Amar AF. College women's experience of stalking: mental health symptoms and change in routines. Arch Psychiatr Nurs 2006;20:108–16.
7. Blaauw E, Winkel FW, Arensman E, Sheridan L, Freeve A. The toll of stalking: the relationship between features of stalking and psychopathology of victims. J Interpers Violence 2002;17:50–63.
8. Blaauw E, Winkel FW, Sheridan L, Malsch M, Arensman E. The psychological consequences of stalking victimisation. In: Boon J, Sheridan L, editors. Stalking and psychosexual obsession. New York, NY: John Wiley and Sons, 2002;23–33.
9. Dressing H, Kuehner C, Gass P. Lifetime prevalence and impact of stalking in a European population: epidemiological data from a middle-sized German city. Br J Psychiatry 2005;187:168–72.
10. Hall DM. The victims of stalking. In: Meloy JR, editor. The psychology of stalking: clinical and forensic perspectives. San Diego, CA: Academic Press, 1998;113–37.
11. Kamphuis JH, Emmelkamp PMG. Traumatic distress among support-seeking female victims of stalking. Am J Psychiatry 2001;158:795–8.
12. Kamphuis JH, Emmelkamp PMG, Bartak A. Individual differences in post-traumatic stress following post-intimate stalking: stalking severity and psychosocial variables. Br J Clin Psychol 2003;42:145–56.
13. Pathé M, Mullen PE. The victim of stalking. In: Boon J, Sheridan L, editors. Stalking and psychosexual obsession. New York, NY: John Wiley and Sons, 2002;1–22.
14. Purcell R, Pathé M, Mullen PE. Association between stalking victimisation and psychiatric morbidity in a random community sample. Br J Psychiatry 2005;187:416–20.
15. Spitzberg BH. The tactical topography of stalking victimization and management. Trauma Violence Abuse 2002;3:261–88.
16. Mullen PE, Pathé M, Purcell R. Study of stalkers. Am J Psychiatry 1999;156:1244–9.
17. Tjaden P, Thoennes N. Stalking in America: findings from the National Violence Against Women survey. Washington, DC: National Institute of Justice and Centers for Disease Control and Prevention, 1998; Publication NCJ 169592.
18. Spitzberg BH, Cupach WR. The state of the art of stalking: taking stock of the emerging literature. Aggress Violent Behav 2007;12:64–86.
19. Fisher BS, Cullen FT, Turner MG. Being pursued: stalking victimization in a national study of college women. Criminol Public Policy 2002;1:257–308.
20. Spitzberg BH, Rhea J. Obsessive relational intrusion and sexual coercion victimization. J Interpers Violence 1999;14:3–20.
21. McEwan TE, Mullen PE, MacKenzie R. A study of the predictors of persistence in stalking situations. Law Hum Behav, doi: 10.1007/s10979-008-9141-0 (accessed July 15, 2008).
22. Mullen P, Mackenzie R, Ogloff J, Pathé M, McEwan T, Purcell R. Assessing and managing the risks in stalking situation. J Am Acad Psychiatry Law 2006;34:439–50.
23. McEwan T, Mullen PE, Purcell R. Identifying risk factors in stalking: a review of current research. Int J Law Psychiatry 2007;30:1–9.
24. Cupach WR, Spitzberg BH. Obsessive relational intrusion: incidence, perceived severity and coping. Violence Vict 2000;15:357–72.
25. Dennison SM, Stewart A. Facing rejection: new relationships, broken relationships, shame and stalking. Int J Offender Ther Comp Criminol 2006;50:324–37.
26. Haugaard JJ, Seri LS. Stalking and other forms of intrusive contact after the dissolution of adolescent dating on romantic relationships. Violence Vict 2003;18:279–97.
27. Sinclair HC, Frieze IH. Initial courtship behavior and stalking: how should we draw the line? In: Davies KE, Frieze IH, Maiuro RD, editors. Stalking: perspectives on victims and perpetrators. New York, NY: Springer Publishing Company, 2002;186–211.
28. Sinclair HC, Frieze IH. When courtship persistence becomes intrusive pursuit: comparing rejecter and pursuer perspectives of unrequited attraction. Sex Roles 2005;52:839–52.
29. Spitzberg BH, Nicastro AM, Cousins AV. Exploring the interactional phenomenon of stalking and obsessive relational intrusion. Commun Rep 1998;11:33–47.
30. Purcell R, Pathé M, Mullen PE. The prevalence and nature of stalking in the Australian community. Aust N Z J Psychiatry 2002;36:114–20.
31. James DV, Farnham FR. Stalking and serious violence. J Am Acad Psychiatry Law 2003;31:432–9.
32. Palarea RE, Zona MA, Lane JC, Langhinrichsen-Rohling J. The dangerous nature of intimate relationship stalking: threats, violence, and associated risk factors. Behav Sci Law 1999;17:269–83.
33. Rosenfeld B. Violence risk factors in stalking and obsessional harassment. A review and preliminary meta-analysis. Crim Justice Behav 2004;31:9–36.
34. Sheridan L, Davies GM. Violence and the prior victim–stalker relationship. Crim Behav Ment Health 2001;11:102–16.
35. Sinwelski SA, Vinton L. Stalking: the constant threat of violence. J Women Soc Work 2001;16:46–66.
36. Logan TK, Leukefeld C, Walker B. Stalking as a variant of intimate violence: implications from a young adult sample. Violence Vict 2000;15:91–111.
37. Acevedo SA. An exploration of victim-initiated interventions and duration of stalking. http://www.lib.umd.edu/drum/handle/1903/4243 (accessed December 11, 2006).
38. Purcell R, Pathé M, Mullen PE. Editorial: when do repeated intrusions become stalking? J Forens Psychiatry Psychol 2004;4:571–83.
39. Sheridan L, Davies GM, Boon JCW. Stalking: perceptions and prevalence. J Interpers Violence 2001;16:151–67.
40. Jagessar JDH, Sheridan LP. Stalking perceptions and experiences across two cultures. Crim Justice Behav 2004;31:97–119.
41. Sheridan L, Gillett R, Davies GM. Stalking: seeking the victim's perspective. Psychol Crime Law 2000;6:267–80.

42. Sheridan L, Gillett R, Davies GM. Perceptions and prevalence of stalking in a male sample. Psychol Crime Law 2002;8:289–310.
43. Meloy JR. The psychology of stalking. In: Meloy JR, editor. The psychology of stalking: clinical and forensic perspectives. San Diego, CA: Academic Press, 1998;1–23.
44. Sheridan L, Davies GM. Stalking: the elusive crime. Legal Criminol Psychol 2001;6:133–47.
45. Muthén LK, Muthén BO. Mplus user's guide, 5th edn. Los Angeles, CA: Muthén & Muthén, 1998–2007.
46. Muthén B, Shedden K. Finite mixture modelling with mixture outcomes using the EM algorithm. Biometrics 1999;55:463–9.
47. Muthén BO. Latent variable mixture modelling. In: Marcoulides GA, Schumacker RE, editors. New developments and techniques in structural equation modelling. Philadelphia, PA: Lawrence Erlbaum Associates, Inc, 2001;1–34.
48. Akaike H. Factor analysis and AIC. Psychometrika 1987;52:317–32.
49. Schwartz G. Estimating the dimension of a model. Ann Stat 1978;6:461–4.
50. Sclove SL. Application of model-selection criteria to some problems in multivariate analysis. Psychometrika 1987;52:333–43.
51. Lo Y, Mendell RN, Rubin DB. Testing the number of components in a normal mixture. Biometrika 2001;88:767–78.
52. McLahlan G, Peel D. Finite mixture models. New York, NY: Wiley and Sons, 2000.
53. Macy RJ, Nurius PS, Norris J. Latent profiles among sexual assault survivors: understanding survivors and their assault experiences. J Interpers Violence 2007;22:520–42.
54. Nurius PS, Macy RJ. Heterogeneity among violence-exposed women: applying person-orientated research methods. J Interpers Violence 2008;23:389–415.
55. Schwalbe CS, Macy RJ, Day SH, Fraser MW. Classifying offenders: an application of latent profile analysis to needs assessment in juvenile justice. Youth Violence Juv Justice 2008;6:279–94.
56. Vaughn MG, DeLisi M, Beaver KM, Howard MO. Toward a quantitative typology of burglars: a latent profile analysis of career offenders. J Forensic Sci 2008;53:1387–92.

Additional information and reprint requests:
Katja Björklund, Ph.D.
Institute of Behavioural Sciences
University of Helsinki
PO Box 9
00014 Helsinki
Finland
E-mail: katja.bjorklund@helsinki.fi

PAPER

TOXICOLOGY; PATHOLOGY AND BIOLOGY

Shane Darke,[1] *Ph.D.; Johan Duflou,*[1,2,3,4] *M.Med.Path. (Forens), F.R.C.P.A.;*
and Michelle Torok,[1] *B.Sc.*

Comparative Toxicology of Intentional and Accidental Heroin Overdose*

ABSTRACT: The demographic and toxicological characteristics of deliberate (SUI, $n = 50$) and accidental (ACC, $n = 927$) fatal heroin overdose cases were examined. SUI cases were more likely to be female, had lower body mass indices, were more likely to be enrolled in treatment and less likely to have hepatic pathology. The median blood morphine concentration of SUI cases was significantly higher than that of ACC cases (0.70 vs. 0.40 mg/L, $p < 0.001$). Blood morphine concentrations of >1 mg/L were seen among 38.0% of SUI cases compared to 13.9% of ACC cases. Being a member of the SUI group remained a significant independent predictor of higher morphine concentrations after controlling for the effects of potential confounders ($p < 0.001$), other significant predictors being the absence of alcohol ($p < 0.001$), the presence of methadone ($p < 0.05$), and the presence of cocaine ($p < 0.05$). The current data are consistent with the view that suicide forms a small, but distinct, category of heroin overdose cases, rather than overdose being a parasuicidal phenomenon *per se*.

KEYWORDS: forensic science, toxicology, opioids, overdose, morphine, suicide

Opioids have the highest death rates of all psychoactive substances and make the single largest contribution to illicit drug-related mortality (1). Two of the major causes of death among opioid users are overdose and suicide (1). Given that rates of major depression are extremely high among opioid users (2) and the risk of completed suicide is 14 times that of the general population (3), it has been hypothesized that large proportions of heroin overdoses are, in fact, parasuicides (4–7). In support of this view, a number of authors have noted strong statistical associations between overdose and suicide histories (4–13) and that overdose has been associated with higher levels of depression and suicidal ideation (6,14–16).

Others, however, dispute this view, arguing that the vast majority of overdoses, both fatal and nonfatal, are accidental (1,8,10–12,17–20). Consistent with this, only small proportions of those who have overdosed report their most recent overdose was deliberate (8,17). Furthermore, only small proportions of fatal overdose cases are classified as suicides (18,20). It could be argued, of course, that the latter finding is an artifact of these being unrecognized suicides.

While the relationship between overdose and suicide is controversial, to date, no study has examined the comparative demographics and toxicology of deliberate and accidental heroin overdose. Are there meaningful demographic and toxicological differences between overdose cases deemed to be accidental and those deemed deliberate? In particular, are there differences in the morphine concentrations of these two groups? If these two types of death are essentially the same phenomenon, then their morphine concentrations should be similar. If suicide cases are a distinct category of overdose, then morphine concentrations should be higher among those who deliberately attempt to overdose. The current study aimed to compare the demographic and toxicological characteristics of deliberate and accidental fatal heroin overdose.

Methods

Case Identification

The Department of Forensic Medicine is located in central Sydney and is the primary forensic pathology center in NSW, conducting approximately 2000 autopsies per annum. In New South Wales (NSW), a case must be reported to the Coroner where a person dies a violent or unnatural death. All cases in this study underwent a standardized forensic autopsy, with examination of all major organs and quantitative toxicological analysis. Cause of death is determined by the forensic pathologist on the basis of circumstances of death, the comprehensive autopsy findings, and the toxicological analyses. Permission to inspect the files had been received from the NSW State Coroner and the Sydney South West Area Health Service human research ethics committee, and all cases were reviewed by the authors.

All cases autopsied at the Department of Forensic Medicine between January 1, 1998, and December 31, 2007, were identified in which the cause of death was attributed to opioid toxicity, attributable to either morphine (the major metabolite of heroin) or the

[1]National Drug and Alcohol Research Centre, University of New South Wales, Sydney, NSW, Australia.

[2]Department of Forensic Medicine, Sydney South West Area Health Service, Sydney, NSW, Australia.

[3]School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia.

[4]Department of Pathology, University of Sydney, Sydney, NSW, Australia.

acute physical sequelae of such toxicity (e.g., hypoxic brain damage). For the purposes of this study, cases in which blood morphine was not detected were excluded. All reported morphine concentrations were of total morphine. A total of 977 heroin overdose cases meeting these criteria were identified. These included 927 cases of accidental overdose (ACC) and 50 cases deemed to be by deliberate heroin overdose (SUI). Suicide was determined by the authors during the file audit on the basis of the presence of suicide notes, verbal statements of intent given to witnesses, police reports, witness statements, and/or concomitant use of other methods (e.g., hose attached to car exhaust, large quantities of pill fragments in stomach). In 34 cases, notes were present, in eight, verbal intent occurred immediately prior to the event, in five cases, the concomitant use of other methods was noted, and in three cases, police statements that the circumstances and personal history indicated suicide were noted.

Toxicology

All presented toxicological analyses were of blood and were conducted by the Division of Analytical Laboratories (DAL). Toxicological data were reported for total body morphine (the primary metabolite of heroin), methadone, alcohol, cannabis (determined by the presence of Δ-9-THC), methamphetamine, cocaine (determined by the presence of cocaine and/or benzoylecgonine), 3,4methylenedioxymethamphetamine (MDMA), benzodiazepines, antidepressants, and antipsychotic medications. Quantitative analyses are presented for total body morphine only, which is the measure employed by DAL, and includes both bound and unbound morphine. In cases where there was prolonged hospitalization prior to death, antemortem toxicology was reported. In all cases, drugs administered by hospital and medical staff were excluded.

Autopsy Findings

Information was collected on age (years), body length (m), and weight (kg), and body mass index (BMI) was calculated. Circumstances of death, and brief case histories, were obtained from accompanying police summaries to the coroner. Information was recorded on all major pathology noted in autopsy reports, including coronary, pulmonary, hepatic, and renal pathology. Hepatic pathology was of particular relevance to the current study, as such pathology may potentially affect morphine concentrations. Data collected included the presence of steatosis, fibrosis, cirrhosis, hepatomegaly, and necrosis.

Statistical Analyses

Where distributions were highly skewed, medians and inter-quartile ranges (IQR) were reported, otherwise means were presented. For bivariate comparisons, t-tests or odds ratios (OR) with 95% confidence interval (CI) were reported. Independent predictors of blood morphine concentration were determined by simultaneous multiple regression. All analyses were conducted using SPSS for Windows (release 15.0) (21).

Results

Case Characteristics

There was no significant group difference in age, but SUI cases were more likely to have been enrolled in drug treatment at the time of death. SUI cases were also more likely to be women, had lower BMIs, and were less likely to have a diagnosis of hepatic pathology (Table 1). The only group difference in hepatic pathology was that lymphocytic infiltrates were less common among the SUI group (22.4 vs. 43.8%, OR 0.37, CI 0.19–0.74). There were no significant group differences in diagnoses of cardiovascular pathology, pulmonary pathology, or renal pathology.

Toxicology

The median blood morphine concentration of SUI cases was significantly higher than that of the ACC cases (Table 1). This difference was particularly evident at concentrations of >1 mg/L (more than twice the median concentration of ACC cases) (SUI: 34.0% v ACC: 13.8%) (Fig. 1).

Alcohol was the most commonly detected drug other than morphine, present in 41.1% of cases ($n = 402$). The median concentration among those in which alcohol was present was 0.12 g/100 mL (IQR 0.13, range 0.01–0.78 g/100 mL). The next most commonly detected substances other than morphine were the benzodiazepines (33.2%, $n = 325$). The most commonly detected benzodiazepines were diazepam (26.4%), oxazepam (6.6%), and temazepam (5.5%).

The only significant difference between SUI and ACC cases in the presence of drugs other than morphine was for antidepressants, more common among the SUI group. There were no group differences in the presence of alcohol, benzodiazepines, cocaine, methadone, methamphetamine, MDMA, antipsychotics, or cannabis.

To determine whether the observed group difference between morphine concentrations was an artifact, a simultaneous multiple regression was conducted using demographic characteristics and factors associated with overdose and blood morphine concentrations

TABLE 1—Characteristics of intentional and accidental fatal heroin overdose cases.

	Suicide ($n = 50$)	Accidental ($n = 927$)	Comparisons
Demographics			
Age	34.3	33.0	ns, $p = 0.19$
Sex (% male)	70.0	82.3	OR 0.50, CI 0.27–0.94
In drug treatment (%)	16.0	6.3	OR 2.83, CI 1.27–6.32
BMI	23.4	25.5	$t_{963} = 2.95$, $p < 0.01$
Pathology (%)			
Hepatic	51.0	65.8	OR 0.54, CI 0.30–0.96
Cardiovascular	40.1	33.8	ns, $p = 0.31$
Pulmonary	24.5	32.0	ns, $p = 0.27$
Renal	8.2	11.0	ns, $p = 0.53$
Toxicology			
Morphine(mg/L)	0.70 (IQR 1.7, range 0.05–35.0)	0.40 (IQR 0.55, range 0.05–34.0)	U = 21,907, $p < 0.001$
Other drugs (%)			
Alcohol	42.0	41.1	ns, $p = 0.90$
Benzodiazepines	38.0	32.9	ns, $p = 0.46$
Cocaine	16.0	16.5	ns, $p = 0.93$
Methadone	12.0	5.9	ns, $p = 0.08$
Methamphetamine	2.0	7.7	ns, $p = 0.14$
MDMA	4.0	2.2	ns, $p = 0.39$
Antidepressants	20.0	10.7	OR 2.01, CI 1.01–4.31
Antipsychotics	4.0	5.1	ns, $p = 0.73$
Cannabis	2.0	2.5	ns, $p = 0.84$

Referent group = "Accidental."
ns, not significant; BMI, body mass index; IQR, inter-quartile ranges.
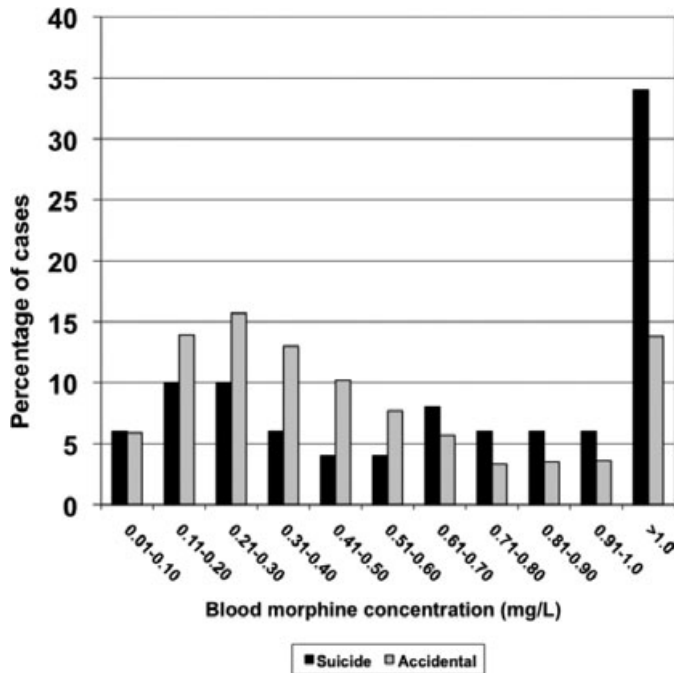
FIG. 1—*Blood morphine concentrations of intentional and accidental fatal heroin overdose.*

in previous research (1) as predictor variables. Variables entered into the model were: group, age, sex, BMI, and the presence/absence of alcohol, methadone, cocaine, antidepressants, and hepatic pathology. After controlling for the effects of potential confounders, being a member of the SUI group remained a significant independent predictor of higher morphine concentrations ($\beta = 0.12$, $t = 3.71$, $p < 0.001$). Other independent predictors of higher morphine concentrations were the absence of alcohol ($\beta = -0.09$, $t = 2.79$, $p < 0.001$), the presence of methadone ($\beta = 0.07$, $t = 2.06$, $p < 0.05$), and the presence of cocaine ($\beta = 0.07$, $t = 2.12$, $p < 0.05$).

Discussion

The major finding of this study was the substantial difference between the blood morphine concentrations of deliberate and ACC. This was particularly pronounced at high morphine concentrations. Importantly, this difference remained significant after potential confounders were controlled. Thus, these differences do not appear to be artifacts of demographic differences, concomitant drug use, or hepatic pathology. The current data are consistent with the view that suicide forms a small, but distinct, category of heroin overdose cases, rather than overdose being a parasuicidal phenomenon *per se*.

The only other toxicological difference was that SUI cases were twice as likely to have antidepressants present, a probable marker for depression. In other regard, the toxicology was typical of fatal heroin overdose, particularly the high prevalence of alcohol (1). Men predominated in both groups, which is typical of overdose series. SUI case were more likely to be women; however, a possible reflection of the fact that men, and male heroin users, are more likely to attempt suicide through means other than self-poisoning (1,22,23). The fact that a minority of cases were in drug treatment at the time of death is consistent with the characteristics of

overdose cases; generally, SUI cases were, however, nearly three times more likely to be enrolled in treatment. This may indicate more distressed heroin users having sought treatment or a greater likelihood of screening and intervention among the treatment population.

The current data also have implications for interventions to reduce overdose mortality. If overdose is a manifestation of depression and suicidal behaviors, then interventions should focus on reducing depression and suicidal ideation. If overdose is, predominantly, a distinct phenomenon from suicide, then the focus should be on known risk factors for overdose, such as concomitant alcohol consumption (1). The current toxicological data are consistent with the latter view.

As with all studies, caveats need to be borne in mind. First, hair analyses were not conducted as part of the standard forensic investigation procedures. As such, no data were available on drug use patterns over the months preceding death. Second, the measure of morphine available was of total body morphine, which includes both bound and unbound morphine. As such, the amount of active morphine at the time of death is uncertain. Third, it is possible that there are other unknown toxicogenomic factors linking depression and hepatic metabolism, which may affect the results. Finally, care should always be taken in extrapolating to other opioid users. The characteristics of these cases, however, were typical of opioid fatalities reported elsewhere (1,18,24,25).

In summary, overdose suicide cases exhibited substantially higher blood morphine concentrations than those deemed accidental. Rather than overdose being a parasuicidal phenomenon, suicide appears to form a toxicologically distinct category of overdose cases.

References

1. Darke S, Degenhardt L, Mattick R. Mortality amongst illicit drug users: epidemiology, causes and intervention. Cambridge: Cambridge University Press, 2007.
2. Teesson L, Havard A, Fairbairn S, Ross J, Lynskey M, Darke S. Depression among entrants for heroin dependence in the Australian Treatment Outcome Study (ATOS): prevalence, correlates and treatment seeking. Drug Alcohol Depend 2005;78:309–15.
3. Wilcox HC, Connor KR, Caine ED. Association of alcohol and drug use disorders and completed suicide: an empirical review of cohort studies. Drug Alcohol Depend 2004;76S:S11–9.
4. Neale J. Suicidal intent in non-fatal illicit drug overdose. Addiction 2000;95:85–93.
5. Neale J, Robertson M. Recent life problems and non-fatal overdose among heroin users entering treatment. Addiction 2005;100:168–75.
6. Rossow I, Lauritzen G. Balancing on the edge of death: suicide attempts and life-threatening overdoses among drug addicts. Addiction 1999;94:209–19.
7. Vingoe L, Welch S, Farrell M, Strang J. Heroin overdose among a treatment sample of injecting drug misusers: accident or suicidal behaviour? J Subst Use 1999;4:88–91.
8. Darke S, Ross J. The relationship between suicide and overdose among methadone maintenance patients in Sydney, Australia. Addiction 2001;96:1443–53.
9. Farrell M, Neeleman J, Griffiths P, Strang J. Suicide and overdose among opiate addicts. Addiction 1996;91:321–33.
10. Johnsson E, Fridell M. Suicide attempts in a cohort of drug abusers: a five year follow-up study. Acta Psychiatr Scand 1997;96:362–6.
11. Kjelsberg E, Winther M, Dahl AA. Overdose deaths in young substance abusers: accidents or hidden suicides? Acta Psychiatr Scand 1995;91:236–42.
12. Kosten TR, Rounsaville BJ. Suicidality among opioid addicts: a 2.5 year follow-up. Am J Drug Alcohol Abuse 1988;14:357–69.
13. Oyefeso A, Ghodse H, Clancy C, Corkery JM. Suicide among drug addicts in the UK. Br J Psychiatry 1999;175:277–82.
14. Best D, Gossop M, Man L, Finch E, Greenwood J, Strang J. Accidental and deliberate overdose among opiate addicts in methadone maintenance treatment: are deliberate overdoses systematically different? Drug Alcohol Rev 2000;19:213–6.

15. Burns JM, Martyres RF, Clode D, Boldero JM. Overdose in young people using heroin: associations with mental health, prescription drug use and personal circumstances. Med J Aust 2004;181:S25–8.
16. Stewart D, Gossop M, Marsden J. Reductions in non-fatal overdose after drug misuse treatment: results from the National Treatment Outcome Research Study (NTORS). J Subst Abuse Treat 2002;22:1–9.
17. Darke S, Ross J, Hall W. Overdose among heroin users in Sydney, Australia I. Prevalence and correlates of non-fatal overdose. Addiction 1996;91:405–11.
18. Darke S, Ross J, Zador D, Sunjic S. Heroin-related deaths in New South Wales, Australia, 1992-1996. Drug Alcohol Depend 2000;60:141–50.
19. Ravndal E, Vaglum P. Overdoses and suicide attempts: different relations to psychopathology and substance abuse? A 5-year prospective study of drug abusers. Eur Addict Res 1999;5:63–70.
20. Darke S, Kaye S, Duflou J. Systemic disease among cases of fatal opioid toxicity. Addiction 2006;101:1299–305.
21. SPSS inc. SPSS for Windows (version 15.0). Chicago: SPSS Inc., 2006.
22. Denning DG, Conwell Y, King D, Cox C. Method choice, intent and gender in completed suicide. Suicide Life Threat Behav 2001;30:282–8.
23. Shenassa ED, Catlin SN, Buka SL. Lethality of firearms relative to other suicide methods: a population based study. J Epidemiol Community Health 2003;57:120–4.
24. Davidson PJ, McLean RL, Kral AH, Gleghorn AA, Edlin BR, Moss AR. Fatal heroin-related overdose in San Francisco, 1997-2000: a case for targeted intervention. J Urban Health 2003;80:261–73.
25. Maxwell JC, Pullum TW, Tannert K. Deaths of clients in methadone treatment in Texas, 1994–2002. Drug Alcohol Depend 2005;78:73–81.

Additional information and reprint requests:
Shane Darke, Ph.D.
Professor
National Drug and Alcohol Research Centre
University of New South Wales
Sydney, NSW 2052
Australia
E-mail: s.darke@unsw.edu.au

Margaret Streeter,[1] *Ph.D.*

A Four-Stage Method of Age at Death Estimation for Use in the Subadult Rib Cortex*,†

ABSTRACT: Age estimation in the subadult skeleton can be rather precise when the epiphyses and dentition are present, but incomplete or commingled remains still present a challenge. Histomorphometric age-at-death estimation methods developed for use on adults are based on the age-associated accumulation of osteons. In the growing skeleton, there is a poor correlation between osteon numbers and age until the latter half of the second decade. As a result, there has been no histological aging method for use in subadults. The analysis of the rib cortex of 72 subadults ranging in age from 2 to 21 years has identified a series of developmental changes in the bone microstructure that can be used to estimate age. This qualitative method utilizes the systematic changes in rib cortical morphology to classify ribs into one of four age phases. This method can be applied to immature skeletons in forensic, archaeological, and paleontological contexts.

KEYWORDS: forensic science, forensic anthropology, age estimation, subadult, histomorphology, rib cortex

Histological age estimation in adults is most often based on the evidence of bone remodeling, which produces an age-associated accumulation of Haversian systems (1–7). Subadult bones present a distinct and complex histomorphology compared to that of the adult. Although remodeling begins in early prenatal life (8,9) the influence of growth and modeling drift on subadult cortical bone overshadows remodeling activity in infancy and early childhood. As a result, osteon accumulations do not correlate well with age until late in the second decade of life (10). A limited amount of histomorphological data is available on the cortex of the immature rib, and the few studies to date that have considered subadult cortical bone include comparatively small samples lumped by age groups (11–15). The use of these broad age categories tends to obscure changes and patterns occurring in relation to developmental stages of growth. While age-at-death determination for subadults presents little difficulty in instances where a mostly complete skeleton is present (especially when the dentition or epiphyses are intact), age estimation in the case of commingled or very fragmentary remains is less precise (16). Analysis of the subadult rib microstructure has identified a pattern of growth and development that can be utilized to estimate age in the immature skeleton. The purpose of this study is to present a new four-phase qualitative method of age-at-death estimation for use on the subadult rib cortex, which is applicable in forensic, archaeological, and paleontological contexts.

Materials and Methods

Cortical bone samples from 72 subadult ribs (45 men and 27 women) ranging in age between 2 and 21 years, with a mean age

of 14.1 ± 5.2 years, were analyzed (Table 1). All rib samples were provided by Missouri Medical examiners obtained at autopsy from traumatic deaths and were judged to be metabolically normal. Of the 72 rib samples, 50 were reported by medical examiners to be of European ancestry, 14 of African ancestry, 5 of Asian ancestry, and 3 were recorded as mixed or unknown ancestry. Bone samples were taken from the middle third of the 5th, 6th, or 7th rib. To ensure the integrity of the bone during sectioning and grinding, the defleshed bone samples were embedded in a resin medium. At least three transverse wafers, each approximately 100 μm thick, were removed from each rib sample using a low-speed wafering saw. The wafers were then manually ground to a thickness of approximately 80 μm to produce translucent wafers suitably thin for microscopic analysis following the method developed by Frost (17). All sections were then mounted on glass microscope slides and cover-slipped following standard histological procedures (6). Rib thin sections were examined using a standard research microscope fitted with paired 10× wide field oculars and 4×, 10×, and 20× objectives. Initially, the 4× objective was used to scan the entire cortical surface to obtain an overall sense of the pattern and relative density of histomorphological structures as they relate to the pleural and cutaneous surfaces. The 10× and 20× objectives were then used to identify specific structures not clearly discernible at lower magnifications. Transmitted and polarized light were employed as necessary for optimal viewing and to distinguish morphological features such as woven bone that are not recognizable using only transmitted light. The amount and location of histological features were noted on the cutaneous and pleural cortices as defined below.

Woven bone—also called fetal or fibrous bone, quickly formed, more highly mineralized and less well organized than lamellar bone and only visible under polarized light (Fig. 1).
Primary lamellar bone—also called circumferential lamellar bone, unremodeled lamellar bone formed during modeling on periosteal or endosteal surfaces in parallel layers (Fig. 2).

TABLE 1—*Age and sex distribution by phase.*

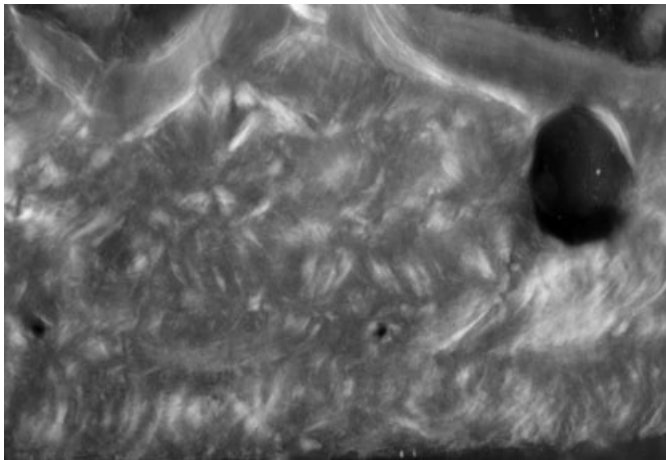| Age (years)	#Males (13.7 ± 5.4)	#Females (14.7 ± 4.8)	#Total (14.1 ± 4.8)
Phase I (<5)	4	1	5
Phase II (5–9)	7	4	11
Phase III (10–17)	22	13	35
Phase IV (18–21)	12	9	21
Total	45	27	72

FIG. 1—*Woven bone (200× polarized).*
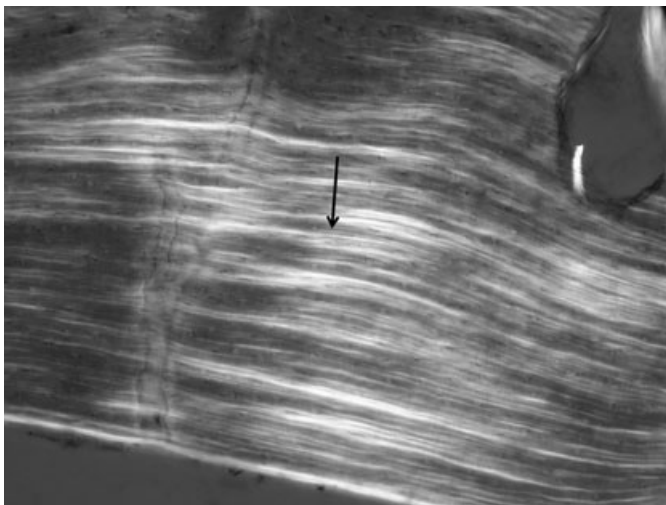
FIG. 2—*Unremodeled primary lamellar bone (arrow) (200× polarized).*

FIG. 3—*Primary vascular canals (arrow) (200×).*
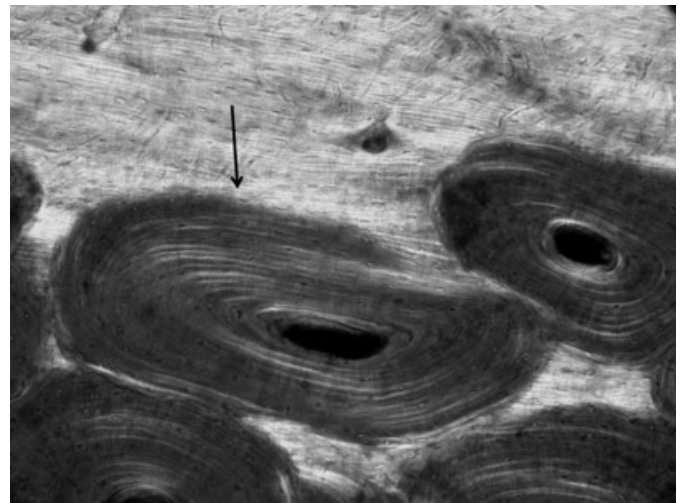
FIG. 4—*Concentric lamellae in a type I osteon (arrow) (200× polarized).*
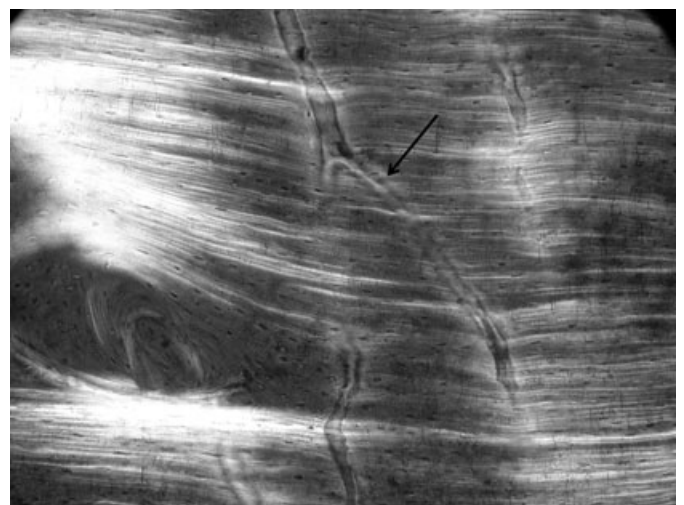
FIG. 5—*Volkmann's canals (arrow) (100× polarized).*

Primary vascular canals—vascular canals originally formed as blood vessels located on the periosteal or endosteal surfaces of bone that have subsequently been incorporated into the cortex as consecutive layers of lamellae were deposited around them during modeling (Fig. 3).

Concentric lamellar bone—lamellar bone formed secondarily during remodeling as osteons (Fig. 4).

Volkmann's canals—non-Haversian transverse vascular canals (Fig. 5).

Type I osteons—also called common osteons, secondary osteons, or Haversian systems. A packet of remodeled bone created by a

FIG. 6—*Drifting osteon (arrow) (200× polarized).*
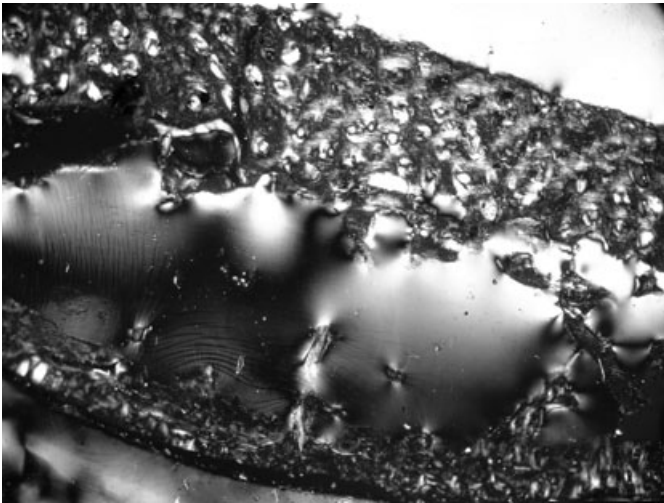
FIG. 7—*Subadult rib cortex with thicker pleural cortex (top) and thinner cutaneous cortex (bottom) both of woven bone in the rib cortex of a 3 year old (40× polarized).*

basic multicellular unit and circumscribed by a reversal line with a central Haversian canal surrounded by concentric rings of lamellae (Fig. 4).

Drifting osteons—also called "waltzing osteons" are similar in appearance to the type I osteon, but they are elongated rather than circular in cross-section, and they have eccentric Haversian canals (Fig. 6).

Characteristic differences between the cutaneous and pleural cortices are easily identified. The pleural cortex is typified by the presence of the costal groove that is established on the inferior cortical surface very early in development. Also, similar to that reported for the normal adult rib (18), there are clearly defined and easily recognizable differences between the widths of the two cortices of the subadult rib (Fig. 7). The pleural cortex is usually thicker than the thinner and more porous cutaneous cortex.

The histomorphology of the subadult rib undergoes a series of developmental phases that can be used to estimate age at death. From the initial primary spongiosa composed of woven bone, the rib cortex is subsequently transformed into primary lamellar bone,

which is remodeled into Haversian systems. However, morphological features in the developing rib cortex are not readily quantifiable in a manner that correlates well with age as they are in adults. Therefore, it was necessary to develop qualitative criteria for estimating age in the subadults comparable to the methods used to estimate age, sex, and ancestry in adults (19–21).

Results

Four phases that characterize age-associated histomorphological changes in the subadult rib were identified and are described (Table 2) and illustrated in Figs. 8–15. The age range for an unknown rib sample is estimated by comparing its histomorphology to the following age-phase descriptions and illustrations.

Phase I (<5 years of age) (Figs. 8 and 9)

The developing rib cortex is initially composed of natal woven bone (Fig. 7). The margins of the marrow cavity are indistinct at first, but the medullary space becomes more clearly defined, and the trabeculae of the marrow cavity become distinct from the cortex in the later years of this phase. Modeling drift is discernible, as the deposition of circumferential lamellar bone (Fig. 2) gradually replaces the natal woven bone (Fig. 1). This transition occurs first on the pleural cortex. Newly deposited primary lamellar bone accumulates first on the pleural endosteal and later on the cutaneous periosteal cortices. Osteoclastic resorption is evident as the scalloped surfaces (Howship's lacunae) along the pleural periosteal and the cutaneous endosteal margins. Evidence of intracortical remodeling in the form of drifting osteons begins late in this phase. The cutaneous cortex is primarily comprised of woven bone (Fig. 8), whereas the pleural cortex is distinguished by the appearance of primary lamellar bone that contains many Volkmann's canals (Fig. 9).

Phase II (5–9 years) (Figs. 10 and 11)

The defining feature of this phase is the evidence of intracortical remodeling in the form of drifting osteons (22), a unique characteristic of the subadult cortex. These drifting osteons form first in the primary lamellar bone near the periosteal surface of the pleural cortex (Fig. 11). They are often aligned in parallel rows drifting on a trajectory from the periosteum toward the endosteum. Many of these drifting osteons are incompletely formed with active bone deposition observable on one side ("the tail") of the wide, eccentrically oriented, vascular canal and resorption indicated by the presence of the scalloped Howship's lacunae on the opposite side of the canal perimeter (22). In phase II, differences in the thickness of the two cortices are marked. The pleural cortex is often two to three times the width of the thinner cutaneous cortex. Another consistent feature of this phase is the large numbers of Volkmann's canals that are dispersed throughout the pleural cortex (Fig. 5), many of them linking osteons and the periosteal and endosteal surfaces. In contrast, the cutaneous cortex is typified by a very porous appearance produced by the presence of large resorptive bays. Inspection at high magnification reveals the characteristic simultaneous resorption and formation fronts of very large drifting osteons (Fig. 8). Cortical drift is evident in the resorption (scalloped surface) on the pleural periosteal and cutaneous endosteal surfaces, while formation (smooth surface) occurs on the pleural endosteal and cutaneous periosteal surfaces. As a result of this modeling drift, the trabeculae that were originally formed as part of the marrow cavity can be seen to have been incorporated into the pleural

TABLE 2—*Summary of rib phases.*

Phase	Primary Lamellar Bone	Remodeling	Woven Bone	Cutaneous Cortex	Pleural Cortex
Phase I (<5 years)	Rare	Rare	Most of both cortices	Thinner, mostly woven bone, many primary vascular canals	Thicker, mostly woven bone, primary lamellae forms initially endosteally
Phase II (5–9 years)	Pleural cortex	Large drifting osteons initially on pleural cortex originating at periosteum	Some areas on cutaneous cortex, rare on pleural cortex	Thinner, mostly intracortical woven bone, fewer primary vascular canals	Thicker, largely primary lamellar bone, few initial drifting osteons and many Volkmann's canals
Phase III (10–17 years)	Both cortices intracortically, and periosteally on cutaneous cortex	Drifting osteons on both cortices	Thin rind on cutaneous periosteal surface	Thinner, mostly lamellar bone with some remodeling, periosteal lamellar or woven bone, large resorptive bays (drifting osteons)	Thicker, denser remodeling, some areas of primary lamellar bone
Phase IV (18–21 years)	Both cortices periosteally	Both cortices, fewer drifting osteons, more Type I osteons	Rare	Thinner, dense remodeling, osteons 2–3 rows deep	Thicker, dense remodeling, osteons 3–4 rows deep, rare areas of primary lamellar bone
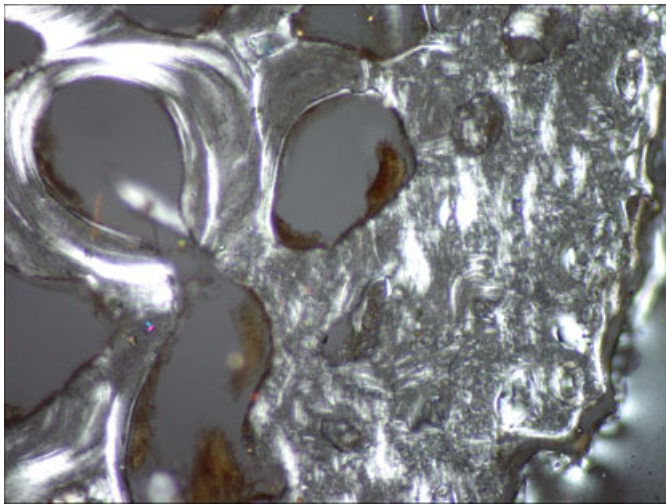
FIG. 8—*Phase I cutaneous cortex of woven bone with primary vascular canals in the rib of a 2 year old (200× polarized).*
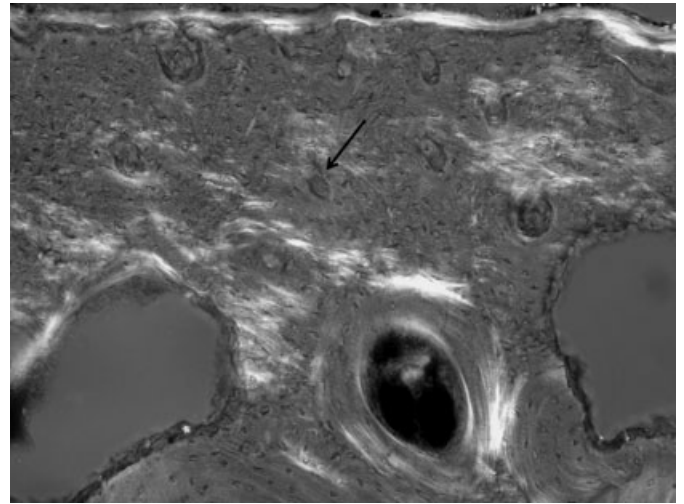
FIG. 10—*Phase II cutaneous cortex of unremodeled primary bone with many primary vascular canals (arrow) in the rib of a 6 year old (200× polarized).*
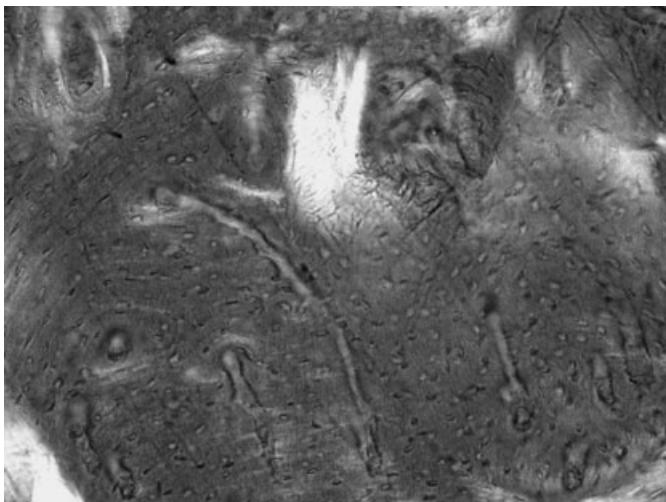
FIG. 9—*Phase I pleural cortex of woven bone with primary vascular and Volkmann's canals in the rib of a 3 year old (200× polarized).*
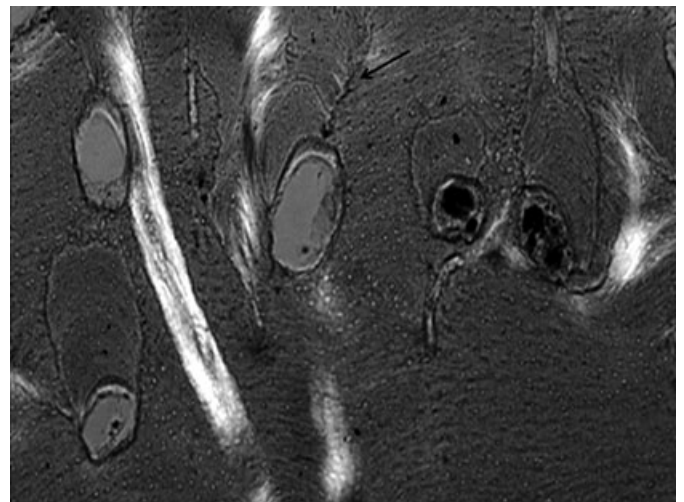
FIG. 11—*Phase II pleural cortex with forming drifting osteons (arrow) at periosteal surface in the rib of a 6 year old (200× polarized).*
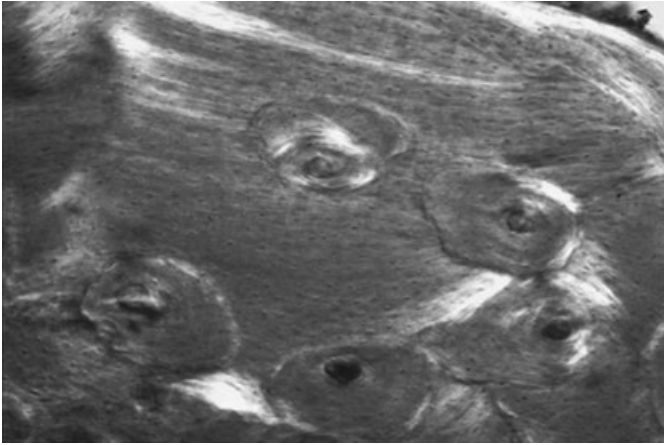
FIG. 12—*Phase III cutaneous cortex with drifting osteons and unremodeled primary bone at the periosteum (top) in the rib of a 17 year old (200× polarized).*
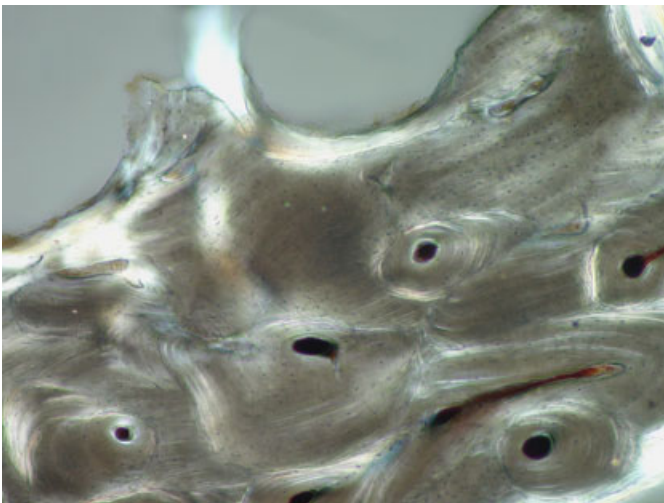
FIG. 13—*Phase III pleural cortex with dense remodeling (drifting osteons) and newly deposited primary bone at the endosteum, as cortical drift continues in the cutaneous direction in the rib of a 17 year old (200× polarized).*
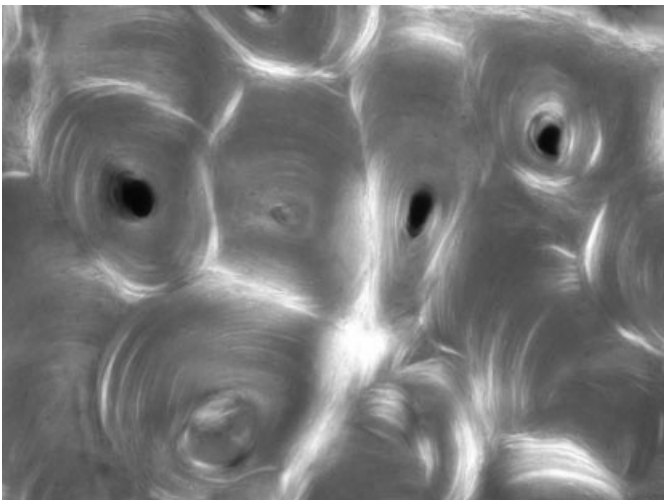
FIG. 14—*Phase IV cutaneous cortex densely remodeled by Type I osteons in the rib of a 19 year old (200× polarized).*
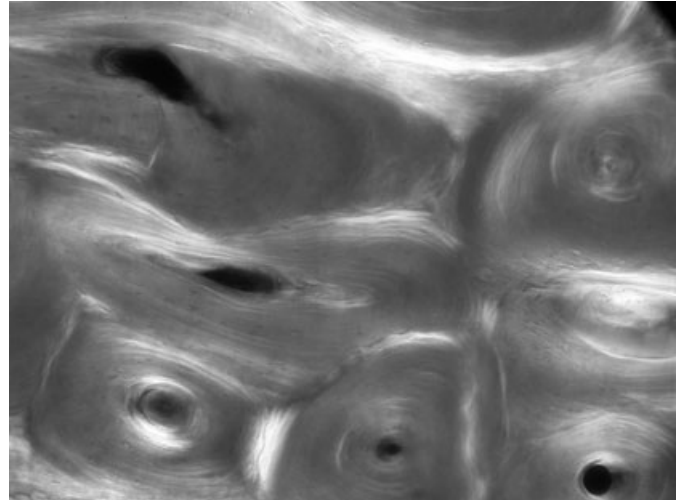
FIG. 15—*Phase IV pleural cortex of a 21 year old with Type I osteons (200× polarized).*

cortex. Scattered patches of natal woven bone may still be observed intracortically especially on the cutaneous cortex.

Phase III (10–17 years) (Figs. 12 and 13)

On both cortices, the primary lamellar bone is remodeled with drifting osteons often two or more rows deep (Fig. 13). A typical characteristic of this phase is the thin rind of primary lamellar or woven bone that can be seen along the cutaneous periosteal surface (Fig. 12). This unremodeled bone is often interspersed with primary vascular canals, a reflection of the accelerated cortical modeling typical of the increase in the growth rate associated with this age range.

Phase IV (18–21 years) (Figs. 14 and 15)

The two rib cortices have a more adult morphology with Type I osteons becoming more common (Figs. 14 and 15). The accumulation of Haversian systems has reached a density three or more rows thick, resulting in the creation of osteon fragments. Volkmann's canals may still be seen on unremodeled areas of the pleural cortex, but primary vascular canals are rare.

Discussion

Clearly defined and easily observable age-associated histomorphological patterns have been identified, which permit the seriation of subadult ribs into four developmental phases. There are several advantages to using this categorical classification system. First, this method uses the rib, which is often more readily accessible than other bones. Second, although this method still relies on histological cross-sections, it does not require the expertise of the other histomorphometric methods based on osteon counts. Each phase is defined by a group of easily identifiable characteristics that facilitate the application of this procedure. Third, this method is less time-consuming than performing osteon counts. This age estimation method identifies normal patterns of growth and development in the subadult rib cortex, providing a baseline of normal morphological characteristics useful for the future identification of pathological conditions.

This histological aging method has been tested on an archaeological subadult sample (23), and although the progression of morphological traits described earlier was observed, the predicted age

ranges did not correlate well with the skeletal ages predicted based on measures of diaphyseal length. This suggests that the archaeological population experienced the same sequence of developmental phases in the rib but at a different rate of growth relative to the modern population upon which this method is based. The differences in the timing of growth stages are possibly the result of genetic, nutritional, or other physiological or environmental influences. It is possible that the age ranges assigned to the phases described here are population-specific. This method requires testing on a sample of ribs from individuals of known age and from diverse populations to assess the age-related variability of these developmental phases.

References

1. Ahlqvist J, Damsten O. A modification of Kerley's method for the microscopic determination of age in human bone. J Forensic Sci 1969;14:205–12.
2. Cho H, Stout SD, Madsen RW, Streeter MA. Population-specific histological age estimating method: a model for known African-American and European American skeletal remains. J Forensic Sci 2002;47:12–8.
3. Ericksen MF. Histological estimation of age at death using the anterior cortex of the femur. Am J Phys Anthropol 1991;84:171–9.
4. Kerley ER. The microscopic determination of age in human bone. Am J Phys Anthropol 1965;23:149–64.
5. Kerley ER, Ubelaker DH. Revision in the microscopic method of estimating age at death in human cortical bone. Am J Phys Anthropol 1978;49:545–6.
6. Stout SD, Paine RR. Histological age estimation using rib and clavicle. Am J Phys Anthropol 1992;87:111–5.
7. Thompson DD. The core technique in the determination of age at death in skeletons. J Forensic Sci 1979;24(4):902–15.
8. Burton P, Nyssen-Behets C, Dhem A. Haversian bone remodeling in human fetus. Acta Anat 1989;135:171–5.
9. Baltadjiev G. Micromorphometric characteristics of osteons in compact bone of growing tibia of human fetuses. Acta Anat 1999;154: 181–5.
10. Streeter M. Histomorphometric characteristics of the subadult rib cortex: normal patterns of dynamic bone modeling and remodeling during growth and development [dissertation]. Columbia (MO): Univ. of Missouri, 2005.
11. Epker BN, Frost HM. The direction of transverse drift of actively forming osteons in human rib cortex. J Bone Joint Surg 1964;47A: 1211–5.
12. Sedlin ED, Frost HM, Villanueva BS. Age changes in resorption in human rib cortex. J Gerontol 1963;18:345–9.
13. Takahashi H, Frost HM. Correlation between body habitus and cross-sectional area of ribs. J Physiol Pharmacol 1965;43:773–81.
14. Takahashi H, Frost HM. A tetracycline-based evaluation of the relative prevalence and incidence of formation of secondary osteons in human cortical bone. Can J Physiol Pharmacol 1965;43:783–91.
15. Takahashi H, Frost HM. Age and sex related changes in the amount of cortex in normal human ribs. Acta Orthop Scand 1996;37:122–30.
16. Saunders SR. Juvenile skeletons and growth-related studies. In: Katzenberg MA, Saunders SR, editors. Biological anthropology of the human skeleton. New York, NY: Wiley-Liss, 2008;117–48.
17. Frost HM. Preparation of thin undecalcified bone sections by rapid manual method. Stain Technol 1958;33:273–7.
18. Landeros O, Frost HM. Comparison of amounts of remodeling activity in opposite cortices of rib in children and adults. J Dent Res 1966;45(1):152–8.
19. Brooks S, Suchey JM. Skeletal age determination based on the os pubis: a comparison of the Acsadi-Nemeskeri and Suchey-Brooks methods. Hum Evol 1990;5:227–38.
20. İşcan MY, Loth SR, Wright RK. Metamorphosis at the sternal rib: a new method to estimate age at death in males. Am J Phys Anthropol 1984;65:147–56.
21. Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. Am J Phys Anthropol 1985;68:15–28.
22. Robling AG, Stout SD. Methods of determining age at death using bone microstructure. In: Katzenberg MA, Saunders S, editors. Biological anthropology of the human skeleton. New York, NY: Wiley-Liss, 2000;187–205.
23. Agnew A, Streeter M, Stout SD. Histomorphological aging of subadults: a test of Streeter's method on a medieval archaeological population. Am J Phys Anthropol Suppl 2007;44:61.

Additional information and reprint requests:
Margaret Streeter, Ph.D.
Department of Anthropology
115 Hemingway WSC
Boise State University
Boise, ID 83725
E-mail: margaretstreeter@boisestate.edu

TECHNICAL NOTE

PHYSICAL ANTHROPOLOGY

Xanth D. G. Mallett,[1] *Ph.D.; Ian Dryden,*[2] *Ph.D.; Richard Vorder Bruegge,*[3] *Ph.D.; and Martin Evison,*[4] *Ph.D.*

An Exploration of Sample Representativeness in Anthropometric Facial Comparison*

ABSTRACT: Faces are assumed to be unique, but their use in court has remained problematic as no method of comparison with known error rates has been accepted by the scientific community. Rather than relying on the assumed uniqueness of facial features, previous research has been directed at estimations of face shape frequency. Here, the influence of age, sex, and ancestry on variation was investigated. Statistical shape analysis was used to examine the necessity for sub-divisions in forensic comparisons, using a large sample of facial images on which 30 anthropometric landmark points had been placed in 3D. Results showed a clear pattern of separation of the sexes in all age groups, and in different age groups in men. It was concluded that sub-division of databases by sex will be necessary in forensic comparisons. Sub-division by age may be necessary in men (although not necessarily in women), and may be necessary by ancestry.

KEYWORDS: forensic science, facial identification, biometrics, anthropometry, principal components analysis, shape analysis

It is a common assumption that faces are unique. They are ubiquitous and—as a potential means of human identification—surpass dermatoglyphic fingerprints, the dentition, and DNA in being remotely detectable. To be of value to the courts, however, a scientifically defensible approach to facial comparison is required which—rather than relying on an assumption of discernable uniqueness—will offer estimates of face shape frequency derived from large database studies (1–3). In previous research (4), the Magna database (5) was used to identify key components of normal face shape variation. In this study, the same database was used to investigate the influence of age, sex, and ancestry on variation. The representativeness of sub-categories of age and sex was examined in further detail, and the necessity for sub-division of databases in forensic facial comparisons was explored.

Materials and Methods

The Magna database (5) consists of facial images collected from healthy volunteers using a digital stereophotogrammetric system (Geometrix® FaceVision 802 Series Biometric Camera; ALIVE Tech, Cumming, GA). Volunteers were classified by age and sex, and ancestry as self-declared according to the United Kingdom Census categories (see Tables 1 and 2). A software tool (Forensic Analyzer® v.1.3; ALIVE Tech) was used to place up to 30 three-

[1]Centre for Anatomy and Human Identification, University of Dundee, Dundee DD1 5EH, U.K.
[2]Department of Statistics, University of South Carolina, Columbia, SC 29208.
[3]Forensic Video Audio and Image Analysis Unit, Federal Bureau of Investigation, Building 27958A, Quantico, VA 22135.
[4]Forensic Science Program, University of Toronto, Mississauga, ON L5L 1C6, Canada.

dimensional (3D) anthropometric landmarks on each facial image. The landmarks selected were those demonstrating maximum inter-subject variation relative to intra-subject variation (see Fig. 1 and Table 3), hence offering the greatest potential for the statistical evaluation of variation in size and shape. Each facial image was landmarked twice by different observers. Only landmark sets with all 30 landmarks present were included in this study (960 women and 2294 men, from 1968 individual subjects).

Patterns of variation in the sample were analyzed using statistical shape analysis (7). Procrustes registration was used to remove differences because of position and rotation. As size differences between subject faces represent real variation, Procrustes registration was undertaken without scaling. Principal components (PC) analysis was then used to decompose differences because of variation in size and shape in different sub-groups in the sample. Statistical programming was undertaken in R (8).

Results and Discussion

Plots of the first three PC scores underlying size and shape variation in men and women in 15–34 (Fig. 2) and 35–54 (Fig. 3) years of age ranges confirm earlier observations (4) of sexual dimorphism, and of age-related trends in face size and shape variation that affect men and women differently.

To assess whether the sample as a whole could be considered representative of the sub-categories of sex and age within it, a closer examination was made of age group categories. Individuals under 16 years of age were excluded to partly mitigate confounding effects of rapid changes in facial morphology affecting juveniles during development.

Analysis of the first two PC scores underlying variation by age group and sex (Fig. 4) in this sample again shows a clear pattern of separation of the sexes in all age groups. Separation increases with age and is underlain by the first PC score, which is attributed

TABLE 1—*Distribution of volunteers in the Magna database by sex and age.*

Age group	Females	Males
14–19	194	188
20–24	71	68
25–29	98	103
30–34	173	188
35–39	291	325
40–44	247	361
45–49	127	181
50–54	53	111
55–59	55	65
60–64	44	54
65+	41	77
Total	1394	1721

TABLE 2—*Distribution of volunteers in the Magna database by sex and ancestry.*

Census category	Females	Males
White British	1265	1553
Other white background	56	78
White and black Caribbean	4	3
White and black African	1	2
White and Asian	4	6
Other mixed background	6	3
Indian	14	17
Pakistani	4	11
Any other Asian background	6	11
Caribbean	2	8
African	7	7
Other black background	3	1
Chinese	18	17
Any other	4	4
Total	1394	1721

FIG. 1—*Illustration of the 30 landmark sites collected from volunteers using the Geometrix FaceVision® FV802 Series Biometric Camera. Bilateral landmarks are in green.*

TABLE 3—*Description of the 30 landmarks used in the study (and shown in Fig. 1).*

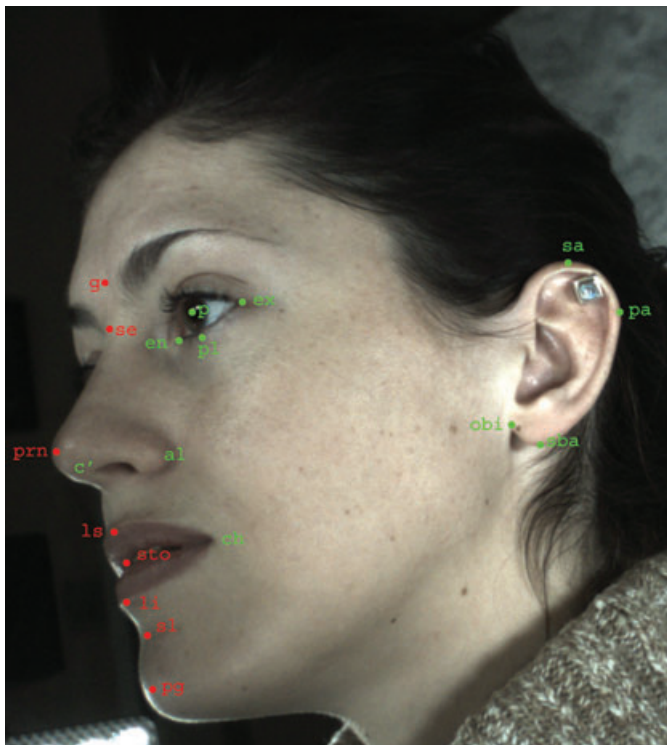| Label	Name	Description—after Farkas 1994 (6)
al	Alar (l, r)	The most lateral point on each alar contour
c'	Highest point of columella (l, r)	The point on each columella crest, level with the tip of the corresponding nostril
ch	Cheilion (l, r)	The point located at each labial commissure
en	Endocanthion (l, r)	The point at the inner commissure of the eye fissure
ex	Exocanthion (l, r)	The point at the outer commissure of the eye fissure
g	Glabella	The most prominent midline point between the eyebrows
li	Labiale inferius	The midpoint of the lower vermillion line
ls	Labiale superius	The midpoint of the upper vermillion line
obi	Otobasion inferius (l, r)	The point of attachment of the ear lobe to the cheek
p	Pupil (l, r)	Determined when the head is in the rest position and the eye is looking straight forward
pa	Postaurale (l, r)	The most posterior point on the free margin of the ear
pg	Pogonion	The most anterior midpoint of the chin
pi	Palpebrale inferius (l, r)	The lowest point in the mid-portion of the free margin of each lower eyelid
prn	Pronasale	The most protruded point of the apex nasi
sa	Superaurale (l, r)	The highest point on the free margin of the auricle
sba	Subaurale (l, r)	The lowest point on the free margin of the ear lobe
se	Sellion	The deepest landmark located in the bottom of the nasofrontal angle
sl	Sublabiale	Determines the lower border of the lower lip and upper border of the chin
sto	Stomion	The imaginary point at the crossing of the vertical facial midline and the horizontal labial fissure between gently closed lips, with the teeth shut in the natural position

to size (4). This size effect is also notable in men considered alone (Fig. 5), where the older age groups have a higher first PC score than the youngest group here. In women (Fig. 6), there is a general tendency for the first PC score to be higher for the older age groups here, but the trend is not remarkable and no clear differentiation between the youngest age group and the older groups is evident. It is tentatively anticipated that male faces in the sample complete the development to the adult form in the younger age group and that further changes as a result of aging are relatively subtle, whereas female faces retain a youthful morphology in adulthood.

To explore the effect of age sub-division in men and women separately, means and standard deviations of the first four PC scores were compared by 3-year age ranges for each sex (see Tables 4 and 5).

Results of the PC analyses (Figs. 2–7) indicate that sub-division of men in the database may be required for use in forensic facial comparison, although the implication is less evident for women. Examination of the means and standard deviations of the first four PC scores underlying face size and shape variation (Tables 4 and 5) may allow ages to be identified at which sub-divisions are most appropriate. For example, means and standard deviations in the first PC scores in men appear to stabilize after 30 years of age. The means and standard deviations in the next three PC scores that underlie size and shape variation do not become comparable at 30 years, however, and first PC score is attributed primarily to size. The significance of size in facial comparison is complicated in that
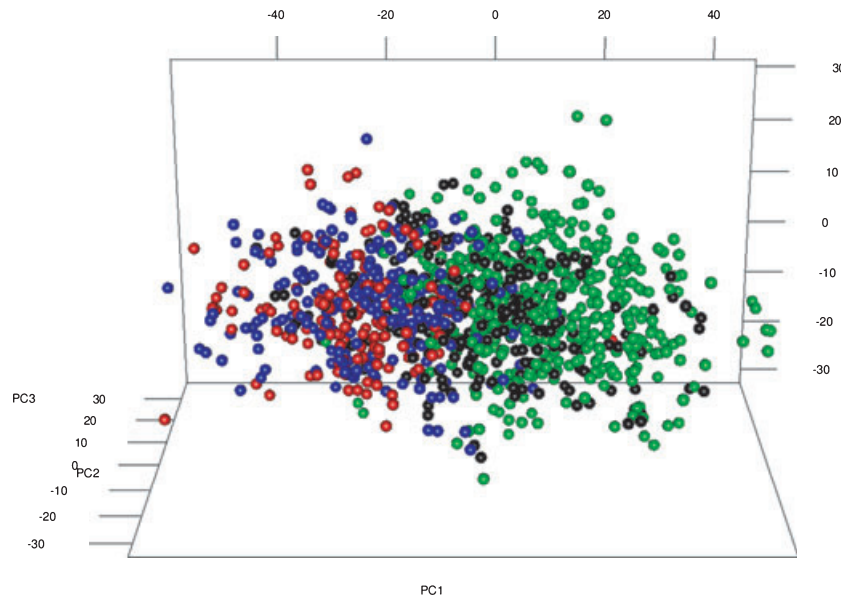
FIG. 2—*Principal components analysis of size and shape variation in males and females for age groups in lower age ranges. Males 15–24: black; females 15–24: red; males 25–34: green; females 25–34: blue.*
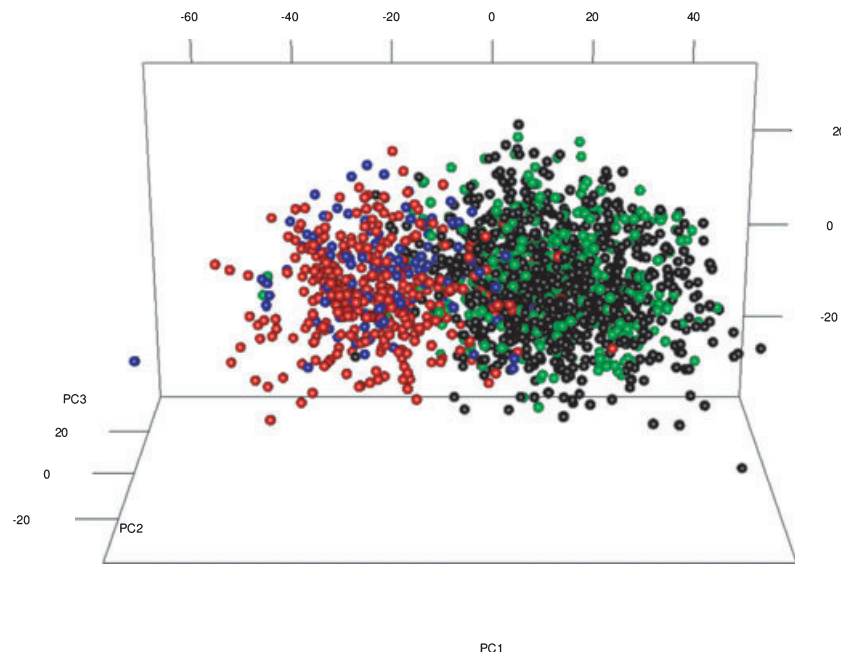

FIG. 3—*Principal components analysis of size and shape variation in males and females for selected age groups in upper age ranges. Males 35–44: black; females 35–44: red; males 45–54: green; females 45–54: blue.*

most contemporary questioned facial images are not of known size. By implication, in the event of a proposed match, estimation of the frequency of the given face shape must be independent of size.

In the PC analysis of men and women aged 35–54 (Fig. 3), two landmark datasets representing independent landmarkings of a single male individual were detected far from the main cluster of men and even on the edge of the female distribution (seen to the left of Fig. 3). This extreme outlier does not appear to be in error, but rather a self-declared male volunteer who may be a transgendered or intersex individual.

Finally, a PC analysis of size and shape variation attributable to ancestry (Fig. 7) reveals some trends affecting the United Kingdom

Census ethnic categories relative to each other, but not relative to the White British component, which predominates numerically in the sample (see Tables 2 and 6).

Conclusions

These results confirm that sex, age, and ancestry affect variation in face size and shape. PC analysis indicates that sub-division of databases by sex will be necessary in forensic facial comparison. Sub-division according to age group may be necessary in men (although not necessarily in women) and may be necessary according to ancestry—although a bias toward older age groups in the
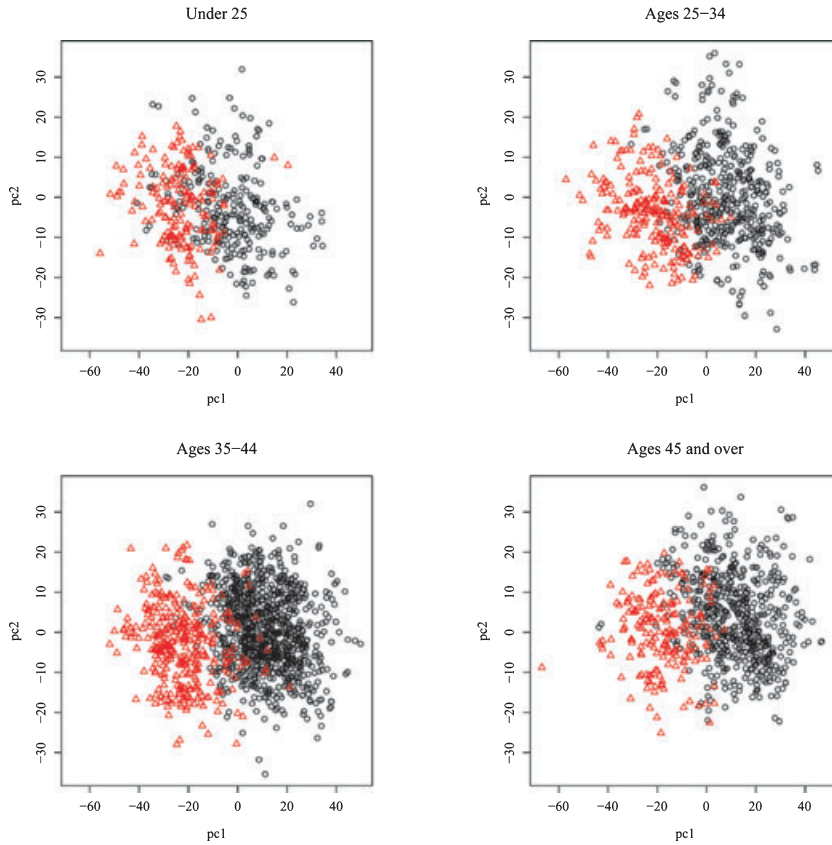
FIG. 4—*Pair wise plots of the first two principal components scores underlying face size and shape variation for males (black circle) and females (red triangle), in four different age groups.*
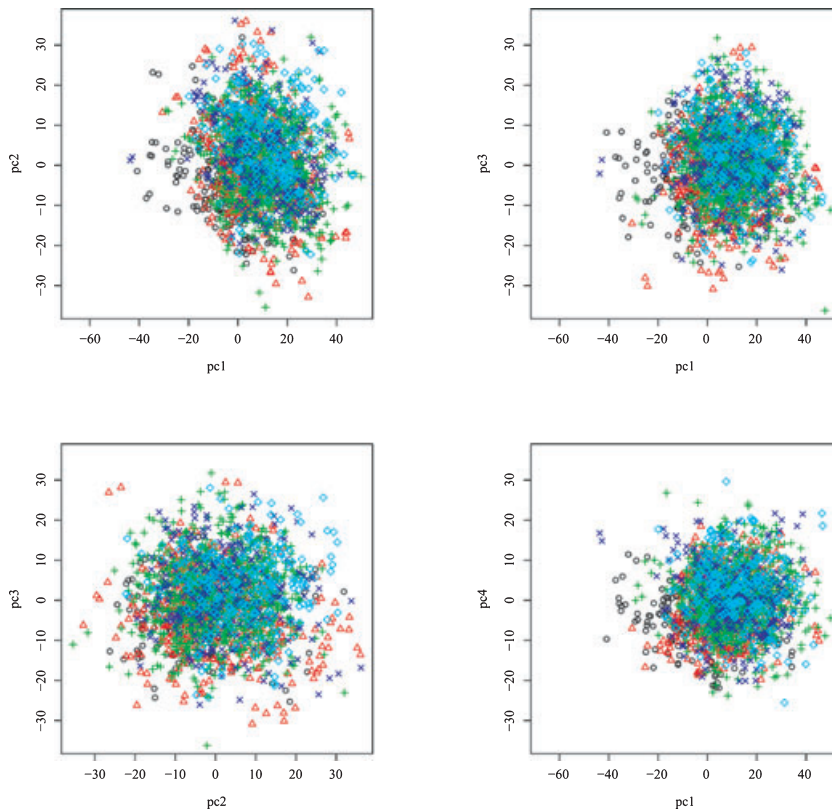

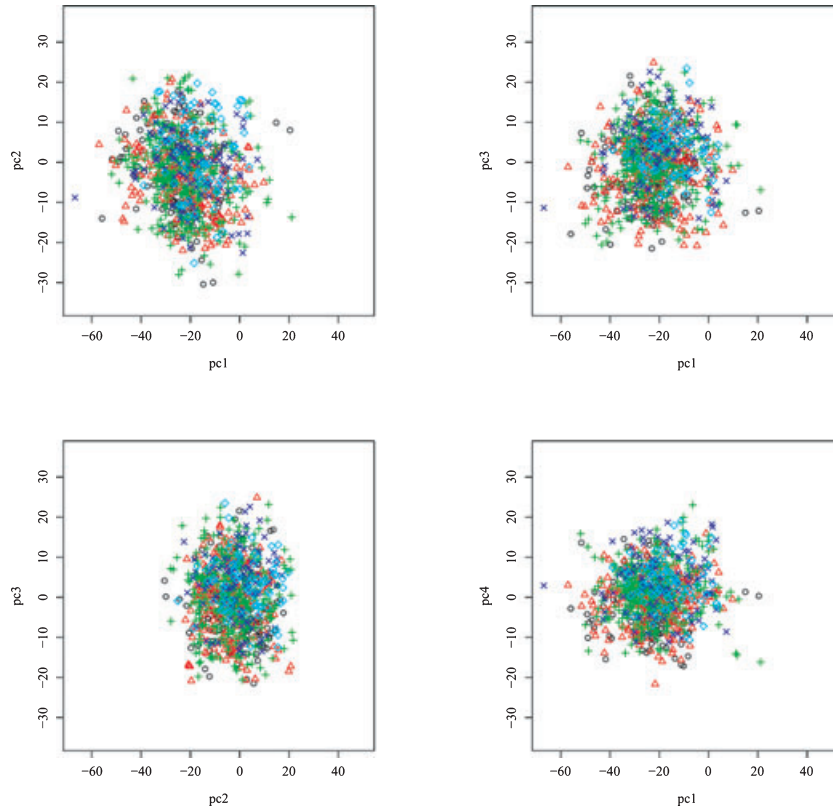FIG. 5—*Pair wise plots of the first four principal components scores underlying face size and shape variation in males for four different age groups: 15–24 years: black circle; 25–34 years: red triangle; 35–44 years: green '+'; 45–54 years: blue '×'; over 55 years: cyan diamond.*
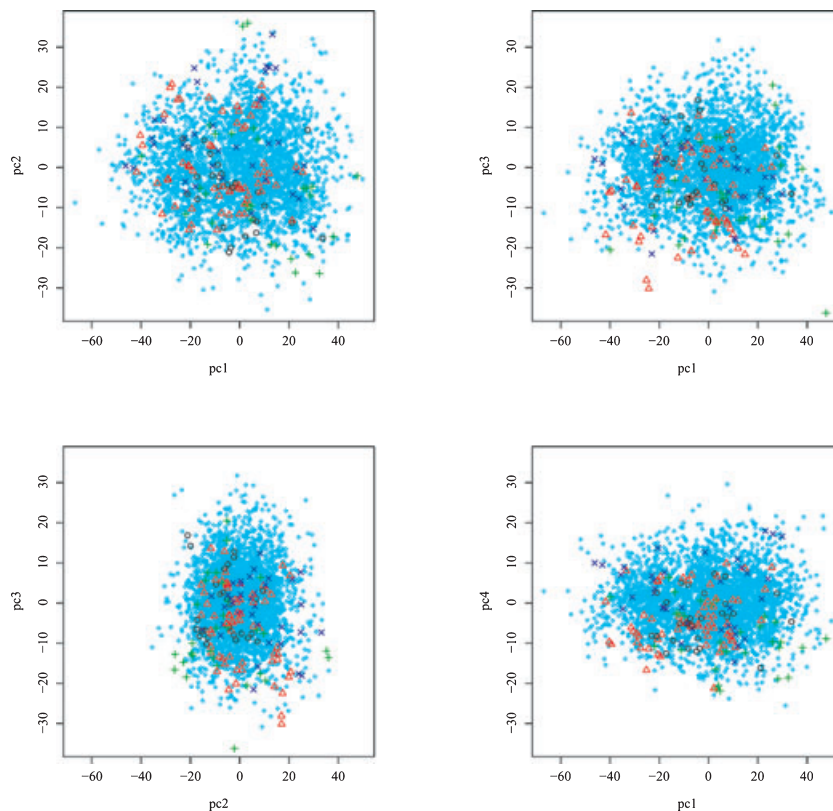
FIG. 6—*Pair wise plots of the first four principal components scores underlying face size and shape variation in females for four different age groups: 15–24 years: black circle; 25–34 years: red triangle; 35–44 years: green '+'; 45–54 years: blue '×'; over 55 years: cyan diamond.*

FIG. 7—*Principal components analysis of overall size and shape variation classified into five ancestry groups. White: small cyan diamond; Mixed Ancestry: black circle; Asian: red triangle; Black: green '+'; Chinese or other: blue '×.'*

TABLE 4—*Means and standard deviations (SD) for the first four principal components (PC) scores in men.*

Age	n	PC1		PC2		PC3		PC4	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
16–18	63	−5.7	13.2	−2.9	8.2	−1.9	7.7	−6.4	5.5
19–21	69	2.8	11.0	−1.9	10.6	−4.0	9.0	−6.3	7.5
22–24	69	2.6	12.8	−5.2	12.0	−3.3	8.4	−2.6	7.2
25–27	100	3.2	12.5	1.9	11.5	−1.4	9.5	−4.8	7.7
28–30	107	6.5	10.5	0.2	12.2	−5.0	11.1	−1.6	8.0
31–33	160	11.7	12.7	−1.4	12.0	−2.8	9.9	−0.5	7.2
34–36	277	10.4	11.5	−0.3	10.3	−0.4	9.6	0.1	7.6
37–39	292	11.0	12.3	0.5	9.2	0.7	8.7	0.0	7.0
40–42	316	11.6	12.6	0.6	10.4	0.2	9.0	0.2	7.5

TABLE 5—*Means and standard deviations (SD) for the first four principal components (PC) scores in women.*

Age	n	PC1		PC2		PC3		PC4	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
16–18	49	−25.6	9.2	−0.2	8.7	−1.7	9.6	−1.3	5.3
19–21	49	−21.6	13.2	−0.1	9.7	−4.0	8.2	−3.7	5.7
22–24	41	−23.0	9.1	−2.4	9.9	−1.8	8.9	−1.3	8.3
25–27	58	−21.9	11.7	−2.5	10.8	−4.5	8.7	−3.1	6.6
28–30	53	−23.1	11.9	−5.4	8.3	0.6	9.1	−1.4	5.8
31–33	70	−19.3	12.1	−4.1	8.3	−1.4	7.7	1.0	5.5
34–36	127	−21.5	12.6	−4.5	8.2	−0.2	7.8	0.2	6.8
37–39	111	−23.2	10.4	−3.4	8.4	0.0	8.5	1.3	6.8
40–42	120	−22.7	11.4	−3.0	10.9	0.8	9.3	1.0	5.8

TABLE 6—Means and standard deviations (SD) for the first four principal components (PC) scores by ancestry.

Ancestry	n	PC1		PC2		PC3		PC4	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
White	3096	0.2	19.0	0.0	10.3	0.2	9.2	0.2	7.5
Mixed	27	−2.7	14.3	−6.3	8.7	−0.8	8.4	−3.8	6.3
Asian	64	−7.3	16.3	0.6	10.3	−5.7	10.0	−4.1	6.8
Black	30	7.5	20.0	−5.6	15.4	−7.3	12.0	−8.5	7.6
Chinese	37	−5.3	21.2	6.7	11.6	−1.9	7.9	2.5	8.5

Magna database (see Table 1) means that sub-samples in the younger age groups were small in this analysis. Persons of non-White ancestry were similarly under-represented (see Table 2).

Caution may be necessary in sub-dividing databases according to sex, however, in view of the complexities of chosen categories of gender and intersex conditions.

One potential source of bias was also identified. Of the 3024 individuals in the sample, only datasets landmarked at all 30 possible points—originating from 1968 individuals—were subjected to PC analysis. Commonly excluded cases include women in which the landmarks of the ear are covered by hair, and men in which the landmarks of the mouth are obscured by beards and mustaches.

Further research using the Magna database will permit the elaboration of prototypic tools for anthropometric matching or exclusion of faces and face shape frequency estimation (9). Research on the Magna database may also permit the incorporation of anthroscopic features such as scars, moles, blemishes, and tattoos—as well as shape features between landmarks—into those tools.

Further data collection of facial images from volunteers of non-White ancestry and from younger individuals may be necessary to ultimately establish database sub-divisions required for forensic facial comparison.

References

1. Saks MJ, Koehler JJ. The coming paradigm shift in forensic identification science. Nature 2005;309:892–5.
2. Mardia KV, Coombes A, Kirkbride J, Linney A, Bowie LJ. On statistical problems with face identification from photographs. J Appl Stat 1996;23(6):655–76.
3. Evison MP. Anthropometry of the face: a review of the traditional methods of craniofacial measurement and their application to the anthropometry of photographic images. Proceedings of the Third UK National Conference on Craniofacial Identification; 2000 May; Manchester (UK). Manchester (UK): University of Manchester Department of Art in Medicine, 2000;7.
4. Evison MP, Dryden IL, Fieller NRJ, Mallett XGD, Morecroft L, Schofield D, et al. Key parameters of face shape variation in 3D in a large sample. J Forensic Sci 2010;55(1):159–62.

5. Evison MP, Vorder Bruegge RW. The Magna Database: a database of three-dimensional facial images for research in human identification and recognition. Forensic Sci Commun 2008;10(2), [Online], http://www.fbi.gov/hq/lab/fsc/backissu/april2008/research/2008_04_research01.htm (accessed April 16, 2009).
6. Farkas LG. Anthropometry of the head and face, 2nd edn. New York, NY: Raven Press, 1994.
7. Dryden IL, Mardia KV. Statistical shape analysis. Chichester, UK: Wiley, 1998.
8. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2009, http://www.r-project.org/ (accessed April 16, 2009).
9. Evison MP, Vorder Bruegge RW. Computer-aided forensic facial comparison: scientific and technical aspects. New York, NY: CRC Press/Taylor and Francis, 2010.

Additional information and reprint requests:
Martin Paul Evison, Ph.D.
Forensic Science Program
University of Toronto Mississauga
3359 Mississauga Road North
Mississauga, ON L5L 1C6
Canada
E-mail: martin.evison@utoronto.ca

TECHNICAL NOTE

PHYSICAL ANTHROPOLOGY

Esther J. Lee,[1] *Ph.D.; Jennifer G. Luedtke,*[1] *M.A.; Jamie L. Allison,*[1] *M.A.; Carolyn E. Arber,*[2] *M.S.;*
D. Andrew Merriwether,[1] *Ph.D.; and Dawnie Wolfe Steadman,*[1] *Ph.D.*

The Effects of Different Maceration Techniques on Nuclear DNA Amplification Using Human Bone

ABSTRACT: Forensic anthropologists routinely macerate human bone for the purposes of identity and trauma analysis, but the heat and chemical treatments used can destroy genetic evidence. As a follow-up to a previous study on nuclear DNA recovery that used pig ribs, this study utilizes human skeletal remains treated with various bone maceration techniques for nuclear DNA amplification using the standard Combined DNA Index System (CODIS) markers. DNA was extracted from 18 samples of human lower leg bones subjected to nine chemical and heat maceration techniques. Genotyping was carried out using the AmpFℓSTR® COfiler® and AmpFℓSTR® Profiler Plus® ID kits. Results showed that heat treatments via microwave or Biz/Na$_2$CO$_3$ in sub-boiling water efficiently macerate bone and produce amplifiable nuclear DNA for genetic analysis. Long-term use of chemicals such as hydrogen peroxide is discouraged as it results in poor bone quality and has deleterious effects on DNA amplification.

KEYWORDS: forensic science, forensic anthropology, maceration, DNA typing, nuclear DNA, short tandem repeat, CODIS markers

Forensic anthropologists typically use skeletal and dental features to establish personal identification, but DNA analysis of a bone sample may be warranted if the results are ambiguous, adequate antemortem records are unavailable, or the presumptive identification requires confirmation. However, DNA begins to degrade soon after the organism's death and samples will exhibit varying degrees of degradation depending on the period of time lapsed and the surrounding environment. Moreover, different processing techniques for soft tissue removal have been shown to have varying effects on DNA degradation (1–4). Degraded DNA strands are fragmented and less likely to be successfully utilized in forensic DNA analysis, especially for markers that target larger loci.

Samples with highly degraded DNA may often have less quantity of DNA. We use the term DNA *quantity* when referring to the actual amount of DNA recovered from a sample, while *quality* refers to the degree of degradation of endogenous DNA. Therefore, both good quality and sufficient quantity of DNA will ensure greater success in DNA recovery for amplification in forensic typing. The difficulty in quantifying DNA for forensic purposes is multilayered. Foreign contaminants such as bacterial and fungal DNA usually reside within the bone sample to be analyzed and are likely to be retained during the extraction process, even after maceration. Quantification via ultraviolet/visible (UV/VIS) spectrometry is relatively inexpensive and easily accessible but will not adequately measure endogenous DNA for forensic analysis because it does not differentiate between human and nonhuman DNA (5).

Recent techniques have improved the discrimination of human DNA from other contaminants using radioactive probes, fluorescence labeling, or via real-time polymerase chain reaction (PCR) (5,6), but these often require special equipment and are costly.

DNA typing in samples that are aged and degraded has had a higher success rate for mitochondrial DNA (mtDNA) than nuclear DNA. This success is often attributed to the high copy number of mtDNA, as it has been estimated that mammalian cells contain 80–680 mitochondria, dependent on cell type, providing about 2.6 DNAs per mitochondrion (7). In addition to gene copy number, it has been suggested that other factors such as cellular location, chromatin structure, and transcriptional activity influence the difference between mtDNA and nuclear DNA degradation (7).

DNA analysis using bone samples has been successful in anthropological studies of ancient populations as well as various forensic cases. Using the mtDNA control region in ancient DNA studies has proven useful to infer past human population histories in various regions across the world (8–10). Newer techniques have even allowed analyses of older skeletal remains such as the Neanderthal and the utilization of whole genomes (11–13). With regard to forensic cases, studies have shown positive results in extracting and amplifying mtDNA and nuclear DNA from skeletal remains (e.g., 14,15). Nuclear DNA is the preferred genome of amplification for forensic purposes as it provides bi-parental kinship information and is individually specific. However, the condition of skeletal remains analyzed for forensic studies is often not ideal for DNA recovery, which may increase the difficulty of amplifying nuclear DNA. For this reason, recent studies have sought different methods of amplifying nuclear DNA from samples that have been heavily degraded because of age or environmental conditions, such as reducing the amplicon size (16–18). While the development of such techniques will enable greater success of DNA analysis, additional challenges

[1]Department of Anthropology, Binghamton University, SUNY, PO Box 6000, Binghamton, NY 13902-6000.
[2]Department of Anatomic Pathology, Hartford Hospital, 80 Seymour Street, Hartford, CT 06102.

may be imposed by the different techniques employed for bone maceration during forensic and anthropological studies.

A few studies have assessed the viability of DNA in bones that have been subjected to different treatments. One study looked at the effect of different maceration techniques employed on pig ribs utilizing three mtDNA and three nuclear loci, ranging in size between 268 and 750 base pairs (bp) (1), while another used two nuclear short tandem repeat (STR) markers to study the effect of different bone processing techniques (3). Others have also evaluated the success rate of extracting DNA from human bone after treating them with different concentrations of bleach (2,19). Overall, these studies have shown that over-exposure to heat, long treatment duration, and harsh chemicals have deleterious effects on DNA amplification.

This study presents the results of nuclear DNA amplification using all 13 genetic markers from the FBI Laboratory's Combined DNA Index System (CODIS) tested on human bone treated with nine different maceration techniques. We focus on the recovery of DNA based on nuclear DNA amplification. In other words, we determine the viability of DNA in terms of *amplifiable* DNA that may have been degraded and/or exist in small quantity because of the different maceration techniques employed. Arismendi et al. (3) previously utilized two nuclear loci, amelogenin and CSF1PO, the latter being one of the 13 CODIS markers and suggested that prolonged exposure to heat may decrease the size and quality of DNA. However, no other study has used the full complement of CODIS markers with a variety of maceration techniques. STR markers are commercially available as kits and have been shown to be suitable even with heavily decomposed material (20,21). In particular, the CODIS markers have been widely established as a common application in forensic cases for purposes of identification (22). The results from this study provide guidance to forensic anthropologists who need to macerate bone to conduct their analyses but must also retain the ability to extract and amplify DNA for genetic analysis.

Materials and Methods

Materials

Four amputated lower legs of four different individuals and two lower legs from one individual were obtained with appropriate IRB permissions from a local hospital. The amputations occurred at the distal femur such that the fibulae and tibiae were present in their entirety. Each long bone was sectioned near the midpoint to increase the sample size. The distal femora were mechanically macerated and served as the controls. The feet, skin, and muscle bellies were removed prior to conducting the experimental macerations.

Maceration Methods

The nine maceration protocols employed are described in Table 1 and only differ from those published by Steadman et al. (1) in that the room temperature procedure is not included in the current study. All solutions used untreated tap water. Heated solutions were placed on a hot plate or within a chaffing dish and the temperature was closely monitored. It was especially important that the temperature of the EDTA/papain solution remained below 45°C to avoid denaturing the papain enzyme. Each experiment was repeated on one tibia and one fibula from different individuals.

The duration of experimental exposure to chemicals and/or heat as well as the temperature of the solutions were strictly controlled. The time and date of initial submergence was noted, as was the time of all subsequent observations and the time and date of the conclusion of each experiment (when the bones were void of soft tissue). Temperatures were also recorded at each observation period. Solutions for heating experiments lasting longer than 1 day were changed daily while room temperature solutions were changed weekly. While the fresh solutions were heated to the desired temperature, the specimens were rinsed under tap water and any loose tissue was gently pulled free using fingertips or plastic forceps. The specimens were out of the solution for a maximum of 5–10 min during solution refreshment, which was not subtracted from the total duration time of the experiment (from immersion to final extraction). The specimens placed in the microwave were observed after each 60-sec heating interval.

Table 2 describes the variables observed and the scoring procedure. Three variables—odor, texture of the soft tissue, and the ease with which flesh could be removed from the bone—were independently scored before initial submergence and at regular intervals during the experiments. Scoring took place hourly for the first 2 days and then at least five times a day thereafter. Odor refers to

TABLE 1—*Maceration methods applied (modified from Steadman et al. 2006).*

Experiment	Protocol
Mechanical removal (Control)	Distal femora manually defleshed using scalpels and forceps to remove meat and strip the periosteum
Hot-water bath	Water bath (~10.5 L) maintained at or just below 90°C
Boiling	Water bath (~10.5 L) maintained around 100°C
Microwave	Specimens placed in microwave-safe dish and covered with plastic wrap; samples heated in a 1300-watt, 2450 MHz microwave oven (Sears, Hoffman Estates, IL) on high power for 1-min intervals until all flesh easily slipped from the bones
Bleach (Sodium hypochlorite)	10% bleach (Clorox, Oakland, CA) solution (1.05 L liquid bleach and 9.45 L water) kept at room temperature (~22°C)
Hydrogen peroxide (H_2O_2)	1.0 L 3% hydrogen peroxide (Cumberland Swan, Smyrna, TN) in 9.5 L water maintained at room temperature (~22°C)
EDTA/Papain	2 ¾ teaspoon (11.25 g) EDTA (Fisher Scientific, Fair Lawn, NJ) and 2 ¾ teaspoon (13.6 g) papain (Sigma, St. Louis, MO) per 10.5 L water maintained below 45°C
Meat Tenderizer/Palmolive®	Six teaspoons (39.4 g) Adolph's® (Lipton, Englewood Cliffs, NJ) nonseasoned meat tenderizer and six teaspoons (29.6 mL) Palmolive® (Colgate/Palmolive, New York, NY) per 10.5 L water maintained at or below 90°C
Detergent/Sodium carbonate (Biz/Na_2CO_3)	Seven tablespoons (100 cc) powdered Biz® (Redox Brands, Inc., West Chester, OH) and seven tablespoons (100 cc) powdered sodium carbonate (Arm and Hammer Super Washing Soda™, Church and Dwight Co., Inc., Princeton, NJ) per 10.5 L water maintained at or below 90°C
Detergent/Sodium carbonate + degreaser	Biz/Na_2CO_3 followed as mentioned earlier, then rinsed and placed in 300 mL liquid sudsy ammonia and 4 L water

TABLE 2—*Maceration scoring system (following Steadman et al. 2006: [12,13]).*

Odor	Soft Tissue Texture	Ease of Soft Tissue Removal	Bone Quality
1. Strong smell that permeates the entire laboratory space	1. Soft tissue is firm and/or quite solid; may feel tougher or more rubbery than when first submerged	1. Adherence to bone is quite strong with little or no flesh removal possible without damaging the bone	1. Brittle, fragile, easily broken
2. Moderate smell in the immediate vicinity of the experiment	2. Soft tissue is as malleable as when originally observed	2. Adherence to bone is moderately strong although large portions can be easily removed; the core of flesh close to the bone is still adherent	2. No cortical erosion but bone is lighter in weight and porous
3. Little to no smell; a slight odor possible around or under the fume hood	3. Soft tissue is considerably softer and looser than when the experiment began; very malleable	3. Adherence is minimal to none as flesh falls off as bones are removed from solution or easily removed with fingertips	3. Softer, more pliable than normal bone but no cortical damage
	4. Soft tissue is nearly liquefied and floats on the surface with little or no connection to bone		4. Cortex eroding and/or flaking but bone will not easily fracture
			5. Strong, normal bone texture and quality

the pungency of the flesh as well as any chemicals and its detectability within the laboratory (e.g., overwhelming the room or only at close distance). Soft tissue texture was documented prior to experimentation and then subsequently observed for the degree of softening (malleability) prior to and including its final removal. The flesh was manipulated by forceps or fingertips to remove any loose tissue during each observation. The ease at which tissue was pulled, pushed, or slid from the bone was then scored. Following final extraction, the bones were dried for a minimum of 72 h and then assessed for bone quality. Scores lower than the starting score indicate that the bone was compromised by the experiment (e.g., brittle, porotic, or softened).

Contamination Control

All molecular procedures were carried out in a sterile facility with separate rooms dedicated to drilling, extracting, and preparing samples for PCR. All rooms in the facility are equipped with high efficiency particulate (HEPA) filtered air and have UV lights over all surfaces. Airflow between the rooms is limited by magnetic interlocks on the doors and by maintaining a serial positive pressure. Disposable gowns, gloves, shoe covers, facial masks, and goggles are required to enter this facility. Exposed surfaces were wiped down with bleach and UV irradiated before each procedure. Genotype information for all researchers that came into contact with the samples at any time during the maceration and molecular procedures were obtained, and negative controls were included during extraction and each PCR.

Molecular Methods

All bones had been stored in −80°C prior to extraction. The outer surfaces of the bone were briefly wiped with bleach before drilling to produce c. 0.1 g of bone powder under a sterile fume hood. Following the FBI bone extraction protocol (23), DNA was extracted from each of the untreated control samples and treated bone samples by decalcification with EDTA, adding Proteinase K and extraction buffers, and followed by a series of alcohol washes using QIAamp® Spin Column (Qiagen, Valencia, CA). DNA concentrations were estimated for the extractions by UV absorbance measurements at 260 nm using SpectraMax®190 microplate spectrophotometer and SoftMax® Pro 4.8 following manufacturer's instructions (Molecular Devices, Sunnyvale, CA).

The AmpF𝓁STR® COfiler® kit and the AmpF𝓁STR® Profiler Plus® ID kit (Applied Biosystems, Foster City, CA) were used to amplify the samples for STR loci, which include the sex identification marker amelogenin and the 13 CODIS markers, D3S1538, vWA, FGA, D16S539, D8S1179, D21S11, D18S51, TH01, TPOX, CSF1PO, D5S818, D13S317, and D7S820. Seven markers are included in the COfiler® kit and ten in the Profiler Plus® ID kit with three overlapping markers between the two kits (D3S1358, D7S820, and amelogenin). It was necessary to use both kits to obtain information for all 13 CODIS markers. PCR amplification for each kit followed the manufacturer's standard protocol containing: 3.8 μL of AmpF𝓁STR® PCR Reaction Mix, 2 μL of AmpF𝓁STR® Primer Set, 0.2 μL of AmpliTaq Gold® DNA Polymerase, 2 μL of DNA template, and sufficient ddH2O to adjust the reaction to a total volume of 10 μL for each reaction (24,25). Thermal cycling was performed on a GeneAmp® PCR System 9700 (PE Applied Biosystems), following the conditions outlined in the manufacturer's protocol: 95°C for 11 min; 28 cycles of successive 1-min holds at 94, 59, and 72°C; 45-min hold at 65°C (24,25).

Data collection was carried out by first diluting amplified products with 30 μL ddH2O, then combining 2 μL of the diluted product with 0.5 μL of GeneScan™ 500 ROX™ (Applied Biosystems) size standard. The combined samples underwent standard alcohol precipitation before being analyzed on an ABI PRISM™ 377 DNA Sequencer (Applied Biosystems), following manufacturer's instructions. All genotyping was performed with ABI PRISM™ GeneScan® Analysis 3.1.2 (PE Applied Biosystems). We used a minimum peak height of 100 relative fluorescence units (RFUs) to detect heterozygous alleles and 300 RFUs for homozygous alleles.

Results

Maceration Results

Table 3 reports the total temporal duration of each experiment as well as the average score for odor, texture, ease of soft tissue removal, final bone quality, and OD_{260} concentration values for each specimen (fibula or tibia) in the nine experiments. The experiments are listed in Table 3 in order of increasing temporal duration. The OD_{260} concentration values are discussed further in the molecular results section.

As expected, the microwave and other procedures that involved heating to 90°C or above cleaned bones faster (in a matter of minutes or hours) than those set at lower temperatures (1). The

TABLE 3—*Maceration results in increasing order of temporal duration of experiment.*

Experiment	Sample	Start Weight (g)	Time*	N Obs[†]	Odor	Texture	Ease	Bone Quality	OD$_{260}$ Concentration[‡]
Microwave (high)	Tibia	164	34 min	34	2.3	1.2	1.2	5	117.3
	Fibula	66	38 min	38	2.3	1.0	1.4	5	46.5
Biz/Na$_2$CO$_3$	Tibia	184	6 h	6	3.0	2.5	2.3	5	22.7
(90°C)	Fibula	66	5 h	4	3.0	2.5	2.3	5	2.32
Meat tenderizer (Adolph's)/Palmolive®	Tibia	210	14 h	16	3.0	1.1	1.3	5	8.4
(90°C)	Fibula	66	12 h	13	3.0	1.0	2.0	4	3.04
Boiling (100°C)	Tibia	283	14 h	16	3.0	1.3	1.1	5	4.9
	Fibula	86	19 h	21	3.0	1.9	1.9	5	53.5
Hot water (90°C)	Tibia	522	42 h	46	3.0	1.4	1.4	5	18.7
	Fibula	70	20 h	23	3.0	1.0	1.9	5	1.7
EDTA/Papain (45°C)	Tibia	366	4 d	29	2.5	2.1	1.3	5	27.2
	Fibula	92	5 d	32	3.0	2.7	1.0	5	5.8
Bleach (22°C)	Tibia	480	8 d	7	3.0	3.0	1.0	4	16
	Fibula	112	5 d	6	3.0	2.7	1.7	4	7.2
Biz/Na$_2$CO$_3$ degreased	Tibia	370	16 d	11	3.0	2.0	2.0	2	1.7
(90°C)	Fibula	66	8 d	10	3.0	2.0	2.1	4	0.9
Hydrogen peroxide (22°C)	Tibia	204	112 + d	7	2.0	3.0	1.0	5	5.8
	Fibula	92	121 + d	35	2.1	3.0	1.0	5	21.4

*d, days; h, hours; m, minutes.
[†]Total number of observations.
[‡]Optical density measured via UV absorbance at 260 nm, shown in ng/μL.

sub-boiling Biz/Na$_2$CO$_3$ and meat tenderizer/Palmolive® treatments were faster than boiling without chemicals, although all still took <24 h. Further, the heat-treated samples generally had better bone quality scores than samples that were processed for days, although bone quality overall was quite good. Degreasing bones with ammonia following maceration by Biz/Na$_2$CO$_3$ took much longer (8–16 days) and resulted in lower bone quality scores (weakened bone). Bones macerated in the bleach solution (5–8 days) exhibited whitening and some cortical flaking but overall the bone quality was good. The high bone quality scores for the hydrogen peroxide treatment are somewhat misleading in that the experiments were terminated before the tissue was entirely removed, so the full extent of any cortical bone change could not be fully appreciated. This is clearly not an efficient maceration treatment and, as demonstrated in the following text, offered the poorest results in terms of nuclear DNA amplification.

For most of the experiments, the tissue remained tightly adhered to the bone until all of the tissue came off at about the same time, resulting in relatively low scores for ease of flesh removal. This is especially true for the experiments that took only hours or a few days. Soft tissue texture also reflected this trend. Most experiments were not particularly odiferous, although the smell of the hydrogen peroxide-soaked specimens became overwhelming near the end of the multimonth experiment. The odor observed in the microwave method was from the flesh being cooked prior to removal from the bone.

Molecular Results

All control sample genotypes were distinct from any of the researchers and negative controls did not produce any amplicons, indicating there was no contamination introduced during the extraction and PCR process. Table 4 shows the genotype information for the five control samples. All amplification reached the minimum RFUs as indicated previously. Overall, the fluorescence signal was weaker for larger amplicons in both the control and experimental samples, even in the positive controls obtained from the researchers. Some larger loci, such as D18S51 and D7S820, for which allele sizes range from 260 to 306 bp, failed to produce detectable

TABLE 4—*Genotype information for control samples using 13 CODIS markers and the amelogenin locus. Amplicon size range is included under each marker name (except for the amelogenin marker).*

Locus/Sample*	1	2	3	4	5
D3S1358 (115–143 bp)	16,17	15,17	16,18	15,15	17,18
vWA (157–197 bp)	18,20	17,18	17,17	16,17	18,19
FGA (222–268 bp)	21,25	23,25	21,22	20,21	19,22
D16S539 (234–274 bp)	11,12	12,12	11,11	11,13	12,13
Amelogenin	107/113	107/113	107/113	107/113	107/107
D8S1179 (138–170 bp)	10,15	11,12	11,13	11,13	8,11
D21S11 (188–244 bp)	26,28.2	28.2,3	28.2,33	27,29.2	29.2,32
D18S51 (274–306 bp)	15,16	16,16	(−)[†]	12,16	15,15
TH01 (170–190 bp)	6,7	6,9	7,7	9,9	6,7
TPOX (219–247 bp)	8,8	8,11	8,8	9,11	11,11
CSF1PO (282–318 bp)	11,13	12,13	12,12	12,12	12,12
D5S818 (136–172 bp)	11,11	12,12	11,13	9,11	11,13
D13S317 (208–236 bp)	9,11	11,12	8,9	11,14	11,11
D7S820 (260–296 bp)	10,12	10,11	11,11	(−)	(−)

*Randomly assigned numbers for control samples.
[†](−), amplification failure.

signals in two of the control samples amplified using the Profiler Plus® ID kit and one control sample using the COfiler® kit. Allele sizes for all treated samples were compared with the original control sample to verify the authenticity and no discrepancies were found.

UV spectrometry is said to provide reliable results for concentrations of at least 3 ng/μL (26). The lowest concentration values that included samples treated with Biz/Na$_2$CO$_3$ followed by degreasing,

the fibula treated with hot water, and the fibula treated with just Biz/Na$_2$CO$_3$ do not meet that threshold (Table 3). Based on the manufacturer's recommendation (24,25), a concentration of 1.0–2.5 ng is the optimum amount of input DNA for the two kits used in this study. This suggests that among the samples listed in Table 3, those with concentration values above 3 ng/µL are sufficient for PCR amplification. However, because this method does not differentiate human and nonhuman DNA, we cannot determine the actual quantity of endogenous DNA. In addition, this measurement does not indicate the degree of degradation, as a high quantity of DNA will show a high absorbance value even if they exist in shorter fragments. Therefore, we cannot predict the success rate of STR amplification from these estimates. Recent techniques using radioactive probes, fluorescence labeling, and real-time PCR have shown success in measuring the quantity of human DNA but could not be performed in this study because of lack of resources and institutional restrictions. Our discussion focuses on the results of STR amplification, which is predicated on good *quality* as well as sufficient *quantity* of DNA.

We have summarized the amplification results of each experimental treatment for both kits in Table 5. Temporal duration for each treatment is included with the results. STR amplification was most successful for the treatment by microwave and Biz/Na$_2$CO$_3$ without degreasing, which showed complete amplification at all loci. These treatments required the shortest amount of time. More than 50% of the STR markers successfully amplified from samples macerated with bleach, EDTA/papain, boiling water, meat tenderizer/Palmolive®, and Biz/Na$_2$CO$_3$ followed by degreasing. Overall, the Profiler Plus® ID kit was less successful than the COfiler® kit, and negative amplification was seen mostly among markers of larger-sized loci.

Treatments using a hot-water bath and hydrogen peroxide were the least successful in STR amplification. Samples macerated in a hot-water bath amplified successfully only at one and two of the smallest loci (amelogenin and D3S1358) for each kit. Amplification was negative for all STR markers analyzed in the samples that had

been treated with hydrogen peroxide. These latter samples also required the longest treatments.

Discussion

The maceration results are similar to those achieved previously with pig ribs (1), with a few caveats. Human leg bones soaked in bleach exhibited better bone quality scores than bleach-treated pig ribs, but this is likely because of the fact that the pig ribs took on average 33 days to macerate, while the human leg bones were clean in 5 days. Conversely, the additional step of degreasing the bones with ammonia did not significantly increase the exposure time in pig ribs, but required 8–16 more days in humans. Thus, the greater cortical thickness and marrow carrying capacity of human long bones appear to be a mitigating factor in degreasing. The experiments on the human leg bones did not seem to be as malodorous as the pig bones except for the hydrogen peroxide treatment, which scored lowest for both studies. The microwave method also differed between studies. While the soft tissue easily loosened and slipped from the pig ribs when put in the microwave with a small amount of water, the tissue was much more resistant in the human samples. This is most likely because of the amount and density of the soft tissue on the leg bones, such that prolonged exposure in the microwave resulted in burning of the soft tissue, particularly at the ends, rather than soft tissue retraction and release from the bone.

Comparison of the amplification results for COfiler® and Profiler Plus® ID kits show that the latter consistently yielded weaker amplification for all samples. For example, bones macerated with EDTA/papain were successfully amplified for the COfiler® kit but were successful only 50% of the time for the Profiler Plus® ID kit. Alleles amplified with the Profiler Plus® ID kit were smaller than 220 bp. Similarly, samples treated in boiling water were less successfully amplified with the Profiler Plus® ID kit than the COfiler® kit. Even among some control samples, loci that were larger than ~230 bp (FGA, D18S51, D7S820) were not detected with the Profiler Plus® ID kit, while the same samples amplified successfully using the COfiler® kit for loci that were similar or larger in size. In particular, the locus D7S820 was successfully amplified for all control samples using the COfiler® kit but was negative in some samples for the Profiler Plus® ID kit. This difference may be attributed to the older age of the Profiler Plus® ID kit such that the primers may have lost their efficiency. This view is supported by an independent observation of PCR amplification using cheek swab samples, which showed similar results. Thus, bone samples that were macerated by EDTA/papain or by boiling water may have retained more endogenous amplifiable DNA than suggested, based on the contrasting results between the COfiler® and Profiler Plus® ID kits.

The purpose of this study was to examine the effectiveness of popular maceration methods and the impact of these techniques on nuclear DNA amplification and final bone quality. This study shows that microwaving or using Biz/Na$_2$CO$_3$ for maceration had the highest success rate of STR amplification, suggesting these techniques have the least deleterious effect on the ability to recover and amplify nuclear DNA. This is consistent with the previous study by Steadman et al. (1), who successfully amplified DNA using primers that created larger amplicons from bone macerated with the same treatments. These treatments not only produced successful DNA amplification, they also required the shortest amount of time (<1 h for microwaving and 5–6 h for Biz/Na$_2$CO$_3$) and maintained good bone quality. Furthermore, treatments with meat tenderizer/Palmolive® or boiling water solutions macerated bones

TABLE 5—*Results of STR analysis for each maceration technique. The experiment duration is again reported in minutes (min), hours (h), or days (d) in parentheses.*

Treatment	Fibula COfiler®/Profiler Plus® ID*	Tibia COfiler®/Profiler Plus® ID
Microwave (high)	+/+ (38 min)	+/+ (34 min)
Biz/Na$_2$CO$_3$ (90°C)	+/+ (5 h)	+/+ (6 h)
Meat tenderizer/Palmolive® (90°C)	+/5 (12 h)	4/5 (14 h)
Boiling (100°C)	5/5 (19 h)	+/5 (14 h)
Hot water (90°C)	2/2 (20 h)	−/2 (42 h)
EDTA/Papain (45°C)	+/6 (5 d)	+/4 (4 d)
Bleach (22°C)	+/7 (5 d)	5/+ (8 d)
Biz/Na$_2$CO$_3$ degreased (90°C)	2/− (8 d)	+/+ (16 d)
Hydrogen peroxide (90°C)	−/− (121 + d)	−/− (112 + d)

*number of loci successfully amplified, (+), all loci successfully amplified, (−), no successful amplification. COfiler® kit contains seven loci and Profiler Plus® ID kit contains ten.

relatively quickly (14–19 h), maintained overall good bone structure, and achieved a relatively good success rate of STR amplification of around 50%. Maceration by hot water alone had the poorest amplification rate among the heating methods, although good bone quality was maintained. The duration required for this treatment (20–42 h) was similar to or higher than those mentioned earlier. Previous studies have shown inconsistent results amplifying DNA from bones macerated in hot-water solution, although duration seems to be a factor (1,3). Overall, these results suggest that the duration of treatment appears to be a strong factor in predicting bone quality as well as the ability to amplify DNA.

The extra step of degreasing samples macerated by Biz/Na_2CO_3 greatly prolonged the procedure and produced mixed results. The amplification success rate for the fibula decreased relative to the nondegreased samples, yet all markers were amplified for the tibia sample with the same treatment. The previous study using pig ribs showed negative results for all three nuclear loci using the same treatment (1). The discrepancies between the fibula and tibia sample from this study may be because of initial differences in the amount of endogenous DNA. In general, the addition of a degreaser to Biz/Na_2CO_3 is not recommended as it not only produces bone of lower quality and is time-consuming, but it also may have a negative effect on nuclear DNA amplification.

Prolonged treatment may also explain the poor performance of bones treated with hydrogen peroxide. This method did not show any successful amplification in our study but weak amplification of mtDNA and nuclear DNA was observed for pig ribs (1). This method of maceration required the longest time in our study (112–121 days), much longer than the previous study (23.5–34 days). Thus, while the chemical itself may not be detrimental to DNA amplification, the prolonged duration seemed to have made amplification more difficult.

The relatively good amplification results for bones treated with EDTA/papain and bleach, which occurred at low temperatures (45 and 22°C) for at least 4 days, demonstrate that factors other than just treatment duration also impact nuclear DNA amplification, namely the nature of the chemicals themselves. Macerating bones in EDTA/papain required 4–5 days (because of the required lower heating temperature at 45°C to avoid denaturation of the enzymes) and produced similar or more successful amplifications than samples heated between 14 and 42 h (meat tenderizer/Palmolive®, boiling and hot water). Also, while the samples from bones treated with EDTA/papain were amplified with relative success at loci as large as ~300 bp in this study, markers used to amplify larger nuclear DNA loci (268–503 bp) showed negative results for the same treatment in the previous study using pig ribs (1). It is notable that the maceration duration for EDTA/papain from the previous study required as much as 10 days. This suggests that short duration of bone maceration with EDTA/papain is relatively effective in obtaining amplifiable DNA as well as maintaining good bone quality. Furthermore, while treatment duration is not the sole factor in determining successful nuclear DNA amplification, prolonged treatment of bone with EDTA/papain for more than 5 days may result in lower quality DNA.

There seem to be varying opinions concerning the effects of bleach on cortical bone quality and endogenous DNA availability (1,2,27). Our results showed that samples treated with bleach amplified successfully at all but four loci (D7S820, CSF1PO, FGA, and D18S51), two of which are the largest markers from the COfiler® and Profiler Plus® ID kits (>230 bp). Steadman et al. (1) showed successful amplification at three mtDNA loci but not for nuclear loci larger than 268 bp. This suggests that bleach treatment of bones tends to lower the quality of DNA, resulting in shorter fragments. This would explain the negative results in larger-sized markers of nuclear DNA. The duration for the bleach treatment was also longer in the previous study, ranging from 12 to 70 days compared to 5–8 days in this study. Therefore, longer treatment of bone with bleach is not recommended, but short washes advocated in one study (17) for purposes of decontamination may be appropriate.

Our study shows that temperature does not have a significant effect on nuclear DNA amplification. The highest temperature used for water maceration was 100°C, and the results showed relatively good amplification. Heat treatments that required a high temperature (90°C) (except hot water alone) produced relatively successful DNA amplification, around 50% or higher. This is comparable to the amplification success rate of samples macerated with EDTA/papain, which required a temperature of 45°C. Thus, the type of chemicals used in maceration and treatment duration seem to be important factors in producing amplifiable DNA.

Researchers cannot always employ the appropriate bone maceration technique that optimizes the chance for DNA genotyping due to a number of reasons. Other applications that focus on amplifying highly fragmented DNA, such as miniplex primer sets or the development of primers to type shorter-length STR loci, may prove useful for samples that did not fare well with some of the maceration methods used in this study (16–18,21,28). One study showed the utility of miniplex primer sets to analyze naturally degraded DNA from human remains that have been exposed to a variety of environmental conditions because of differing storage conditions (16). That study showed that samples kept in cold storage for up to 14 years were less prone to degradation. While these provide helpful guidelines for researchers, we must also acknowledge the difficulty in accounting for treatment of remains prior to storage. In forensic and anthropological studies, the treatment and storage of human remains differ at each institution, and these conditions may also need to be taken into consideration when determining the likelihood of DNA recovery and successful genotyping.

Conclusions

This study substantiates previous findings that the treatment of skeletal remains prior to molecular analysis can have a direct impact on successful extraction and analysis of DNA. We have focused on the practical aspect of nuclear DNA amplification via STR analysis utilizing the CODIS markers commonly used in forensic studies. Our study suggests that heat treatment via microwave or Biz/Na_2CO_3 in sub-boiling water shows the highest efficacy in terms of maximizing nuclear DNA amplification as well as sufficiently macerating bones for the purpose of osteological analysis. Treatments that require longer durations are discouraged as they seem to accelerate DNA degradation, resulting in a low success rate of DNA amplification, and also poorer bone quality. Maceration at temperatures over 90°C may produce good bone quality, but results in less-amplifiable DNA when treated for more than 10 h. Methods using meat tenderizer/Palmolive® and boiling water show good bone quality and relative success of DNA amplification, but using hydrogen peroxide or a hot-water bath has the most deleterious effect on nuclear DNA amplification despite showing decent bone quality.

References

1. Steadman DLW, DiAntonio L, Wilson J, Sheridan KE, Tammariello S. The effects of chemical and heat and maceration on the recovery of nuclear and mitochondrial DNA from bone. J Forensic Sci 2006;51: 11–7.
2. Fenton TW, Birkby WH, Cornelison J. A fast and safe non-bleaching method for forensic skeletal preparation. J Forensic Sci 2003;48:274–6.
3. Arismendi JL, Baker LE, Matteson KJ. Effects of processing techniques on the forensic DNA analysis of human skeletal remains. J Forensic Sci 2004;49:1–5.
4. Rennick SL, Fenton TW, Foran DR. The effects of skeletal preparation techniques on DNA from human and nonhuman bone. J Forensic Sci 2005;50:1016–9.
5. Nicklas JA, Buel E. Quantification of DNA in forensic samples. Anal Bioanal Chem 2003;376:1160–7.
6. Niederstätter H, Köchl S, Grubwieser P, Pavlic M, Steinlechner M, Parson W. A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA. Forensic Sci Int Genet 2007;1:29–34.
7. Foran DR. Relative degradation of nuclear and mitochondrial DNA: an experimental approach. J Forensic Sci 2006;51:766–70.
8. Lertrit P, Poolsuwan S, Thosarat R, Sanpachudayan T, Boonyarit H, Chinpaisal C, et al. Genetic history of Southeast Asian populations as revealed by ancient and modern human mitochondrial DNA analysis. Am J Phys Anthropol 2008;137:425–40.
9. Shook BAS, Smith DG. Using ancient mtDNA to reconstruct the population history of Northeastern North America. Am J Phys Anthropol 2008;137:14–29.
10. Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R, Heyd V, et al. Ancient DNA, strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. Proc Natl Acad Sci USA 2008;105:18226–31.
11. Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, et al. Sequencing and analysis of Neanderthal genomic DNA. Science 2006;314:1113–8.
12. Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, et al. Analysis of one million base pairs of Neanderthal DNA. Nature 2006;444:330–6.
13. Gilbert MTP, Kivisild T, Grønnow B, Andersen PK, Metspalu E, Reidla M, et al. Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. Science 2008;320:1787–9.
14. Gojanović MD, Sutlović D. Skeletal remains from World War II mass grave: from discovery to identification. Croat Med J 2007;48:520–7.
15. Nelson K, Melton T. Forensic mitochondrial DNA analysis of 116 casework skeletal samples. J Forensic Sci 2007;52:557–61.
16. Opel KL, Chung DT, Drábek J, Tatarek NE, Jantz LM, McCord BR. The application of miniplex primer sets in the analysis of degraded DNA from human skeletal remains. J Forensic Sci 2006;51:351–6.
17. Parsons TJ, Huel R, Davoren J, Katzmarzyk C, Miloš A, Selmanović A, et al. Application of novel "mini-amplicon" STR multiplexes to high volume casework on degraded skeletal remains. Forensic Sci Int Genet 2007;1:175–9.
18. Grubwieser P, Mühlmann R, Berger B, Niederstätter H, Pavlic M, Parson W. A new "miniSTR-multiplex" displaying reduced amplicon lengths for the analysis of degraded DNA. Int J Legal Med 2006;120:115–20.
19. Kemp BM, Smith DG. Use of bleach to eliminate contaminating DNA form the surface of bones and teeth. Forensic Sci Int 2005;154:53–61.
20. Crainic K, Paraire F, Leterreux M, Durigon M, de Mazancourt P. Skeletal remains presumed submerged in water for three years identified using PCR-STR analysis. J Forensic Sci 2002;47:1025–7.
21. Hill CR, Kline MC, Coble MD, Butler JM. Characterization of 26 MiniSTR loci for improved analysis of degraded DNA samples. J Forensic Sci 2008;53:73–80.
22. Butler JM. Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci 2006;51:253–65.
23. Isenberg AR. Forensic mitochondrial DNA analysis. In: Saferstein R, editor. Forensic science handbook, Vol. II, 2nd edn. Upper Saddle River, NJ: Pearson, 2005;297–327.
24. Applied Biosystems. AmpFℓSTR® Profiler Plus® PCR Amplification kit user's manual. Foster City, CA: Applied Biosystems, 2006.
25. Applied Biosystems. AmpFℓSTR® COfiler® PCR Amplification kit user's manual. Foster City, CA: Applied Biosystems, 2006.
26. Nielsen K, Mogensen HS, Hedman J, Niederstätter H, Parson W, Morling N. Comparison of five DNA quantification methods. Forensic Sci Int Genet 2008;2:226–30.
27. Hangay G, Dingley M. Biological museum methods, Vol. 1. Orlando, FL: Academic Press, 1985.
28. Hill CR, Kline MC, Mulero JJ, Lagacé RE, Chang C-W, Hennessy LK, et al. Concordance study between the AmpFℓSTR MiniFiler™ PCR amplification kit and conventional STR typing kits. J Forensic Sci 2007;52:870–3.

Additional information and reprint requests:
Dawnie Wolfe Steadman, Ph.D.
Department of Anthropology
Binghamton University
PO Box 6000
Binghamton, NY 13902-6000
E-mail: osteo@binghamton.edu

TECHNICAL NOTE
PHYSICAL ANTHROPOLOGY

Chang Liu,[1] *B.Sc.; Hyun M. Park,*[2] *Ph.D.; Maria V. Monsalve,*[3] *Ph.D.; and David D. Y. Chen,*[1] *Ph.D.*

Free Fatty Acids Composition in Adipocere of the Kwäday Dän Ts'ìnchí Ancient Remains Found in a Glacier*

ABSTRACT: Adipocere is a postmortem decomposition product consisting of mostly a mixture of free fatty acids (FFAs) that are formed because of the hydrolysis of triglycerides in adipose tissues. This article describes a simple and robust method for the extraction, identification, and quantification of FFA commonly found in adipocere using gas chromatography–mass spectrometry (GC/MS). This method was applied to analyze tissues from Kwäday Dän Ts'ìnchí, ancient remains discovered in a retreating glacier in the Tatshenshini-Alsek Park, British Columbia, Canada in August 1999. The lyophilized tissues were grinded and extracted with hexane. The trimethylsilyl fatty acid derivatives were analyzed by GC/MS, and the relative abundances of myristic acid, palmitic acid, oleic acid, and stearic acid were determined. Milligram per gram levels of saturated fatty acids were found in the tissues of the ancient remains, while the levels of unsaturated fatty acids, such as palmitoleic acid, were found to be negligible. The results provided further evidence of the existence of adipocere found during forensic examination of the Kwäday Dän Ts'ìnchí ancient remains.

KEYWORDS: forensic science, free fatty acids, adipocere, extraction method, gas chromatography–mass spectrometry, ancient human remains

On August 14, 1999, the frozen remains of a young man were discovered by hunters in a glacier in the Tatshenshini-Alsek Park of northwestern British Columbia (BC), Canada. The preserved body was the first ancient corpse discovered in a melting glacier in North America. The local Champagne and Aishihik First Nations people named the remains "Kwäday Dän Ts'ìnchí", which means Long-Ago Person Found. Several items were discovered with the ancient body, including a robe-style fur garment, a plant fiber hat, and some wooden artifacts. Initial radiocarbon dating of the plant fibers on his hat and the animal skin clothing indicated an age of about 500 years (1). However, more recent radiocarbon data obtained directly from the analysis of his body suggested that Kwäday Dän Ts'ìnchí more likely died between 1670 and 1850 AD (2), still predating Europeans entering this region (circa 1890). The local people allowed the collection of small amounts of samples enabling limited scientific research to continue after the cremation of the Kwäday Dän Ts'ìnchí, and these collected samples provided a unique opportunity to assess the tissues in frozen remains of this age for study (3–6).

The name adipocere is derived from Latin, meaning fatty wax, which is also known as grave wax or mortuary wax. This grayish-white water-insoluble material is the late-stage postmortem hydrolysis product of triglycerides in adipose tissues. A number of studies on the characterization of the chemical composition and properties of adipocere have been reported (7–13). The characterization of the preservation process provides important information about the surrounding environment in a forensic context. The conditions under which this process can occur vary, but humid and anaerobic conditions must be present. Corpses released by glaciers and bodies found in aquatic environments often have adipocere present. However, the formation of this substance is not well understood because of insufficient knowledge about the specific factors leading to adipocere. As such, it is of considerable interest to have a reliable analytical method that enables the identification and quantification of adipocere to evaluate the state of decomposition of a body and to assess the influence from the surrounding environment.

Beattie et al. (1) reported adipocere formation on the back, the vertebral spine, and the intrinsic muscles of Kwäday Dän Ts'ìnchí. The formation of adipocere usually takes about 3–6 months, but a complete transformation of the whole body could take many years (7). The formation of adipocere is initiated by intrinsic lipases, which convert the triacylglycerides (TAG) into their corresponding saturated and unsaturated fatty acids. The unsaturated fatty acids may then be hydrogenated, facilitated by degradative anaerobic bacteria, to form their corresponding saturated fatty acids in humid and micro-aerobic conditions (14).

The major constituents of adipocere have been identified as saturated free fatty acids (FFAs) with even numbers of carbon atoms, such as myristic acid, palmitic acid, and stearic acid. Among these, palmitic acid is usually the most abundant acid, followed by stearic acid then myristic acid (11,12). Unsaturated fatty acids, including palmitoleic, oleic, and linoleic acid, may be present in smaller amounts (15,16). In some cases, minor components such as TAG, 10-hydroxystearic acid, 12-hydroxystearic acid, phytanoic acid, and the calcium and magnesium salts of fatty acids are also found (8,15–21).

[1]Department of Chemistry, University of British Columbia, Vancouver, BC V6T 1Z1, Canada.

[2]Advanced Analysis Center, Korea Institute of Science and Technology, Seoul 130-650 Korea.

[3]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC V6T 1Z1, Canada.

TABLE 1—*Optimum detection conditions for GC/MS analysis of FFA.*

Parameter	Condition
GC parameters	
Column type	DB-5MS (J&W Scientific, Santa Clara, CA) Fused-silica capillary column (30 m × 0.25 mm × 0.25 μm 5% phenyl, 95% dimethylpolysiloxane)
Inlet pressure	100 kPa
Carrier gas	Helium
Injection volume	1 μL
Injection type	Pulsed splitless mode
Temperature	250°C
Initial oven temperature	100°C
Rate of change	15°C/min to 210°C and 5°C/min to 290°C held 3 min
MS parameters	
Acquisition mode	Scan
Scan parameters	50–650 *m/z*
Solvent delay	3 min
Quadrupole temperature	150°C
Source temperature	230°C

GC, gas chromatography; MS, mass spectrometry; FFA, free fatty acids.

There have been some methods reported that analyze the FFA components of adipocere. Stuart et al. and Bereuter et al. (15,17) have reported a rapid screening method to identify adipocere formation using infrared spectroscopy. A study of adipocere using high-performance liquid chromatography (HPLC) was demonstrated by Yan et al. (18). Gas chromatography (GC) and GC combined with mass spectrometry (GC/MS) were also used to the chemical characterization of adipocere (10,12). This work reports a rapid and simple method of extraction in combination with GC/MS for quantification of FFAs from adipocere tissues.

Materials and Methods

Chemicals and Materials

The protocol and consent form for the study of Kwäday Dän Ts'ìnchí was approved by the Clinical Research Board Ethics Committee at the University of British Columbia. Trapezius (0.5 g) and bicep (0.5 g) muscles from the Kwäday Dän Ts'ìnchí (kept at −80°C) and from an unembalmed male cadaver donated for education and research to the Faculty of Medicine at the University of British Columbia were excised with a sterile scalpel. The cadaver tissues, which also were stored at −80°C, were used as controls.

The lipid standards (GC grade), including triolein, myristic acid, palmitic acid, palmitoleic acid, heptadecanoic acid, linoleic acid, oleic acid, and stearic acid, were purchased from Sigma–Aldrich (Oakville, ON, Canada), as was the silylating agent, *N,O*-bis(trimethylsilyl) trifluoroacetamide containing 1% v/vtrimethylchlorosilane (BSTFA + 1% TMCS). Individual stock solutions (10 mg/mL) of lipid standards were prepared in hexane (HPLC Grade; Fisher Scientific, Nepean, ON, Canada).

Extraction of Lipids

Prior to extraction, moisture in the tissues was removed. For adipocere samples, 20 mg of tissue was weighed out and lyophilized (Flexi-Dry™ MP Freezing Dryer, SP Scientific, Stone Ridge, NY) overnight. The dried tissue was then ground into a fine powder. The resulting powder was transferred into a glass vial with a polytetrafluoroethylene (PTFE) cap. Three milliliters hexane was added and mixed with the powder. At this point, heptadecanoic acid (2 mg/mL, 200 μL) was added to the mixture as an internal standard.

The sample was then sonicated for three 30-min sessions and centrifuged (Sorvall GLC-1 General Laboratory Centrifuge, Thermo Scientific, Waltham, MA) for 45 min at 1000×g. This procedure was then repeated two more times, resulting in a total volume of approximately 9 mL of extractant. The supernatant was concentrated by rotary evaporator (Heating Bath B490, Buchi Laboratory Equipment, Flawil, Switzerland) until about 1 mL of solvent was left. The remaining solvent was aspirated to dryness using a nitrogen stream. The dried sample was then redissolved in 1-mL hexane.

Derivatization of Individual Fatty Acids

FFAs contained in the 1-mL hexane were converted to their trimethylsilyl (TMS) derivatives by BSTFA with 1% TMCS.
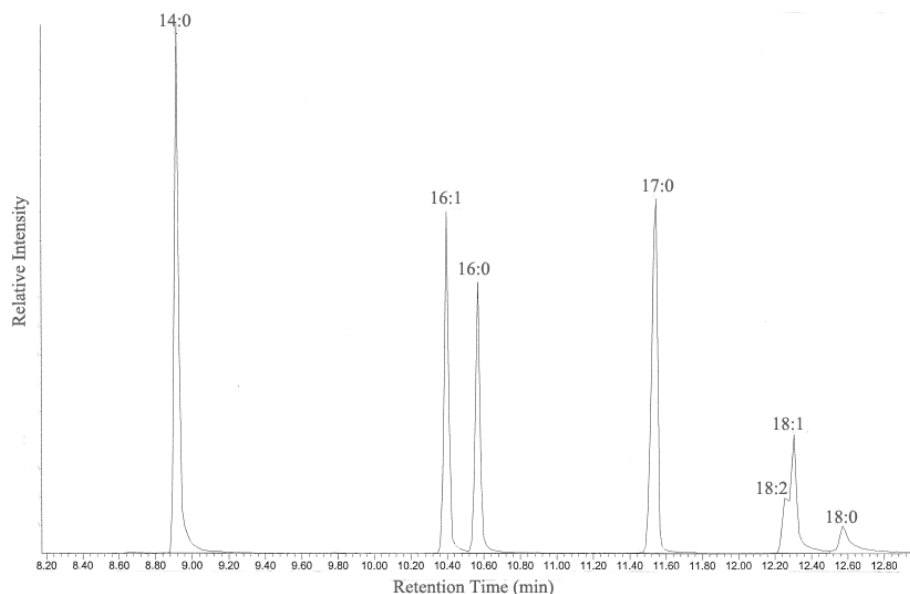

FIG. 1—*Partial gas chromatograms of trimethylsilyl derivatives of the standard mixture of lipids (500 μg/mL). The numerical notation associated to each peak shows the numbers of carbons and the number of double bonds, respectively, and the identity of the compounds are: 14:0-myristic acid, 16:1-palmitoleic acid, 16:0-palmitic acid, 17:0-heptadecanoic acid (internal standard), 18:2-linoleic acid, 18:1-oleic acid, 18:0-stearic acid.*
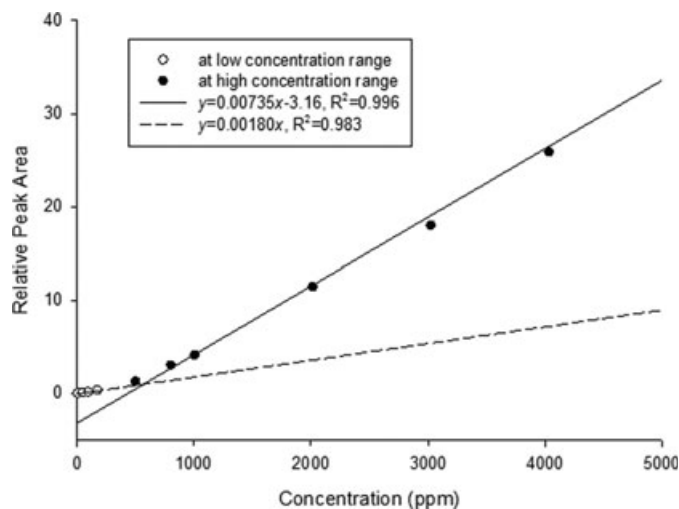
FIG. 2—*Relationship between relative peak area ($A_{Myristic}/A_{IS}$) and the concentration of myristic acid: (a) concentration from 500 to 5000 ppm, (b) concentration from 10 to 200 ppm.*

Excess silylation reagent was added (200 μL), and the vials were tightly capped with PTFE-sealed screw cap lids. The vials were then placed in a water bath at 60°C for 30 min. The TMS esters formed during this procedure were injected into the GC/MS for analysis.

Fatty Acid Analysis

The TMS fatty acid derivatives were analyzed by an Agilent 6890 Series GC coupled to an Agilent 5973 Network mass spectrometer. The analysis was monitored in total ion scan mode, and the fatty acids known to be present in adipocere were identified. The saturated fatty acids considered were myristic acid, palmitic acid, and stearic acid. The unsaturated fatty acids, palmitoleic acid, oleic acid, and linoleic acid, were also considered because of their possible presence in low concentrations. Peaks relating to the TMS

esters of fatty acids were identified by comparison of their retention times and mass spectra with standards. The optimized conditions for the operation of the GC/MS system are listed in Table 1.

Results and Discussion

Calibration

A standard mixture of lipids (oleic acid, linoleic acid, palmitoleic acid, stearic acid, palmitic acid, and myristic acid) at a number of concentrations (10, 20, 50, 100, 200, 500, 800, 1000, 2000, 3000, 4000, and 5000 μg/mL) were dissolved in 1-mL hexane, and 200-μL heptadecanoic acid (2 mg/mL) was added as the internal standard. The TMS fatty acid derivatives in the mixture were efficiently separated under the optimized GC/MS conditions, as shown in Fig. 1.

The calibration curves are not linear throughout the large concentration range tested, and the concentrations of individual FFA in adipocere were significantly different. Therefore, the calibration curves to be used were constructed with a rather narrow range in the vicinity of the concentration of the fatty acids present in the sample. As shown in Fig. 2, while the slope of the calibration curve for larger concentrations of myristic acid is 0.0018, the slope of the calibration curve for smaller concentrations is 0.0073. As a result, it was necessary to create calibration curves for each of the individual fatty acids studied with concentrations relevant to the peak areas obtained from the tissue analysis.

Lyophilization

To increase the recovery of the FFAs in the tissues, it was necessary to grind the tissues into fine powders prior to the extraction of FFA with hexane. Because both the fresh muscle tissue and adipocere tissue contain moisture, the tissues were lyophilized first to dryness before grinding. Lyophilization also causes less damage to the material than other dehydration methods using higher temperatures (22). Additionally, the boiling points of fatty acids are relatively low, and lyophilization could reduce the loss of the fatty acids in the tissues.
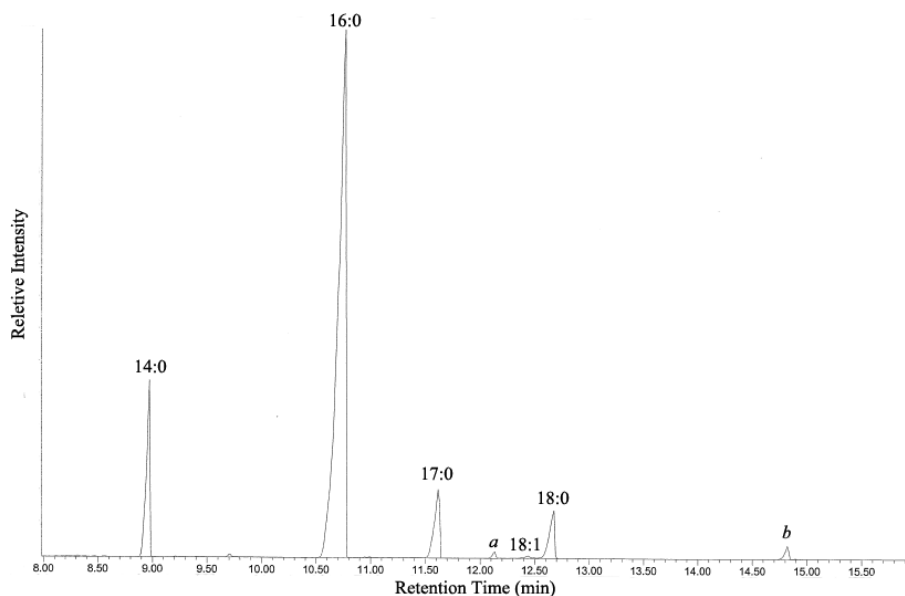

FIG. 3—*Gas chromatography/mass spectrometry chromatogram of fatty acid in Kwäday Dän Ts'ìnchí adipocere: 14:0-myristic acid, 16:0-palmitic acid, 17:0-heptadecanoic acid (internal standard), 18:1-oleic acid, 18:0-stearic acid, a-phytanoic acid, b-12-hydroxystearic acid.*

TABLE 2—*FFA concentration by weight (WT [μg/g]) and FFA relative composition (RC [%]) of fresh muscle tissue and adipocere of Kwäday Dän Ts'ìnchí (KDT).*

FFA		Fresh Muscle Tissue (Control)				KDT Sample			
		Biceps		Trapezius		Biceps		Trapezius	
FFA Name	Notation	WT	RC	WT	RC	WT	RC	WT	RC
Myristic acid	14:0	0	0.0	8	3.3	9405	9.0	9906	10.9
Palmitoleic acid	16:1	0	0.0	7	3.0	0	0.0	0	0.0
Palmitic acid	16:0	301	100.0	215	93.8	64434	61.4	54891	60.4
Oleic acid	18:1	0	0.0	0	0.0	8338	7.9	4433	4.9
Stearic acid	18:0	0	0.0	0	0.0	22778	21.7	21690	23.8

FFA, free fatty acids.

Application of the Method to the Samples

Figure 3 shows the GC/MS chromatogram of the total lipid extract (following trimethylsilylation) isolated from adipocere of Kwäday Dän Ts'ìnchí. The relative compositions of fatty acids in the tissues of Kwäday Dän Ts'ìnchí and the control subject are listed in Table 2. The concentrations of FFA in the control tissues, obtained from a recently deceased cadaver, are extremely low, and only palmitic acid could be detected with certainty. For the adipocere tissue of Kwäday Dän Ts'ìnchí, even-numbered, straight-chain saturated fatty acids were detected in abundance. Palmitic acid was the most abundant, with a relative composition of about 60% of the total FFA. This is because when the lipids were hydrolyzed by degradative anaerobes, they are converted from the neutral fat into FFA. This result is similar to those obtained from previous studies of adipocere (6,10,23). Some of the FFAs observed in this work have also been found in the bone and skin samples of Kwäday Dän Ts'ìnchí. Although the relative abundance varies somewhat depending on the types of tissue, the similarity between these results should be enough to suggest the presence of adipocere in the preservation processes (5).

The unsaturated fatty acids could be hydrogenated by bacterial enzymes to corresponding saturated fatty acids in humid and micro-aerobic conditions (14). As a result, the concentration of monounsaturated acids was low compared to the saturated fatty acids. Although oleic acid was observed, the peaks for palmitoleic acid and linoleic acid could not be seen in significant abundance.

Other Fatty Acid Components

As shown in Fig. 3, a few other peaks were seen in the chromatograms. The structures of these components were identified by their mass spectra to be 3,7,11,15-tetramethylhexadecanoic acid (phytanic acid) and 12-hydroxystearic acid. These fatty acids are also common constituents of adipocere (20). Because of the lack of appropriate standards, these components were not quantified in this study.

Conclusions

A simple and robust analytical method for the extraction, separation, and detection of FFAs from adipocere was developed and applied to the frozen human remains of Kwäday Dän Ts'ìnchí. Lyophilization of the tissues before extraction facilitated grinding of the tissue into fine particles to release the fatty acid components during the hexane extraction process. Using GC/MS, the relative fatty acid composition of adipocere in the Kwäday Dän Ts'ìnchí samples was determined. The results provided further evidence of adipocere in ancient tissues preserved in the glacier.

References

1. Beattie O, Apland B, Blake EW, Cosgrove JA, Gaunt S, Greer S, et al. The Kwäday Dän Ts'ìnchi discovery from a glacier in British Columbia. Can J Archaeol 2000;24:129–47.
2. Richards MP, Greer S, Corr LT, Beattie O, Mackie A, Evershed RP, et al. Radiocarbon dating and dietary stable isotope analysis of Kwäday Dän Ts'ìnchi. Am Antiquity 2007;72:719–33.
3. Monsalve MV, Humphrey E, Walker DC, Cheng C, Vogl W, Nimmo M. Brief communication: state of preservation of tissues from ancient human remains found in a glacier in Canada. Am J Phys Anthropol 2008;137:348–55.
4. Monsalve MV, Stone AC, Lewis CM, Rempel A, Richards M, Straathof D, et al. Brief communication: molecular analysis of the Kwäday Dän Ts'ìnchi ancient remains found in a glacier in Canada. Am J Phys Anthropol 2002;119:288–91.
5. Corr LT, Richards MP, Jim S, Ambrose SH, Mackie A, Beattie O, et al. Probing dietary change of the Kwäday Dän Ts'ìnchi individual, an ancient glacier body from British Columbia: I. Complementary use of marine lipid biomarker and carbon isotope signatures as novel indicators of a marine diet. J Archaeol Sci 2008;35:2102–10.
6. Dickson JH, Richards MP, Hebda RJ, Mudie PJ, Beattie O, Ramsay S, et al. Kwäday Dän Ts'ìnchí, the first ancient body of a man from a North American glacier: reconstructing his last days by intestinal and biomolecular analysis. Holocene 2004;14:481–6.
7. Bereuter TL, Lorbeer E, Reiter C, Seidler H, Unterdorfer H. Post mortem alterations of human lipids—Part I: evaluation of adipocere formation and mummification by desiccation. In: Spindler K, Wilfing H, Rastbichler-Zissernig E, zur Nedden D, Nothdufter H, editors. Human mummies. New York and Vienna: Springer-Verlag, 1996;265–73.
8. Evershed RP. Chemical composition of a bog body adipocere. Archaeometry 1992;34:253–65.
9. Fiedler S, Graw M. Decomposition of buried corpses, with special reference to the formation of adipocere. Naturwissenschaften 2003;90:291–300.
10. Forbes SL, Stuart BH, Dent BB. The identification of adipocere in grave soils. Forensic Sci Int 2002;127:225–30.
11. Kahana T, Almog J, Levy J, Shmeltzer E, Spier Y, Hiss J. Marine tophonomy—adipocere formation in a series of bodies recovered from a single shipwreck. J Forensic Sci 1999;44:897–901.
12. Takatori T. The mechanism of human adipocere formation. Leg Med 2001;3:193–204.
13. Makristathis A, Schwarzmeier J, Mader RM, Varmuza K, Simonitsch I, Chaves JC, et al. Fatty acid composition and preservation of the Tyrolean iceman and other mummies. J Lipid Res 2002;43:2056–61.

14. Polson CJ. The essentials of forensic medicine, 2nd edn. New York, NY: Pergamon Press, 1965.

15. Stuart BH, Forbes SL, Dent BB, Hodgson G. Studies of adipocere using diffuse reflectance infrared spectroscopy. Vib Sepctrosc 2000;24:233–42.

16. Adachi J, Ueno Y, Miwa A, Asano M, Nishimura A, Tatsuno Y. Epico-prostanol found in adipocere from five human autopsies. Lipids 1997;32:1155–60.

17. Bereuter TL, Mikenda W, Reiter C. Iceman's mummification—implica-toins from infrared spectroscopical and histological studies. Chem Eur J 1997;3:1032–8.

18. Yan F, Mcnally R, Kontanis EJ, Sadik OA. Preliminary quantitative investigation of postmortem adipocere formation. J Forensic Sci 2001;46:609–14.

19. Takatori T. Investigations on the mechanism of adipocere formation and its relation to other biochemical reactions. Forensic Sci Int 1996;80:49–61.

20. Notter SJ, Stuart BH, Dent BB, Keegan J. Solid-phase extraction in combination with GC/MS for the quantification of free fatty acids in adipocere. Eur J Lipid Sci Technol 2008;110:73–80.

21. Gill-King H. Chemical and ultrastructural aspects of decomposition. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press, 1997; 93–108.

22. Kennedy JF, Joaquim MSC. Recovery processes for biological materials. New York, NY: John Wiley & Sons Ltd, 1993.

23. Ruttan RF, Marshall MJ. The composition of adipocere. J Biol Chem 1917;29:319–27.

Additional information and reprint requests:
David D. Y. Chen, Ph.D.
Department of Chemistry
University of British Columbia
Vancouver
BC V6T 1Z3
Canada
E-mail: chen@chem.ubc.ca

TECHNICAL NOTE
CRIMINALISTICS

Nicole H. Hoffman,[1,2] *M.S. and Terry Fenger,*[1] *Ph.D.*

Validation of Half-Reaction Amplification Using Promega PowerPlex® 16*

ABSTRACT: DNA amplification is a fundamental yet costly process used in DNA analysis. This study evaluated half-reaction amplification (12.5, 12, and 13 uL) using the Promega Powerplex® 16 Kit with the hope of reducing sample analysis costs by half. A sensitivity study was completed, along with the testing of various blood stain samples including those with low (<0.40 ng) and high DNA concentrations (>3.0 ng), peak height imbalances, and allelic drop-out. Also, 467 samples submitted to the MUFSC laboratory for testing were analyzed. Results indicate that half-reaction amplification produced higher quality profiles than full-reactions. Average peak heights increased by 85%, peak height imbalances improved, and drop-out was eliminated in 75.8% of samples. Only eight of 467 case samples required re-amplification, a success rate of 94% was observed, and the repeat rate decreased significantly. Finally, a DNA input of 0.25–1.0 ng is ideal for half-reaction amplification.

KEYWORDS: forensic science, forensic DNA, validation, DNA amplification, PowerPlex® 16, CSF, D13S317, D16S539, D18S51, D21S11, D3S51358, D5S818, D7S820, D8S1179, FGA, Penta D, Penta E, THO1, TPOX, vWA

Over the last decade, laboratories have seen an increase in backlogged cases and an increase in the number and types of samples being submitted for analysis (1). To support this increased demand, laboratories can expect an increase in operational costs stemming from the purchase of more products for DNA testing, and the hiring of more analysts to handle extra caseload volumes. While the forensic community has seen an increasing success rate in sample analysis, laboratories continue to search for faster, better, and less-expensive ways to analyze samples without compromising their quality of work (2). The DNA amplification process is a possible area for such cost reduction. Specifically, the Promega (Madison, WI) PowerPlex® 16 Amplification Kit currently costs $5,995 and supports the amplification of 400 DNA samples (approximately $15 per sample). This study investigated the possibility of reducing the amount of amplification reagents used to increase the total reactions of each kit to 800 samples. This would reduce the cost per sample by half. The overall objective was to optimize and evaluate the performance of a half-reaction amplification protocol using the Promega PowerPlex® 16 Amplification Kit.

The DNA amplification step was chosen as an option for cost reduction because it has proven to be the most costly step throughout the DNA analysis process. A recent in-house cost analysis study conducted at the Marshall University Forensic Science Center (MUFSC) calculated the total cost of analyzing a single sample to be $24.30. Of that total, $15.92 is spent on PowerPlex® 16

amplification alone, accounting for nearly 66% of the total costs. Not only is DNA amplification one of the most expensive steps but it is a highly essential one as well. Many database and casework samples with limited DNA and/or of poor quality could not be analyzed without proper PCR amplification (3). Two factors were considered for this validation; a total reaction volume that is suitable for amplification and the volume of each reaction component that will support DNA amplification. First, Promega recommends the amplification of a DNA template in a total volume of 25 μL (4). However, amplification can be performed using a total volume in the range of 5–100 μL (3). A reduced total amplification volume of 12.5 μL falls within this range and should be suitable for successful amplification. Second, manufacturers' recommended protocols suggest reaction volumes of each component based on optimization studies (5). These studies have demonstrated the optimal ratio of reagents and the amplification parameters to use for the production of quality DNA profiles. Therefore, the overall goal was to keep all variables constant except the total reaction volume to collect data that was comparable to original full-reaction data.

The effects of half-reaction PCR were assessed through the analysis of five different studies using samples with various issues seen at the MUFSC laboratory. First, samples previously producing full profiles and requiring no repeat analysis when amplified using the full-reaction protocol were evaluated to determine whether amplification in 12.5 uL would be successful on quality samples. Second, a variety of samples commonly seen at the MUFSC laboratory and identified as "problem samples" were evaluated. These included samples with high and low DNA concentrations (>3.0 and <0.4 ng, respectively), allelic drop-out, and peak height imbalances at the D5S818 locus. Total reaction volumes of 12 and 13 uL were also assessed to allow for an easier amplification set-up. A sensitivity study was completed to determine the range of DNA template most suitable for half-reaction amplification. Finally, the half-reaction amplification procedure was used to test 456 buccal swab samples outsourced to the MUFSC laboratory for analysis.

[1]Marshall University Forensic Science Center, DNA Laboratory, 1401 Forensic Science Drive, Huntington, WV.

[2]Indiana State Police, Evansville Regional Laboratory, 19411 Hwy 41 North, Evansville, IN.

Materials and Methods

DNA Samples and Preparation

This study utilized single-source blood stain samples previously analyzed in the MUFSC laboratory. Original samples were spotted on FTA® cards (Fitzco Inc., Spring Park, MN), extracted on a Beckman Coulter Biomek® 2000 using Promega's DNA IQ™ Extraction Kit, amplified with the Promega Powerplex® 16 Amplification Kit (full-reaction protocol) on a GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA), run on an ABI PRISM® 3100 Genetic Analyzer and analyzed using ABI PRISM® Data Collection Software v.1.1, Genescan® v.3.7.1, and Genotyper® v 3.7 software. All procedures were performed using manufacturer's recommended protocols. Samples used in this study were either the original extracts (not previously quantified) or the normalized samples that were previously quantified, according to manufacturer's protocol, using the Quantifiler® Human DNA Quantification Kit (Applied Biosystems). Normalization is the automated process of diluting extracted samples to the same concentration. In this study, 25 uL of extract product was added to a variable amount of water that would bring all samples to a concentration of 1 ng/uL. Original extracts were not quantified because MUFSC validation studies of the Beckman Coulter Biomek® 2000 indicated that consistent quantities of DNA, sufficient for amplification, were obtained from FTA® cards extracted with the Promega DNA IQ™ Extraction Kit. Only a handful of samples were originally quantified; therefore, the concentration of most extracts used in this study was unknown. Known in-house DNA samples with known concentrations were used for the sensitivity study.

PCR STR Amplification

Three PowerPlex® 16 "half-reaction" volumes were used to amplify samples, 12 μL, 12.5 μL, and 13 μL. All three reaction volumes consisted of the following master mix components per sample: 1.25 μL Gold ST*R 10× Buffer, 0.4 μL AmpliTaq Gold® DNA Polymerase, and 1.25 μL PowerPlex® 16 10× Primer Pair Mix. Table 1 summarizes the volumes of master mix, DNA template, and molecular biology grade (MBG) water (Thermo Fisher Scientific, Waltham, MA) used for the three total amplification volumes. The volume of DNA template amplified remained constant at 1 μL to reflect the treatment of original samples as well as assess the efficacy of half-reaction volumes on PCR and not the variation of

TABLE 1—*Amplification components.*

Component	Total Amplification Volumes		
	12 uL Amp	12.5 uL Amp	13 uL Amp
Master mix	3.0	2.9	3.0
DNA template	1.0	1.0	1.0
MBG water	8.0	8.6	9.0

DNA template. Samples were loaded into a MicroAmp™ Optical 96-well Reaction Plate (Applied Biosystems), sealed with MicroAmp™ 8-cap strips, and amplified on a GeneAmp® PCR System 9700 thermal cycler. All samples were amplified using thermal cycling parameters as follows: 95°C for 11 min, 96°C for 1 min; 10 cycles of 94°C 30 sec (ramp 100%), 60°C for 30 sec (ramp 29%), and 70°C for 45 sec (ramp 23%); 22 cycles of 90°C for 30 sec (ramp 100%), 60°C for 30 sec (ramp 29%), and 70°C for 45 sec (ramp 23%); and 60°C hold for 30 min.

Amplification of Various Sample Types

Five studies were conducted utilizing various samples including those of good quality, low DNA concentrations, high DNA concentrations, samples with peak height imbalances, and samples with allelic drop-out. Table 2 summarizes the samples used in each of the five studies.

Sensitivity

A sensitivity study was performed according to SWGDAM validation guidelines (6). Four in-house, known DNA samples previously extracted from buccal swabs on the Beckman Coulter Biomek® 2000 were serially diluted and quantified, according to manufacturer's protocol, using the Quantifiler® Human DNA Quantification Kit. Each dilution series, consisting of eight DNA concentrations ranging from 0.03125 to 5.0 ng, was amplified twice using the PowerPlex® 16 kit in a final volume of 12.5 uL. Data was analyzed at a threshold of 25 rfu (minimum), 50 rfu, and 100 rfu (maximum) to determine the optimal range of input DNA that would produce a complete profile. Results were compared to those originally obtained from the validation of PowerPlex® 16 using the full-reaction method.

Application to Case Samples

Finally, we evaluated the performance of the half-reaction amplification on 467 samples submitted to the MUFSC laboratory for analysis and research purposes. These samples were buccal swabs samples collected on cotton swabs. They were processed in the same manner as the blood stain samples used for this study.

CE and Data Analysis

Amplified products were prepared for capillary electrophoresis using 8.5 uL Hi-Di formamide, 0.5 uL Genescan 500 ROX internal size standard (Applied Biosystems), and 1 uL of the amplified product. Samples were injected on an ABI PRISM® 3130-*xl* Genetic Analyzer using a 3 kV, 5 sec injection with a run time of approximately 45 min. ABI PRISM® Data Collection Software v.1.1, Genescan® v.3.7.1, and Genotyper® v 3.7 software were used for data collection and analyzed at a maximum threshold of 100 rfu.

TABLE 2—*Summary of the five studies conducted.*

Study	Sample Types	Number of Samples	Amp Vol. (uL)	Number of Amps.
12.5 uL reaction volume	Good quality	12	12.5	1
Extreme DNA concentrations	<0.4 ng concentration	5	12.5	2
	>3.0 ng concentration	10	12.5	2
12 uL vs. 13 uL	Good quality	14	12, 13	1 each vol.
Peak height imbalances	D5S818 peak ratio <50%	9	12, 12.5, 13	1 each vol.
Allelic drop-out	Allelic drop-out	22	12, 13	1 (12 uL); 2 (13 uL)

Data and Statistics

The effectiveness of amplifying in a half-reaction volume was assessed primarily by evaluating peak height data. First, the mean peak height at each locus was calculated for all samples, as well as across all loci for each reaction volume. The quality of the DNA profile (drop-out, pull-up, stutter, off-scale data, etc.) was noted for samples containing high or low DNA concentrations. To assess samples with peak height imbalances, the peak height ratio at D5S818 for each sample and average ratios for each reaction volume were calculated and compared to full-reaction amplification data. The standard deviation (SD) for each reaction volume was calculated using the mean peak height ratio. Several samples in the "peak height imbalance study" were not available for testing at various times during this validation because they had yet to be submitted for analysis or were still being processed. Therefore, they are noted and were not included in calculations. To assess the success of half-reaction amplification on samples originally containing allelic drop-out, the percentage of full profiles obtained using the half-reaction method was calculated. This was carried out by dividing the number of full profiles obtained by the total number of samples tested. Data from the sensitivity study was used to determine the ideal concentration range that would yield quality profiles containing no pull-up, drop-out, off-scale peaks, or peak height imbalances of <50%.

Results and Discussion

Results indicate that half-reaction amplification (12.5 μL) using PowerPlex® 16 produced similar and often times improved profiles compared to full-reaction amplification (25 μL). Average peak heights increased across all but three of the 16 STR loci (D18S51, Penta E, and D8S1179), and by an average of 155.72 rfu (Fig. 1). Only one sample amplified with the half-reaction method produced drop-out that was not observed in the full-reaction amplification.

An increase in peak heights using half-reaction amplification may be useful when working with low level DNA samples. Samples amplified at full-reaction and exhibiting drop-out just below the threshold levels could be amplified with the half-reaction method and possibly result in higher peak heights. This may be possible because average peak heights tended to increase when amplifying in a lower volume. An increase in peak heights could be significant enough to raise the original drop-out allele above threshold and result in a full profile. If low level samples were initially amplified at half-reaction, the need for re-amplification because of drop-out could be eliminated. This, in turn, could allow for the use of less DNA, which is critical when little is available.
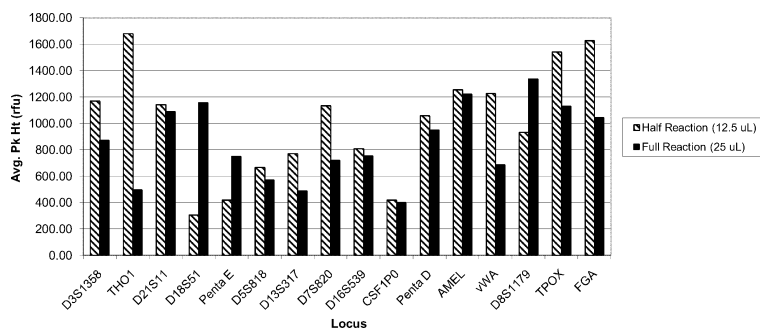
"Problem samples" containing low and high DNA concentrations were also analyzed using the half-reaction protocol. All samples with a low DNA concentration (<0.4 ng) produced full profiles when amplified in 12.5 uL. PowerPlex® 16 DNA target studies conducted by MUFSC have indicated that "a 1–4 ng DNA amplification range would be ideal when using the PowerPlex® 16 amplification system." Therefore, amplifying in half-reaction volumes may allow the system to work using lower concentrations of template DNA. This possibility was assessed during the sensitivity study (see further results). The ability to amplify less DNA could eliminate the need to concentrate low level samples, and in turn, the opportunity for contamination would be lessened because of reduced handling of samples.

Eight of the ten samples with a high DNA concentration produced full profiles in the first amplification while all ten had full profiles in the second amplification. The higher amount of input DNA did not result in pull-up, off-scale peaks, split peaks, enhanced stutter effects, or baseline noise often associated with the amplification of large concentrations of template DNA (7). The ability to amplify higher concentrations of DNA template would lessen the need for dilutions and, thus, the number of sample handling steps that could potentially increase the risk of contamination.

Varying Total Reaction Volumes

After initial findings demonstrated the success of half-reaction amplification in 12.5 uL, another study was performed to determine whether rounding the total reaction volume to 12 uL or 13 uL would be as effective. The half-reaction protocol using 12.5 uL requires the addition of 2.9 μL master mix to 8.6 μL DNA plus water. However, Rainin repeat pipettes (Rainin Instrumentation LLC, Oakland, CA) used at the MUFSC laboratory do not have a 2.9 uL setting, making manual amplification set-up more difficult. Therefore, this study investigated whether 3 uL of master mix could be used in a final volume of 12 uL or 13 uL. Overall, results indicate that amplification volumes of either 12 uL or 13 uL yielded similar and often times higher quality profiles than full-reaction amplifications. Average peak heights at each locus increased as the total reaction volume decreased. Average peak heights across all loci for 25, 13, and 12 uL reactions were 426.26, 1689.82, and 1925.50 rfu, respectively (Fig. 2). Because results for 12 and 13 μL reaction volumes were quite similar, it should be left to the discretion of each individual laboratory as to which volume is more suitable for their amplification needs.

D5S818 Peak Height Imbalances

A significant number of samples previously analyzed at the MUFSC laboratory have shown a peak height imbalance at the
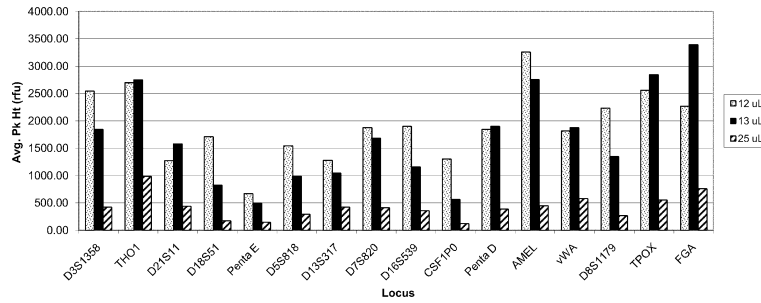

FIG. 1—*Average peak heights obtained when using half-reaction vs. full-reaction amplification.*

FIG. 2—*Relationship between varying total amplification volumes and average peak heights.*

D5S818 locus when a "10" allele was present within a heterozygote. In many cases, the "10" allele was not called in Genotyper® but was noted as a possible true allele because it was in the stutter position and above the 12.3% stutter threshold set forth by Promega (5). In other instances, the allele in question was noted because peak height ratios were very low or allelic drop-out was seen but a small peak was still visible. Re-amplification using PowerPlex® 16 produced similar, reproducible results. Amplification with the AmpFℓSTR Identifiler® Amplification Kit (Applied Biosystems) using the manufacturer's recommended protocol, produced two balanced peaks with proper allele calls. These observations may indicate a possible primer binding site mutation at D5S818 when amplifying with PowerPlex® 16.

Original extracts of nine samples were amplified using the three half-reaction volumes to determine whether reduced amplification volumes would have an effect on peak height imbalances at the D5S818 locus. Peak height ratios increased slightly as reaction volumes decreased (Fig. 3). The range of peak height ratios for full-reaction amplifications was 18.1–34.1%, while half-reaction amplifications ranged from 26.1% to 57%. Standard deviations decreased slightly as sample volumes increased but rose using a 25 uL volume (Fig. 3). Gaines et al. (8) conducted a similar study in which they investigated reduced reaction volumes using the AmpFℓSTR Profiler Plus Kit™. Their study showed that as reaction volumes decreased (with a constant DNA input of 2.0 ng), average peak height ratios did not vary significantly, and SD values remained constant. Discrepancies in results obtained from the two

studies may be because of sample size and the amount of template DNA used. Our study utilized only nine samples of unknown quantities while theirs used five replicates of each sample amplified at five different reaction volumes (5, 10, 15, 25, and 50 uL), using 2 ng of DNA. A greater sampling in this study as well as using quantified samples may yield results supporting the conclusions of Gaines et al. (8).

Samples with Allelic Drop-Out

Another study examined the effects of half-reaction amplification on samples containing drop-out when originally amplified at full-reaction. Forty-one percent (9 of 22) of the samples in the first amplification (13 uL) produced full profiles. However, 95.5% (21 of 22) of samples in the second amplification (also 13 uL) produced full profiles. This discrepancy may indicate that the first amplification simply was not as successful because of a variety of factors (i.e., laboratory temperature fluctuations, poor pipetting techniques of master mix and/or sample, low template volume). The third amplification (12 uL volume) produced full profiles for 91% (20 of 22) of the samples. In total, 66 amplifications were performed resulting in 50 full profiles (75.8%). Figure 4 displays profiles for one sample originally amplified at full-reaction and again after using the half-reaction protocol. Peak heights (rfu) are displayed as the second number below each allele. Average peak heights across all loci were similar for reaction volumes of 12 uL (979.01 rfu) and 13 uL (782.75 rfu) and, again, average peak
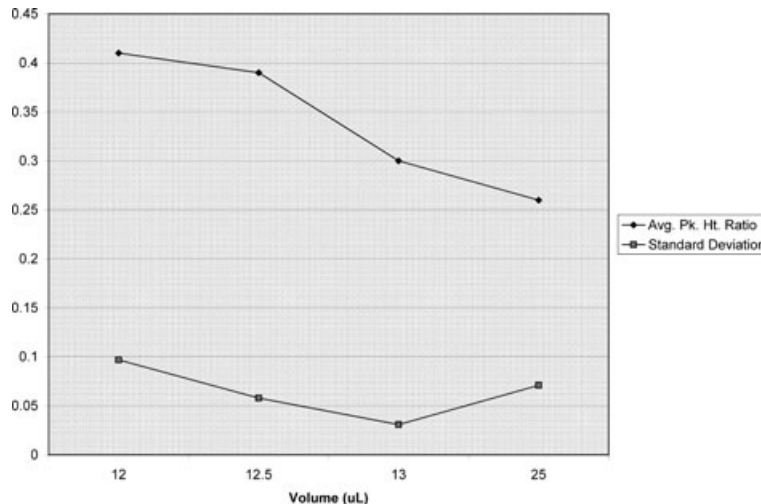

FIG. 3—*Effects on peak height ratios and standard deviations when using reduced volume amplification on samples with a peak height imbalance at the D5S818 locus.*

FIG. 4—PowerPlex® 16 electropherograms for sample ''7035'' amplified at full-reaction (A) and half-reaction (B). Panel B does not include an rfu scale; however, these values are noted as the second number below each allele.

heights increased as total amplification volumes decreased. Previous studies have found that lowering the reaction volume increases the concentration of amplified products as well as the total volume of PCR product (8). These increases allow for the amplification of less DNA template while still obtaining a full quality profile.

It is worth noting that the second amplification (13 uL) of sample "6915" contained a two-person mixed profile. The contributor of the minor profile could not be determined. Neither the extract, amplification negative control nor the first or third amplifications of the sample were contaminated. This indicates that contamination may have occurred during the second amplification set-up step, poor lab technique was involved, or samples were of low quality. This incident stresses the importance of closely following laboratory protocol to prevent contamination. It also demonstrates the high sensitivity and susceptibility to contamination associated with DNA amplification procedures.

Sensitivity

A sensitivity study was conducted after initial findings indicated that using a reduced reaction volume may allow for the amplification of less DNA template. Data showed a DNA input range of 0.25–1.0 ng was optimal for half-reaction amplification using PowerPlex® 16. Amplifying more than 1.0 ng of DNA produced peak heights above 5000 rfu that resulted in pull-up, split peaks, and/or increased stutter. Amplifying <0.25 ng of DNA resulted in allelic drop-out and peak height imbalances with ratios less than 50%. Original sensitivity studies conducted at MUFSC showed an input range of 1–4 ng was ideal for PowerPlex® 16 full-reaction amplifications. Therefore, this study demonstrates that half-reaction amplification may be more sensitive than full-reaction amplification and may allow for the amplification of less template DNA. Results from the sensitivity study also support results seen in our initial studies in which low level DNA samples (<0.4 ng) produced full profiles. A lower sensitivity is helpful for the analysis of samples with low levels of DNA or when sample quantities are limited. Less sample volume may be consumed and more of the sample will remain for additional testing if need-be.

Application to Case Samples

The half-reaction protocol established in this study was applied to 467 single-source samples outsourced to the MUFSC laboratory for analysis and research purposes. Unlike original blood punch samples used in the study, these samples were buccal swabs and were quantified and normalized to 1 ng. According to earlier findings in this study, 1 ng of template DNA should be sufficient for amplification using a half-reaction volume. The amplification of 467 samples produced 438 quality profiles using a 13 uL reaction volume. Of the 29 samples that failed, only six (1.3% of the total samples) required re-amplification because of drop-out, and two other samples were re-amplified because of peak height imbalances. The remaining 21 failed samples required re-setup because of low internal lane standards (ILS) in the original injection. The 1.3% repeat rate observed in this study is a significant decrease from the 24.5% repeat rate currently seen at the MUFSC laboratory for single-source samples amplified using the full-reaction amplification. Overall, the use of the half-reaction amplification method resulted in a 94% success rate for obtaining full, quality DNA profiles in one amplification.

Conclusion

In an effort to reduce laboratory costs, this study investigated the possibility of reducing PowerPlex® 16 amplification volumes by half. Results indicate that a reaction volume of 12.5 uL produced similar if not higher quality profiles than full-reaction amplifications. In fact, the final volume can be altered to 12 uL or 13 uL, and successful amplification parameters can still be maintained. Drop-out rates decreased and peak height ratios at the D5S818 locus increased. Sensitivity results show that half-reaction amplifications may be more sensitive and allow for the amplification of less DNA template. The ideal input range of DNA was found to be 0.25–1.0 ng compared to a 1.0–4.0 ng range determined for full-reaction amplification.

The results of this study could have several positive impacts on the forensic science community and the analysis of DNA. First, half-reaction amplification is more sensitive and requires the use of less DNA template. This would allow analysts to consume less DNA for testing and retain a larger amount of sample for additional testing if needed. Second, obtaining quality DNA profiles from a single half-reaction amplification step would reduce the number of samples needing re-amplification because of drop-out or imbalanced peak heights. Because samples would not require a second or third re-amplification step, the chance for contamination is also decreased because of less sample handling. A decrease in the repeat rate and an increased success rate would result in a less time-consuming DNA analysis process and, subsequently, current DNA backlogs may be reduced. Most importantly, the use of a half-reaction protocol would allow laboratories to amplify more samples with a single amplification kit, thus lessening operational costs. Lessened costs for the laboratory could result in lower costs for the client and could possibly allow for the analysis of samples that previously proved too costly.

References

1. Frappier R, Calandro L, Shade LL. Improving forensic DNA laboratory throughput: enhanced data analysis and expert systems capability. Forensic Mag 2008;5(1):25–31.
2. Swango KL, Hudlow WR, Timken MD, Buoncristiani MR. Developmental validation of a multiplex qPCR assay for assessing the quantity and quality of nuclear DNA in forensic samples. Forensic Sci Int 2007;170(1):35–45.
3. Butler JM. Forensic DNA typing: biology and technology behind STR markers, 2nd edn. San Diego, CA: Academic Press, 2003.
4. The Promega Corporation. PowerPlex® 16 amplification kit technical manual. USA: Promega Corporation, 2002.
5. Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, et al. Validation of a 16-locus fluorescent multiplex system. J Forensic Sci 2002;47(4):773–85.
6. http://www.fbi.gov/hg/lab/fsc/backissue/july2004/standards/ss04_03_standards02.htm (accessed on December 12, 2008).
7. Westring CG, Kristinsson R, Gilbert DM, Danielson PB. Validation of reduced-scale reactions for the Quantifiler® Human DNA Kit. J Forensic Sci 2007;52(5):1035–43.
8. Gaines ML, Wojtkiewicz PW, Valentine JA, Brown CL. Reduced volume PCR amplification reactions using the AmpFISTR® Profiler Plus™ Kit. J Forensic Sci 2002;47(6):1224–37.

Additional information—reprints not available from author:
Nicole H. Hoffman, M.S.
Forensic Scientist
Indiana State Police
19411 Hwy 41 North
Evansville, IN 47725
E-mail: Nhoffman2@isp.in.gov

TECHNICAL NOTE
CRIMINALISTICS

Jon Wilson,[1] *M.S.F.S.; Valerie Fuller,*[1] *Ph.D.; Gifty Benson,*[2] *M.S.F.S., D.D.S.; Denise Juroske,*[2,†]
M.S.F.S.; Eric Duvall,[2] *M.S.F.S.; Jun Fu,*[2] *Ph.D.; Jane Pritchard,*[2] *M.T.(ASCP), C.L.S.(MB);*
and Robert W. Allen,[2] *Ph.D.*

Molecular Assay for Screening and Quantifying DNA in Biological Evidence: The Modified Q-TAT Assay*

ABSTRACT: A method is described for the quantitation of total human and male DNA. Q-TAT utilizes end-point, multiplex polymerase chain reaction (PCR) amplification of the amelogenin and SRY loci to quantify DNA and incorporates a cloned nonhuman template to detect PCR inhibition. Standard curves of fluorescence from amelogenin or SRY amplicons were generated from amplification of known amounts of NIST traceable SRM-female or SRM-male DNA. Curves showed good linearity up to 500 pg of SRM-template ($R^2 > 0.99$) and reliably estimated total and male DNA content in casework samples. The nonhuman pRL_{null} template included in each PCR was a sensitive indicator of known PCR inhibitors including EDTA, hemin, blue denim dye, and humic acid. Finally, the SRY amplicon was a sensitive indicator of male DNA and, in mixtures, could reliably estimate male DNA present in an excess of female DNA. The Q-TAT multiplex is a reliable quantitation method for forensic DNA typing.

KEYWORDS: forensic science, human DNA quantitation, PCR inhibitors, gender typing, STR analysis, amelogenin, capillary electrophoresis

Quantitation methods used in forensic DNA typing laboratories have evolved with changes in typing methods to make them faster and more sensitive to enhance turn-around times and the overall quality of results. Underscoring the importance of accurate, human-specific quantitation methods, standards mandating quantitation of human DNA recovered from evidentiary samples were added to accreditation programs such as ASCLD/LAB and FQS (1). Reasons for the mandate include a compromise on the quality of STR/DNA profiles produced when too much or too little DNA template is amplified (2), as well as the requirement for a forensic laboratory to preserve as much DNA as possible for confirmatory testing by a different laboratory, if desired. For these reasons, laboratories have moved away from nonspecific quantitation methods such as spectrophotometry/fluorometry and yield gels to human-specific methods involving quantitative hybridization of slot blots (i.e., Quantiblot) (3), and more recently, to quantitation methods employing polymerase chain reaction (PCR) amplification with concentration estimates made either postamplification (4–7) or in real-time fashion (qPCR) (8–14). We described one such PCR-based assay known as Q-TAT (7) that quantified human DNA in a sample by amplifying the amelogenin locus on the X and Y chromosomes using fluorescent primers included in the Sex Typing kit available from Promega Corp. (Madison, WI). Following

amplification, products were separated using capillary electrophoresis, and fluorescence contained within amelogenin amplicons from unknowns was compared with fluorescence in amplicons produced from serial dilutions of known amounts of human DNA and used to prepare a standard curve (7).

An ideal DNA quantitation method should provide sufficient information to an analyst about the DNA recovered from a forensic sample, so that informed decisions can be made about how best to proceed with typing to obtain as much probative information from the item as possible. For example, knowing that a sexual assault sample does not contain male DNA could allow an analyst to ignore processing that sample rather than wasting time and resources processing it for little or no promise of developing the profile of an assailant. Likewise, knowing a sample contains low amounts of male DNA might direct an analyst to decide to perform Y-STR typing rather than autosomal STR typing on the sample and greatly enhance the chances of obtaining probative information for the prosecution of a suspect. Finally, an assay that reveals the presence of PCR inhibitors or degraded DNA template allows an analyst to take extraordinary purification steps to remove inhibitors or decide to amplify STR loci using one of the newer "mini-STR" kits more appropriate for degraded DNA (15,16), again enhancing the chances of obtaining probative information for the investigation and prosecution of the case.

This article presents further enhancements to the Q-TAT assay originally published in 2006 (7) that simply measured fluorescence incorporated into amplicons from the amelogenin locus. Among the enhancements are the inclusion of an additional PCR target harbored within the SRY gene on the Y chromosome and also the inclusion of the luciferase gene from the marine coelenterate known as the Sea Pansy (*Renilla rentiformis*). The latter template was added to

[1]Police Laboratory, Tulsa Police Department, Tulsa, OK.
[2]Department of Forensic Science, Center for Health Sciences, Oklahoma State University, Tulsa, OK.

PCRs as an internal amplification control supplied as a recombinant plasmid to detect PCR inhibitors. The reliability and sensitivity of this new improved multiplex are presented in this report.

Materials and Methods

DNA Extraction

DNA isolated from blood samples from men were sometimes used to create standard curves for the Q-TAT assay. DNA samples recovered from blood provided by alleged fathers participating in parentage testing by RFLP analysis in the 1990s had been archived for possible research use with anonymity for the donor assured. DNA was extracted using routine methods employing SDS (2% wt/vol) and proteinase K (200 ug/ml) in TNE buffer (10 mM Tris-Cl, pH 8.0 with 1 mM EDTA and 0.2 M NaCl) overnight at 37°C. The digest was extracted with phenol:CHCl$_3$:isoamyl alcohol (9:0.96:0.04 vol/vol/vol), and the aqueous phase was further extracted with chloroform:isoamyl alcohol (24:1 vol/vol) (17). DNA was precipitated with 95% ethanol, and the nucleic acid was recovered using a disposable inoculating loop prior to resuspension in 10 mM Tris-Cl pH 8.3 + 0.1 mM EDTA (TE-4). Several such samples of DNA were pooled and quantified using spectrophotometry at 260 and 280 nm as well as using rough quantitation based upon ethidium bromide mediated fluorescence of DNA samples from unknowns compared to quantitation standards in yield gels.

NIST traceable standards of DNA from male to female donors (SRM-male or SRM-female; NIST#2372) were purchased from the National Institute for Standards in Technology and were used for the preparation of standard curves and as internal controls run with each assay. The DNA concentration listed on the certificate of analysis that accompanied each sample (Sample A is from a man, and Sample B is from a woman) was considered accurate and used as the basis for dilutions needed for controls or standard curve samples.

PCR Amplification and Genetic Analysis

Samples were amplified using an ABI 9700 thermal cycler (Applied Biosystems, Foster City, CA). The concentration estimates for DNA in casework samples were derived from reactions containing either 1 μL of undiluted sample or from a second amplification with 1 μL of template diluted 10-fold. If fluorescence from amelogenin or SRY amplicons in both reactions was above 50 RFU but below 7000 RFU, estimates from both reactions were averaged to estimate DNA concentration. Otherwise, only the sample with on-scale fluorescence was used for the estimate.

The cycling parameters for the Q-TAT multiplex assay consist of a three temperature cycle, comprised of 10 sec at 98°C, 60 sec at 55°C, and 30 sec at 72°C. This 3-step cycle was repeated 30 times and was followed by a 10-min hold at 60°C, and then a final infinite hold at room temperature (25°C). The cycling parameters were applied to 12.5 μL amplifications using Taq DNA polymerase supplied as one of two forms: One Taq polymerase used was supplied as a twofold concentrated PCR mixture containing dNTPs, buffering components, and Taq DNA polymerase (2× GoTaq® Colorless Master Mix; Promega Corporation). When using the 2× GoTaq® polymerase, PCRs consisted of 1.25 μL of 10× primers (1 μM final amelogenin and SRY primers, and 0.1 μM final primers for the pRL$_{null}$ plasmid), 1 μL of DNA, 1 μL of pRL$_{null}$ plasmid DNA at 0.5–1.0 pg/μL, 7.5 μL of 2× GoTaq® polymerase, and water to 12.5 μL final volume. PCRs prepared using AmpliTaq Gold® (Applied Biosystems) contained primers and template DNA in the same concentrations as described for use with the 2×

GoTaq® polymerase, but in addition, Gold-STR reaction buffer (Promega Corp.) at 1× final concentration was included in PCRs, and AmpliTaq Gold® was added at 0.25 U per 12.5 μL reaction.

The pRL$_{null}$ plasmid (Promega Corp.) consists of a cDNA copy of the mRNA (~932 nucleotides) encoding the luciferase enzyme of the marine coelenterate known as the Sea Pansy (*R. rentiformis*) (18) cloned into an *Escherichia coli* plasmid vector conferring ampicillin resistance to the host and containing a multiple cloning site enhancement as well as promoters and other sequences in the plasmid to facilitate expression. A region of the cDNA insert was chosen as the pRL$_{null}$ amplicon based upon primers that exhibited a G:C ratio, and hence T_m, that were comparable to the other primers for amelogenin and SRY templates and also direct the amplification of a ~200-bp amplicon that is similar yet distinct from the amelogenin products. The plasmid was added to PCRs as supplied from Promega Corp. without restriction digestion or any other pretreatment other than dilution with TE-4.

All six of the primer sequences were synthesized by Invitrogen, Inc. (Carlsbad, CA). They included three reverse primer sequences labeled with the fluorophore dye FAM and three unlabeled forward primer sequences. The sequences for the forward and reverse primers were:

> AMEL forward: 5′ – ACC TCA TCC TGG GCA CCC TGG – 3′
> AMEL reverse (FAM labeled): 5′ – AGG CTT GAG GCC AAC CAT CAG – 3′
> SRY forward: 5′ – ACG AAA GCC ACA CAC TCA AGA AT – 3′
> SRY reverse (FAM labeled): 5′ – CTA CAG CTT TGT CCA GTG GC – 3′
> pRL forward: 5′ – AAG GTG GTA AAC CTG ACG TTG – 3′
> pRL reverse (FAM labeled): 5′ – TTC ATC AGG TGC ATC TTC TTG – 3′

One μL of amplification products from the multiplex PCR were mixed with 25 μL formamide (HiDi; Applied Biosystems) containing GS500 size standards labeled with LIZ dye (Applied Biosystems) according to recommendations of the supplier. The samples were then subjected to electrophoresis and fluorescent analysis with the aid of an ABI Prism 310 Genetic Analyzer (Applied Biosystems). Analysis of fluorescence was performed using either Gene-Scan/Genotyper (ver. 2.5.2) or GeneMapper ID software (version 3.2, Applied Biosystems, Inc.). Electropherograms contained products of 110 bp (SRY amplicon), 200 bp (pRL$_{null}$ amplicon), 210 bp (Amel-X amplicon), and 216 bp (Amel-Y amplicon). Detection threshold for amplicon fluorescence was set at 50 RFU.

In the initial work of Allen and Fuller (7), a standard curve was prepared for each assay of unknowns. However, recent experiments (19) have demonstrated that standard curves are sufficiently reproducible so that a single curve (or average of several curves) can be used repeatedly as long as the curve provides estimated quantities of DNA in replicate internal male and female NIST traceable controls (located throughout each Q-TAT run) that are within ±30% of their known values; the operational coefficient of variation (CV) for the assay (7,19).

Standard curves were prepared by amplifying amounts of SRM DNA ranging from 32.5 to 1000 pg. Peak height fluorescence contained within the Amel-X amplicon from SRM-female DNA or Amel-X + Y from SRM-male DNA (for total human estimates) or from the SRY amplicon from SRM-male DNA (for male-only estimates) was used to create the standard curve. The only time a new standard curve was produced was when new lots of primers were received by the laboratory.

Data Analysis

For estimation of the DNA concentrations in unknowns, the relative fluorescence in the Amel-X (or Amel-X + Y) or SRY amplicons was compared with the fluorescence in Amel-X (or Amel-X + Y) and SRY products amplified from known amounts of SRM DNA. An Excel spreadsheet was created to automate much of this process (i.e., importing of the peak height data from the genetic analyzer, reporting the R^2 value of the resulting standard curve, and checking the internal SRM male and female controls against target values to ensure their estimates are within the 30% window of acceptability).

Inhibitor Studies

Known inhibitors were diluted with TE-4 and added to Q-TAT reactions containing 500 pg of male DNA prior to cycling. Following capillary electrophoresis, peak heights associated with the pRL amplicon were quantified. Inhibitors tested included EDTA (disodium salt #ED2SS from Sigma, St. Louis, MO), hemin (#51280 from Fluka Chemicals, Milwaukee, WI), humic acid (#H16752 from Sigma), and denim blue dye (RIT-Denim Blue[36] dye; Phoenix Brands, LLC, Stamford, CT) supplied as a "concentrate" in liquid form but without any reference to the actual dye concentration.

Results

Considerations for the Standard Curve

In the study of Allen and Fuller (7), the basic parameters associated with quantifying human DNA recovered from a sample using the Q-TAT assay were investigated. The first-generation assay targeted the amelogenin locus, and the strategy for quantitation was to use fluorescence contained within amelogenin amplicons produced from known quantities of DNA to create a standard curve that could be used to estimate DNA in unknowns (7).

With the addition of two additional PCR targets to the assay (i.e., SRY and pRL), the characteristics of standard curve(s) produced with the multiplex were re-evaluated. Electropherograms of DNA fragments amplified from SRM-female and SRM-male DNA samples using the multiplex Q-TAT assay are shown in Fig. 1 with the different amplicons labeled with the Genotyper macro. For total human DNA, we elected to quantify the fluorescence in the Amel-X amplicon amplified from SRM-female genomic DNA (Fig. 1). However, it should be noted that the SRM-male DNA could also be used to generate the standard curve for total human DNA by summing fluorescence in the Amel-X and Amel-Y amplicons with little to no consequence to the accuracy of the Q-TAT assay (7,19–21). A standard curve for total DNA is shown in Fig. 2, and it is seen that the curve deviates from linearity between 400 and 500 pg of SRM-female DNA (R^2 value of only 0.93). The less than desirable R^2 value can be traced to the plateau that occurs in the curve between the 500 and 1000 pg standards because of saturation of the CCD camera with fluorescence (as evidenced by truncated peaks in the electropherogram) (not shown). If the 1000-pg point is eliminated from the curve, the R^2 value increases to >0.99 (Fig. 3A), indicating a good fit of the data to a line. Similar results are observed for male DNA when fluorescence in the SRY amplicon produced from concentrations of SRM-male DNA from 63 to 500 pg is used (Fig. 3B). Based upon the data in Fig. 3, standard curves can be produced from SRM-female to SRM-male DNA that are suitable for estimating total and/or male DNA in unknown samples. It should also be noted that reducing the electrokinetic injection time on the genetic analyzer from 5 to 3 sec improved the R^2 value of the standard curve that includes the 1-ng sample.
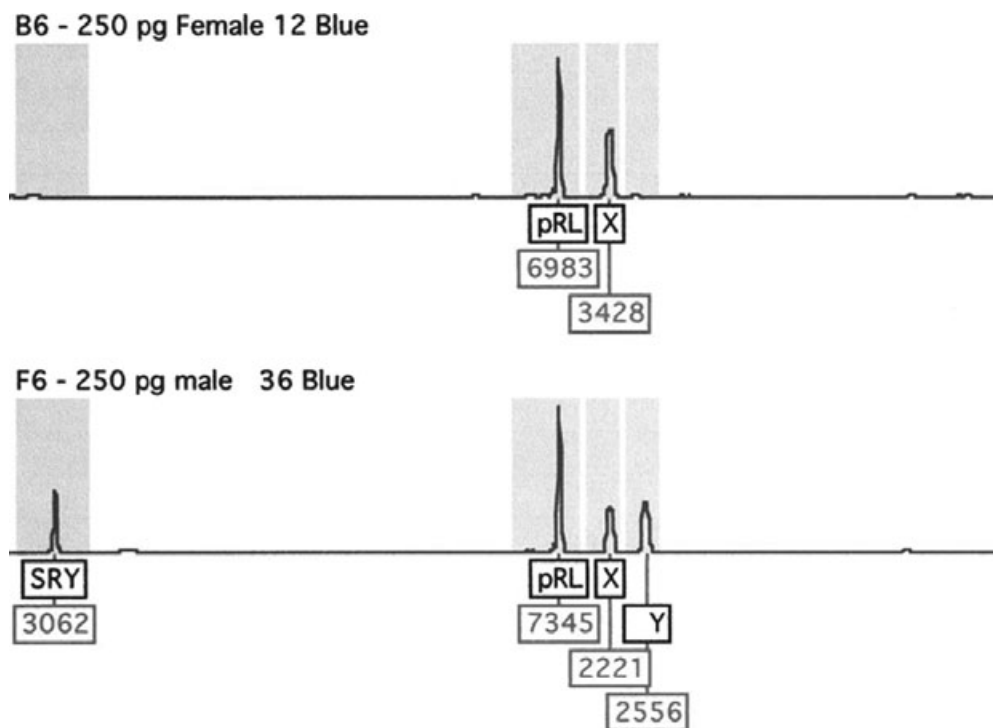

FIG. 1—*Electropherogram of amplicons produced using the Q-TAT assay from SRM-female to SRM-male DNA. A typical electropherogram produced using a Genotyper® 2.5.2 macro created to correctly estimate the size and to label each of the four amplicons produced with the Q-TAT assay. Amplicons shown were produced from 250-pg SRM-male DNA (top panel) and 250-pg SRM-female DNA (bottom panel). Amplicons shown include: SRY, pRL_{null}, Amel-X, and Amel-Y.*

The improved R^2 is likely because of the reduction in amplicon (and therefore fluorescence) entering the capillary during the reduced injection time thereby potentially extending the dynamic range of the assay. However, a reduced injection time will also likely reduce the sensitivity of the assay (not shown).
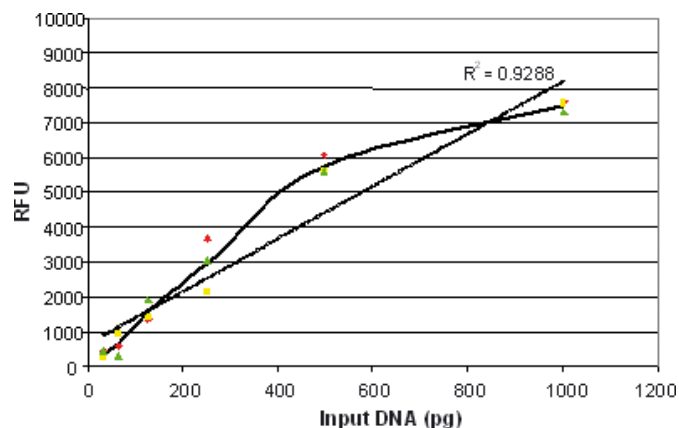

FIG. 2—*Standard curve produced from SRM-female DNA. Fluorescence gathered from known concentrations of SRM-female DNA using the Amel-X amplicon. SRM-female DNA was serially diluted twofold from 1000 to 32 pg/μL. One μL aliquots of each diluted template were amplified three different times using the Q-TAT assay. Data points were fit to a curve and to a straight line (R^2 value of 0.9288).*
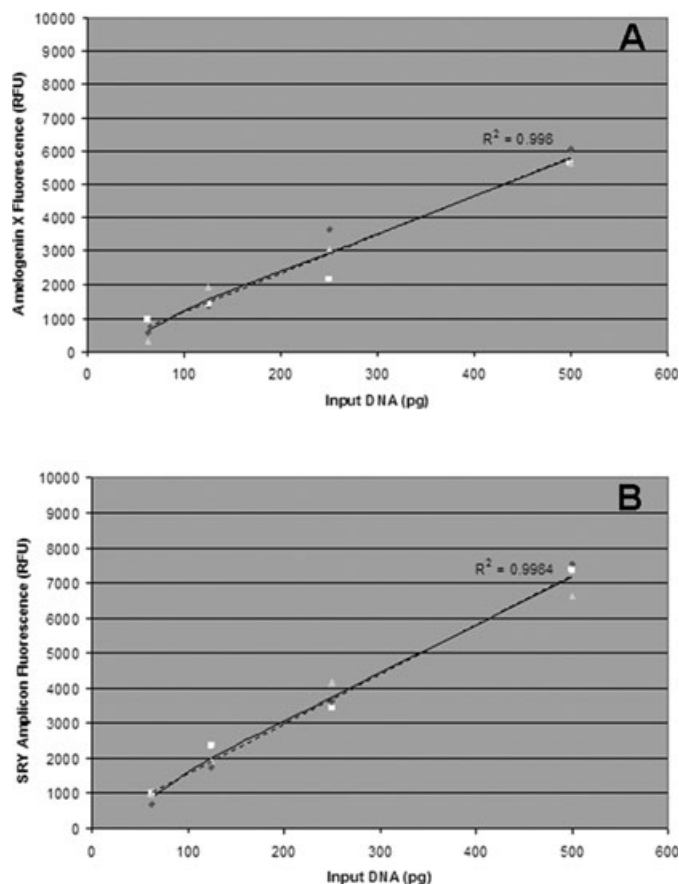
FIG. 3—*Standard curves omitting the 1000-pg sample produced for total and male DNA. SRM-female DNA (Panel A) or SRM-male DNA (Panel B) was serially diluted twofold from 500 to 63 pg/μL. One μL aliquots of each dilution of template were then amplified three different times using the Q-TAT assay, and the RFU data points were fit to a straight line.*

Inhibitor Studies

The rationale for the pRL_{null} plasmid/primers system is that a compound that will inhibit amplification of the STR loci in a human DNA sample will also inhibit the amplification of the pRL_{null} plasmid present in each PCR. Therefore, if fluorescence associated with the pRL_{null} amplicon is visualized in normal amounts in an electropherogram that lacks SRY and/or amelogenin products, one could conclude that the tested sample lacks human DNA. However, if the pRL_{null} amplicon is missing as well, it would suggest amplification did not occur possibly because of the presence of an inhibitor of PCR. Known inhibitors of PCR include EDTA, hemin present in blood, blue denim dye, and humic acid, a constituent of soil. Each of these compounds was tested in different concentrations for its ability to inhibit the production of Amel-X, Amel-Y, SRY, and especially, the pRL_{null} amplicon, included in the Q-TAT assay as an indicator of inhibition. Results (Fig. 4) demonstrate that the pRL_{null} indicator system is the most sensitive target in the multiplex to inhibition by any of the known inhibitors. Interestingly, amplification of the SRY target is least inhibited by compounds such as hemin, blue denim dye, and low concentrations of humic acid (Fig. 4).

Quantitation of Casework Samples

The reliability of the Q-TAT multiplex for quantifying total human and male-only DNA was evaluated in several ways. In one series of experiments, precision of the Q-TAT assay was investigated. SRM-male and SRM-female samples were serially diluted twofold. Each dilution was amplified multiple times in the Q-TAT assay, and the amount of total or male DNA was quantified (Fig. 5).

Each data point in Fig. 5 represents a separate amplification of one dilution of input DNA template, and the number of replicates for each dilution varied between 9 and 23. For quantitation of DNA in each dilution, estimates were made using standard curves prepared from Amel-X fluorescence (SRM-female DNA) for total DNA or SRY (SRM-male DNA) for male DNA. Estimates of DNA in the dilutions of SRM-male and SRM-female samples were made using fluorescence in the Amel-X (for total female), Amel-X + Y (for total male), or the SRY amplicons (for total male) (Fig. 5). Comparison of the spread of data points for male estimates made using fluorescence in amelogenin markers versus SRY reveals a larger spread of points when amelogenin is used (Fig. 5). However, statistical analysis of the two methods using ANOVA showed no significant difference in the estimates for male samples at lower input amounts of DNA ($p > 0.05$ for the 1:8 and 1:16 dilutions). There was a significant difference in the variability of concentration estimates for the 1:2 and 1:4 dilutions. Whether calculating total human DNA in male or female samples at every dilution using amelogenin fluorescence or calculating the male-only DNA content of samples using fluorescence in SRY, the CV averaged 18%.

The ability of the Q-TAT assay to provide reliable estimates of DNA extracted from casework samples ($N = 47$) was assessed by amplifying 1 ng of DNA based upon the Q-TAT estimate for total DNA using Amel-X + Y for profiling with the Identifiler® (Applied Biosystems) multiplex ($N = 34$) or SRY for profiling with the Y-filer® multiplex ($N = 14$) according to manufacturer's instructions. Samples subjected to STR typing included only those with concentration estimates of at least 100 pg/μL. Among the 34 samples analyzed at the D3S1358 locus, the average fluorescence was 3081 RFU/allele (±1753 RFU), whereas for DYS393 alleles,
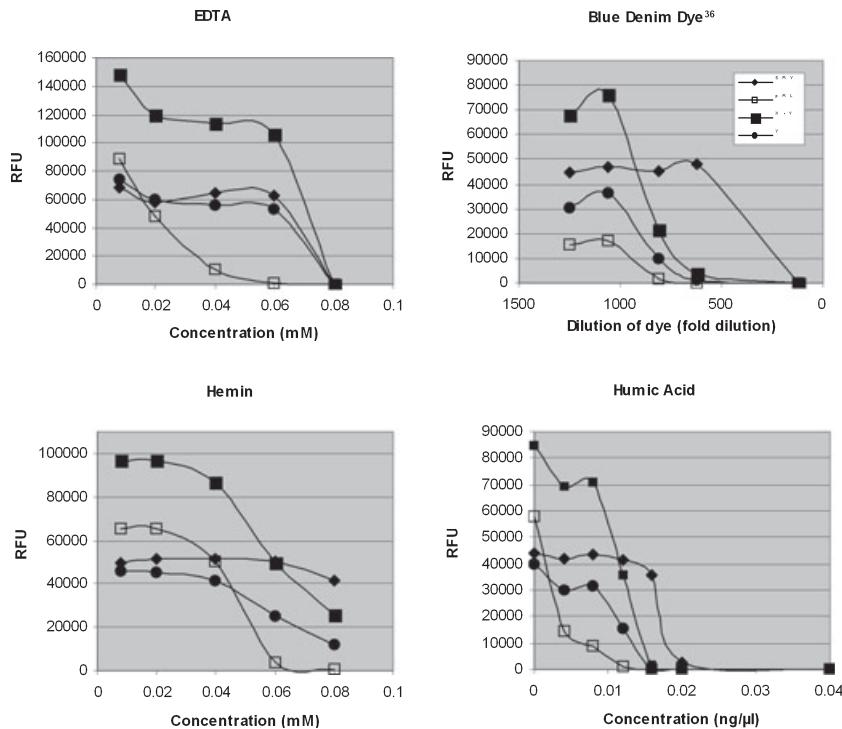
FIG. 4—*Inhibition of Q-TAT assay by different compounds. Five hundred pg of well characterized male DNA extracted from peripheral blood samples and 0.5 pg of pRL$_{null}$ plasmid were amplified in the Q-TAT assay with or without different inhibitory compounds added in a 1 μL volume to make the final concentrations shown. Amplicons were analyzed on the genetic analyzer, and fluorescence in each product was quantified against a standard curve prepared from pooled male DNA isolated from peripheral blood from several male donors. Products quantified included: SRY amplicon (solid diamonds), Amel-X + Y amplicons (solid squares), Amel-Y amplicon (solid circles), and pRL$_{null}$ (open squares).*
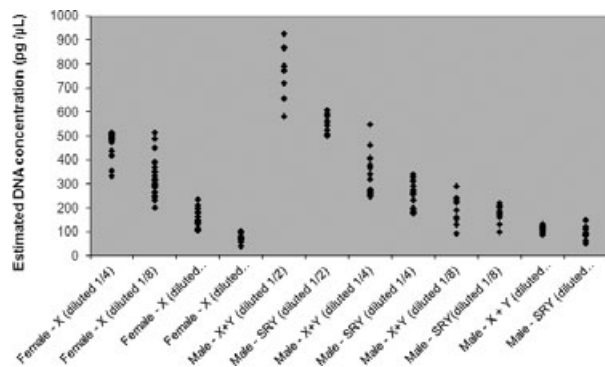
FIG. 5—*Precision of the Q-TAT assay for total and male DNA quantitation. The scatter plot shows the ability of the Q-TAT assay to reproducibly estimate the DNA concentration of male and female samples by comparing the Amel-X, Amel-X + Y, or SRY amplicons to a total human or male-only standard curve prepared using SRM-female DNA or SRM-male, respectively. SRM DNA samples were serially diluted twofold, and each dilution was amplified multiple times. The minimum number of amplifications performed for any dilution was 9, and the maximum number for any dilution was 23. The SRM DNA tested, its dilution, and the amplicon used for quantitation (i.e., Amel-X alone, Amel-X + Y, or SRY) are shown along the X axis.*

FIG. 6—*Fluorescence (RFU) in STR amplicons produced from DNA from casework quantified by Q-TAT. A total of 34 casework samples (16 evidentiary and 18 reference) were quantified using Q-TAT for total human DNA for analysis using the Identifiler® multiplex. Thirteen casework samples (seven evidentiary and six reference) were quantified for male DNA for analysis using the Y-Filer® multiplex. In all cases, 1 ng (based upon a Q-TAT estimate) of template was amplified, and fluorescence in STR alleles was quantified. Shown are the average RFU/allele for the D3S1358 locus (solid diamonds) and the DYS393 locus (open squares).*

average fluorescence was 2180 RFU (±1586 RFU) (Fig. 6). Over 85% of the autosomal profiles exhibited RFU/allele values between 1000 and 6000 RFU, whereas over 60% of DYS393 alleles fell within this range. In no case was the presence of PCR inhibitors suggested from fluorescence contained within the pRL$_{null}$ amplicon, and no sample failed to produce a complete STR profile with the respective STR typing kit (not shown). In addition, all results met reporting criteria (above 150 RFU threshold and no truncated peaks).
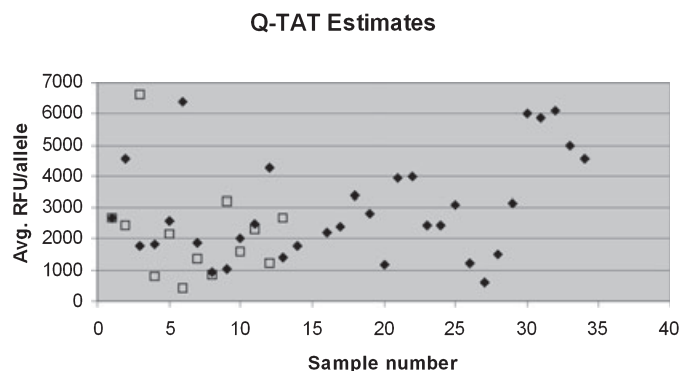
Quantitation of Male and Female DNA in Mixtures

In earlier work with the basic Q-TAT assay, the ability of the assay to accurately quantify male DNA in mixed male:female samples was evaluated (20). Two approaches were taken to evaluate the ability of the multiplex Q-TAT method to detect male DNA in the presence of an excess of female DNA. Figure 7 shows the data obtained when 125 pg of purified SRM-male DNA is mixed with

an increasing amount of SRM-female DNA. Whereas the Amel-Y amplicon fluorescence exceeds the 150 RFU threshold in the presence of a 10-fold excess of SRM-female DNA, the SRY amplicon exceeds threshold even when a 400-fold excess of SRM-female DNA is present in the sample (Fig. 7).

In a second approach, male and female whole blood was mixed in varying ratios, and then a constant volume of the mixture was spotted onto FTA paper. We then amplified 1.5-mm punches of the stain and evaluated the fluorescence in the Amel-Y and SRY amplicons (Fig. 8). Above a 20-fold excess volume of female blood, the Amel-Y fluorescence drops below the 150 RFU threshold, whereas SRY fluorescence drops below the threshold in the presence of 100-fold volume excess of female blood (Fig. 8).
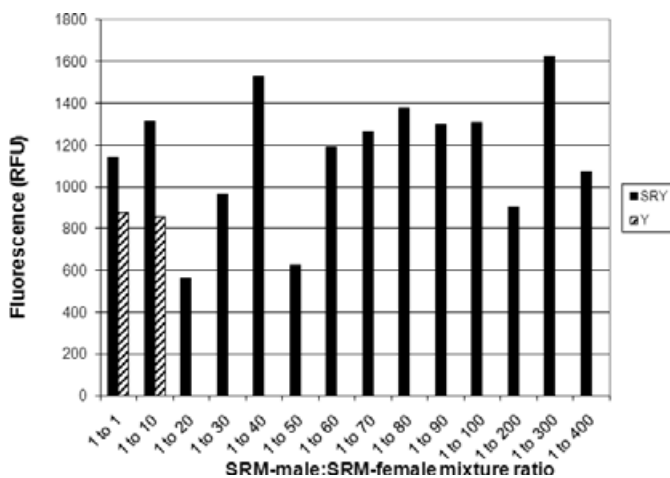
FIG. 7—*Sensitivity of the Improved Multiplex Assay to detect male DNA in the presence of excess female DNA. A constant amount of SRM-male DNA (125 pg) was mixed with increasing amounts of SRM-female DNA and then amplified with the Q-TAT assay. The solid bars in the figure represent fluorescence incorporated into the SRY amplicon. The cross-hatched bars represent fluorescence incorporated into the Amel-Y amplicon.*
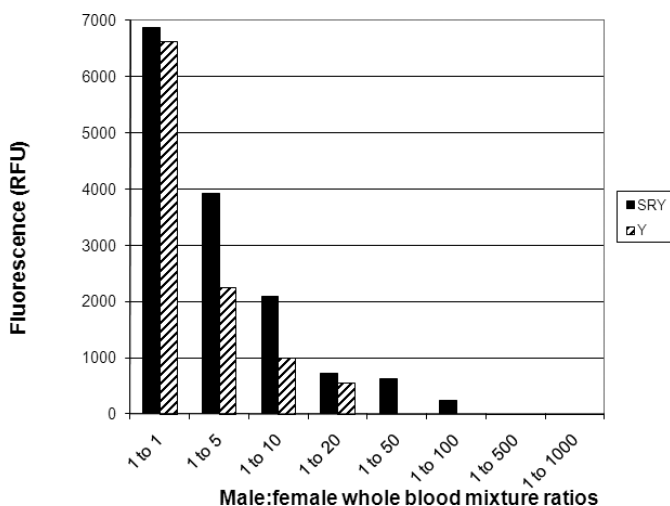
FIG. 8—*Sensitivity of the Q-TAT assay for detecting male DNA in the presence of excess female DNA in blood mixtures. Male and female whole blood was mixed in varying ratios, dried on FTA paper, and extracted from a constant sample size (1.5-mm hole punch) from each mixture. Each punch was amplified in the Q-TAT assay, and the male DNA in the sample quantified. The solid bars represent detected levels of SRY fluorescence. The cross-hatched bars represent detected levels of Amel-Y fluorescence.*

Discussion

In this study, the Q-TAT assay (7) was modified to include two additional amplification targets: One added target is the SRY gene (22,23), located on the Y chromosome and encoding a factor directing primordial gonads to develop into testes (23). Inclusion of the SRY gene in the assay served two functions. First, it provided for amplification of an additional Y chromosomal target thereby ensuring DNA from the 1–2% of men with mutations at the amelogenin locus will still be revealed as men through the SRY locus and vice versa (24–27). Second, in mixed male:female samples containing excess female DNA, quantitation of the male DNA would still be feasible, because there is no homologue to SRY on the X chromosome to compete for primers, resulting in less chance for "dropout" of the SRY amplicon (20).

Also included in the modified Q-TAT assay is a plasmid harboring the luciferase gene of the marine organism known as the Sea Pansy (*R. rentiformis*) (18) along with primers that direct the amplification of a 200-bp region of the gene. This target was included to serve as an inhibition indicator, because a constant amount of template and primers are present in each amplification reaction. When an inhibitor is detected, extraordinary steps can be taken to further "clean" the DNA sample to remove the inhibitor, or the sample may simply be diluted with the hope of reducing the inhibitory effect sufficiently to enable an STR profile to be produced.

There are several reasons for quantifying human DNA recovered from a forensic specimen. Perhaps the most important reason is to amplify an optimal amount of genomic DNA template so as to produce an optimal STR profile suitable for routine analysis and interpretation. Profiles that contain allele dropout or other artifacts resulting from the amplification of suboptimal DNA template result in additional time spent on interpretation, repeated amplifications with adjusted amounts of template, or even further processing of the DNA sample to enhance its purity and promote the production of a profile. With the backlog of casework that continues to grow, efficient processing of casework is a priority. A second reason for quantitation is the possibility of court-ordered confirmatory testing of samples by a different laboratory in any particular case. Finally, accurate characterization of human DNA in a sample can suggest the testing methodology most likely to yield probative results.

DNA typing methods in current widespread use include autosomal STR typing, autosomal mini-STR typing, Y-STR typing, and mitochondrial DNA sequencing (2). Of the current methods in use, the discriminatory power is maximal for autosomal STRs and decreases from mini-STR to Y-STR, and mitochondrial DNA sequencing. For any evidentiary sample, an analyst must decide which method holds the greatest promise of producing the best investigative answer for a case. A quantitation method that also provides information that assists in the decision making is therefore desirable. For example, information about whether male DNA exists in a sexual assault sample or not, and if so, how much is present, can provide the analyst with information that will assist in choosing whether to process that sample or not. If the sample is to be processed, choosing between autosomal versus Y-STR analysis may benefit from Q-TAT results. In this regard, as little as 1% male DNA in the presence of excess female DNA in whole blood mixtures can be quantified using the SRY amplicon. The enhanced sensitivity of SRY undoubtedly results from the lack of a comparable PCR target on the X chromosome thereby eliminating an X-linked locus to compete for primers during amplification. Such is not the case with the amelogenin locus that exists on both X and Y. In an evidentiary sample with a small amount of male DNA not amenable to enrichment through differential extraction, Y-STR

analysis might be the method of choice for profiling to ensure probative results are obtained. Likewise, if the sample contains inhibitors that were co-extracted with the DNA, a quantitation method that alerts an analyst to their presence will allow extraordinary clean-up steps to be taken to separate DNA template from inhibitors, or the sample may simply be diluted and allow a profile to be produced.

DNA quantitation methods in widespread use that meet accreditation standards include blotting methods (Quantiblot from Applied Biosystems, Inc.) and qPCR. Because Applied Biosystems is no longer making the Quantiblot kits, there has been a rush to adopt and validate qPCR technology and available quantitation kits to replace blotting as the principal quantitation method used by forensic laboratories. qPCR kits currently available include Quantifiler® (for total human DNA), Quantifiler-Y® (for male DNA), and Quantifiler Duo® (for both total and male DNA) from Applied Biosystems and Plexor® (for both total and male DNA) from Promega Corp.

To establish qPCR technology in the laboratory, monies must be devoted to instrumentation and kits used for validation. In addition, additional bench space, quality assurance, and training will be needed to support the use of the validated technology.

The Q-TAT methodology offers an approach for DNA quantitation that differs from qPCR in several significant ways: First, the assay involves post-PCR quantitation. Quantitation occurs only after a PCR amplification process that consists of 30 cycles and requires about 1.5 h to complete, followed by 17–20 min of capillary electrophoresis to separate the important amplicons and allow for their quantitation. The completion time for the assay can be reduced however with the substitution of an all-inclusive PCR mix available from Promega Corp. (GoTaq) in place of GoldSTR buffer (Promega Corp.) and AmpliTaq Gold (Applied Biosystems). These modifications have trimmed the total time required to amplify and analyze Q-TAT reaction products to about 1.5 h at a cost of about \$1.00/sample excluding costs associated with capillary electrophoresis (19).

The modified Q-TAT assay, like its predecessor, is not as sensitive as qPCR for detecting very low amounts of human DNA. The lowest reliable concentration of DNA amenable to accurate quantitation with Q-TAT is about 50 pg, which is about 10-fold less sensitive than qPCR (7). However, this limit in sensitivity is adequate in our opinion as samples containing lower amounts of human DNA would be unlikely to produce a profile with any of the normal methods used in the forensic laboratory. Samples containing a large amount of DNA pose an additional problem for the Q-TAT assay because of its limited dynamic range, with a ceiling of about 500 pg of DNA as the maximum amount that can be accurately quantified. However, if DNA from a sample is tested undiluted and also diluted 10–20-fold initially, the effective dynamic range for quantitation is extended well into the ng/μL range. With the inexpensive cost of the assay, the availability of robotics or at least multi-channel pipetting devices, and the widespread use of multicapillary genetic analyzers, quantifying each unknown DNA sample undiluted and diluted 10–20-fold would create little difficulty for a forensic laboratory. We have now used the Q-TAT assay in practice for almost a year and find this strategy workable and effective for processing samples and producing quality STR results.

The SRY amplicon, because of its reduced size (110 bp) relative to the Amel-Y product (216 bp), may also serve as an indicator system for DNA degradation. Although not a sensitive indicator of degradation, fluorescence in the SRY amplicon disappears more slowly than Amel-Y in evidentiary samples in which autosomal STR typing results suggested degraded template as evidenced by reduced RFU or dropout of high molecular weight loci in the profile (B. Smith, unpublished observations).

Besides supplying a forensic laboratory with a reliable and inexpensive assay for DNA quantitation, the Q-TAT assay could have application as a high throughput screening tool for sexual assault evidence in backlog casework. A quick and easily automated PCR setup strategy coupled with a multicapillary electrophoresis platform would enable large numbers of samples to be quickly screened for the presence and relative abundance of male DNA in sexual assault evidence, whether inhibitors are present in any sample and if there are samples of questioned integrity. Preliminary studies have shown that Q-TAT PCRs can be produced robotically, confirming the feasibility of high throughput sample characterization (28).

References

1. DNA Advisory Board, Federal Bureau of Investigation. Quality assurance standards for forensic DNA testing laboratories. Washington DC: Federal Bureau of Investigation, 2000, http://www.fbi.gov/hq/lab/html/testinglab.htm (accessed March 13, 2009).
2. Butler JM. Forensic DNA typing, biology, technology and genetics of STR markers, 2nd edn. Burlington, MA: Elsevier Academic Press, 2005.
3. Walsh P, Valaro J, Reynolds R. A rapid chemiluminescent method for quantitation of human DNA. Nucl Acids Res 1992;20:5061–5.
4. Sifis ME, Both K, Burgoyne LA. A more sensitive method for the quantitation of genomic DNA by Alu amplification. J Forensic Sci 2002; 47:589–92.
5. Nicklas JA, Buel E. Development of an Alu based, QSY7-labeled primer PCR method for quantitation of human DNA in forensic samples. J Forensic Sci 2003;48:282–91.
6. Walker JA, Kilroy GE, Xing J, Shewale J, Sinha SK, Batzer MA. Human DNA quantitation using Alu element based polymerase chain reaction. Anal Biochem 2003;315:122–8.
7. Allen RW, Fuller V. Quantitation of human genomic DNA through amplification of the amelogenin locus. J Forensic Sci 2006;51:76–81.
8. Andreasson H, Gyllensten U, Allen M. Real-time PCR quantification of nuclear and mitochondrial DNA in forensic analysis. BioTechniques 2002;33:402–11.
9. Nicklas JA, Buel E. Development of an Alu based, real-time PCR method for quantitation of human DNA in forensic samples. J Forensic Sci 2003;48:936–44.
10. Alonso A, Marten P, Albarran C, Garcia O, Fernandez de Simon L. Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. Forensic Sci Int 2004;139:141–9.
11. Green RL, Roinestad IC, Boland C, Hennessy LK. Developmental validation of the Quantifiler real-time PCR kits for the quantification of human nuclear DNA samples. J Forensic Sci 2005;50:809–25.
12. Shewale JG, Schneida E, Wilson J, Walker JA, Batzer MA, Sinha S. Human genomic DNA quantification system, H-Quant: development and validation for use in forensic casework. J Forensic Sci 2007;52:364–70.
13. Richard ML, Frappier RH, Newman JC. Developmental validation of a real-time quantitative PCR assay for automated quantification of human DNA. J Forensic Sci 2003;48:1041–6.
14. Horseman KM, Hichey JA, Cotton RW, Landers JP, Maddox LO. Development of a human specific real-time PCR assay for simultaneous quantitation of total genomic and male DNA. J Forensic Sci 2006; 51:758–65.
15. Butler JM, Shen Y, McCord BR. The development of reduced size STR amplicons as tools for the analysis of degraded DNA. J Forensic Sci 2003;48:1054–64.
16. Chung DT, Drabek J, Opel KL, Butler JM, McCord BR. A study on the effects of degradation and template concentration on the efficiency of the STR miniplex primer sets. J Forensic Sci 2004;49:733–40.
17. Maniatis T, Fritsch EF, Sambrook J. Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1982.
18. Sherf BA, Navarro SL, Hannah RR, Wood KV. Dual-Luciferase™ reporter assay: an advanced co-reporter technology integrating firefly and Renilla luciferase assays. Promega Notes 1996;57:2–8.
19. Wilson J. Developmental validation of an improved multiplex assay for use in forensic casework: the simultaneous gender-typing and quantification of total human and male-only DNA within forensic samples (MS Thesis). Tulsa, OK: Oklahoma State University, 2008.

20. Juroske D. Quantitation of male and female DNA in mixed biological samples using quantitative amplification of the human amelogenin locus (MS Thesis). Tulsa, OK: Oklahoma State University, 2006.

21. Benson GA. Improved quantitation of human DNA using quantitative template amplification technology (MS Thesis). Tulsa, OK: Oklahoma State University, 2006.

22. McKeown B, Strickley J, Riordan A. Gender assignment by PCR of the SRY gene: an improvement on amelogenin. Proceedings of the 18th International ISFH Congress; 1999 Aug 17–21; San Francisco, CA. Prog In Forensic Genet 1999;8:433–5.

23. Berta P, Hawkins JR, Sinclair AH, Taylor A, Griffiths BL, Goodfellow PN, et al. Genetic evidence equating SRY and the testis determining factor. Nature 1990;348:448–50.

24. Santos FR, Pandya A, Tyler-Smith C. Reliability of DNA-based sex tests. Nat Genet 1998;18:103.

25. Chang YM, Peruman R, Keat PY, Yong RY, Kuehn DL, Burgoyne L. A distinct Y-STR haplotype for amelogenin negative males characterized by a large Y(p)11.2 (DYS458-MSY1-Amel-Y) deletion. J Forensic Sci 2007;166:115–20.

26. Chang YM, Burgoyne LA, Both K. Higher failure rates of amelogenin sex test in an Indian population group. J Forensic Sci 2003;48:1309–13.

27. Shewale J, Richey SL, Sinha SK. Anomalous amplification of the amelogenin locus typed by AmpFLSTR Profiler Plus amplification kit. Forensic Science Communication 2000; 2. Available at: http://www.fbi.gov/hq/lab/fsc/backissu/oct2000/shewale.htm (accessed March 13, 2009).

28. Duvall EJ. Automation of a human DNA quantitation technique using a Biomek® 2000 robotic platform (MS Thesis). Tulsa, OK: Oklahoma State University, 2008.

Additional information—reprints not available from author:
Robert W. Allen, Ph.D.
Department of Forensic Science, Center for Health Sciences
Oklahoma State University
1111 West 17th St.
Tulsa, OK 74107
USA
E-mail: robert.w.allen@okstate.edu

TECHNICAL NOTE

CRIMINALISTICS

Mark Barash,[1,†] *M.Sc.; Ayeleth Reshef,*[1] *M.Sc.; and Paul Brauner,*[1] *(ret.), M.Sc., F.F.S.Soc.*

The Use of Adhesive Tape for Recovery of DNA from Crime Scene Items

ABSTRACT: The selection of the appropriate method of collection of biological material from crime scene items can be crucial to obtaining a DNA profile. The three techniques commonly used for sampling items are: cutting, swabbing, and taping. The tape sampling technique offers an advantage, in that it enables the collection of a potentially highly informative source of DNA, shed epithelial cells, from selected areas on crime scene items (the inside fingers of a glove, for instance). Furthermore, surface collection of biological material by taping reduces co-sampling of known PCR inhibitors such as clothing dyes. The correct choice of tape for crime scene item sampling is important. Not all tapes are suitable for biological trace evidence collection as well as DNA extraction. We report on one tape that met both these criteria. Three different cases are presented which demonstrate the usefulness of adhesive tape sampling of crime items. Finally, the advantages of the tape collection technique are discussed and guidelines for preferred areas of tape sampling on various casework items are presented.

KEYWORDS: forensic science, adhesive tape sampling, trace evidence, short tandem repeats, AmpFlSTR SGM Plus, DNA typing

Apart from blood, semen, and saliva, casework items may contain a concealed source of highly informative DNA—shed epithelial cells. Items commonly submitted for profiling from crime scenes include clothing, ski masks, socks, weapons, cell phones, flashlights, glasses, watches, and jewelry, all of which may have been worn or handled, and therefore potentially contain epithelial cells.

Swab collection of biological material from human skin (1), bottles and cans (2), a car steering wheel, a writing pen, a hot dog, an electric cord, a contact lens, and a pair of gloves (3) has yielded sufficient DNA for profiling. However, the superiority of using adhesive tape for recovering epithelial cells over other methods—swabbing and cutting—has been reported for shoe insoles (4) and a baseball cap, a jacket (5), and cadavers (6).

Of four different tapes that we tested for the collection of biological material, only one was found satisfactory for DNA extraction and amplification.

We present three cases in which taping was used as the method of sampling of various crime scene items. Shed epithelial cells (and possibly other biological materials) were collected from presumed areas of high friction between the body and the worn garment or the handled object. In two cases, selected taping of different areas of the same item (a gun and a glove) yielded two different, identifiable DNA profiles.

Since its implementation in our laboratory in 2004, tape sampling has been found to be an efficient, non-destructive method for obtaining concealed DNA from routinely submitted crime scene items.

Materials and Methods

Preliminary taping experiments were carried out using four different tapes:

- 2″ × 2″ GripLifters™ (Sirchie® Fingerprint Laboratories, Youngsville, NC).
- Handi-Lifts™ (Lightning Powder® Brand Clear Tapes, Crestwood, IL).
- Frosted Lifting Tape (Sirchie® Fingerprint Laboratories, Youngsville, NC).
- Three-layer adhesive tape (Industrial Self Adhesives Limited [ISA], Nottingham, UK).

The first two tapes contained an excessive amount of glue, which prevented easy manipulation in collection, while the third formed a viscous mass ("porridge") at 95°C during the Chelex extraction (7). Only the three-layer adhesive tape enabled the trouble-free lifting of biological material that could be subsequently extracted and profiled.

To use this tape, the ruled acetate layer was first removed from the adhesive PVC and silicone paper layers of the tape. The remaining two layers of the tape were treated by UV cross-linking (6 J/cm) and cut into approximately 0.5 × 2 cm strips.

Prior to sampling, these two layers were separated. Biological material was collected with the adhesive layer, by applying a strip to the relevant area on the case item. Light pressure was applied to the tape with a gloved finger or the back of a pair of sterile forceps to maximize sampling. These samplings were carried out on areas of presumed highest friction between the body and the casework item worn (internal glove fingertips, shirt collars and seams, etc.),

[1]Forensic Biology Laboratory, Division of Identification and Forensic Science (DIFS), Israel Police National Headquarters, Jerusalem, Israel.
[†]Present address: Faculty of Health, Science and Medicine, Bond University, Gold Coast, QLD 4229, Australia.

or handled (gun and knife handles, rope ends, etc.). Each tape strip was removed and reapplied to adjacent areas, until fully "saturated" with collected material. The tape was then cut into small pieces using sterilized scissors, inserted into a 1.5-mL conical tube, and centrifuged for 20 sec at 20,000×g (spindown).

The presence of cells on tapings was verified by eosin/hematoxylyn staining of a strip of adhesive tape sampled from a glove. Nucleated cells that had adhered to the tape were microscopically visible (data not shown).

DNA extraction was carried out by the Chelex procedure (7) in a final volume of 200 μL. The extracted DNA was further purified and concentrated using the Microcon filter device, Ultracel YM-30 membrane (Millipore Corporation, Billerica, MA).

DNA yields in samples were estimated using the Real Time PCR Quantifiler™ Human DNA Quantification kit (8) and the ABI PRISM® 7000. Approximately 0.9 ng of the DNA from case items and reference samples was amplified using the AmpFlSTR® SGM Plus™ kit following the standard procedure (9). (In some cases, as little as 0.5 ng of DNA gave a full profile.) Amplified PCR products were electrophoretically separated on an ABI PRISM® 3100 Genetic Analyzer fitted with a 50 μm by 36 cm capillary, loaded with POP-4 polymer. The running conditions were as follows: 5 sec injection time, 30 min running time, and 60°C running temperature. Genemapper® 3.2 software (PE Applied Biosystems, Foster City, CA) was used for genotyping. The detection threshold was set at 60 RFU. A stochastic threshold of 200 RFU was set for the designation of homozygotes.

Cases

Case 1

After receiving a number of death threats, the mayor of a city in Israel had a grenade hurdled into his home. An eyewitness saw a man in a knitted ski mask running from the scene of the crime. Later, police investigators recovered a ski mask along the route taken by this suspect.

The ski mask was not tested using the preliminary Phadebas procedure, rather separate tapings from the *inside* of the mask were collected for DNA profiling from:

- The left corner of the mouth hole in the mask.
- The right corner of the mouth hole in the mask.
- The central area, approximately 1 cm above the mouth hole in the mask (the assumed nostril area).
- The tear duct area of both eye holes in the mask.

The following results were obtained: Tapings 1 and 2 contained a mixture of at least two DNA profiles. The major profile was of male origin (Fig. 1, taping 1). Taping 3 contained the profile of a single male and matched the major profile, from tapings 1 and 2 (Fig. 2). Taping 4 did not contain enough DNA for profiling analysis.

Buccal swabs samples were submitted from two suspects in the case. The DNA profile from the ski mask matched the profile of suspect 2. In a later case, a glove was recovered from the scene of an attempted murder of the mayor's son. This glove was also
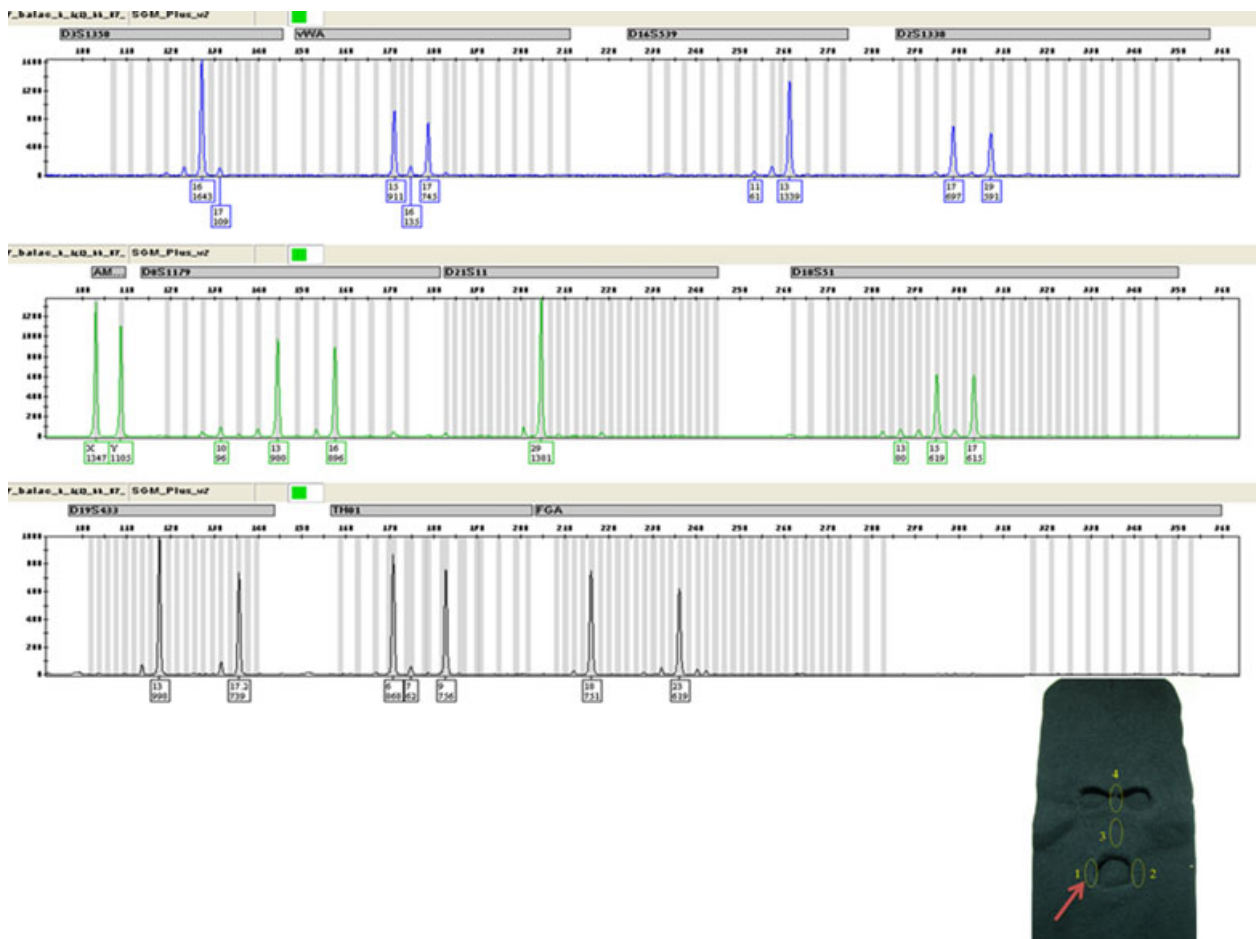

FIG. 1—*Electropherogram of the profile obtained from the tape sampling of the inside mouth hole of the ski mask indicated by the arrow.*
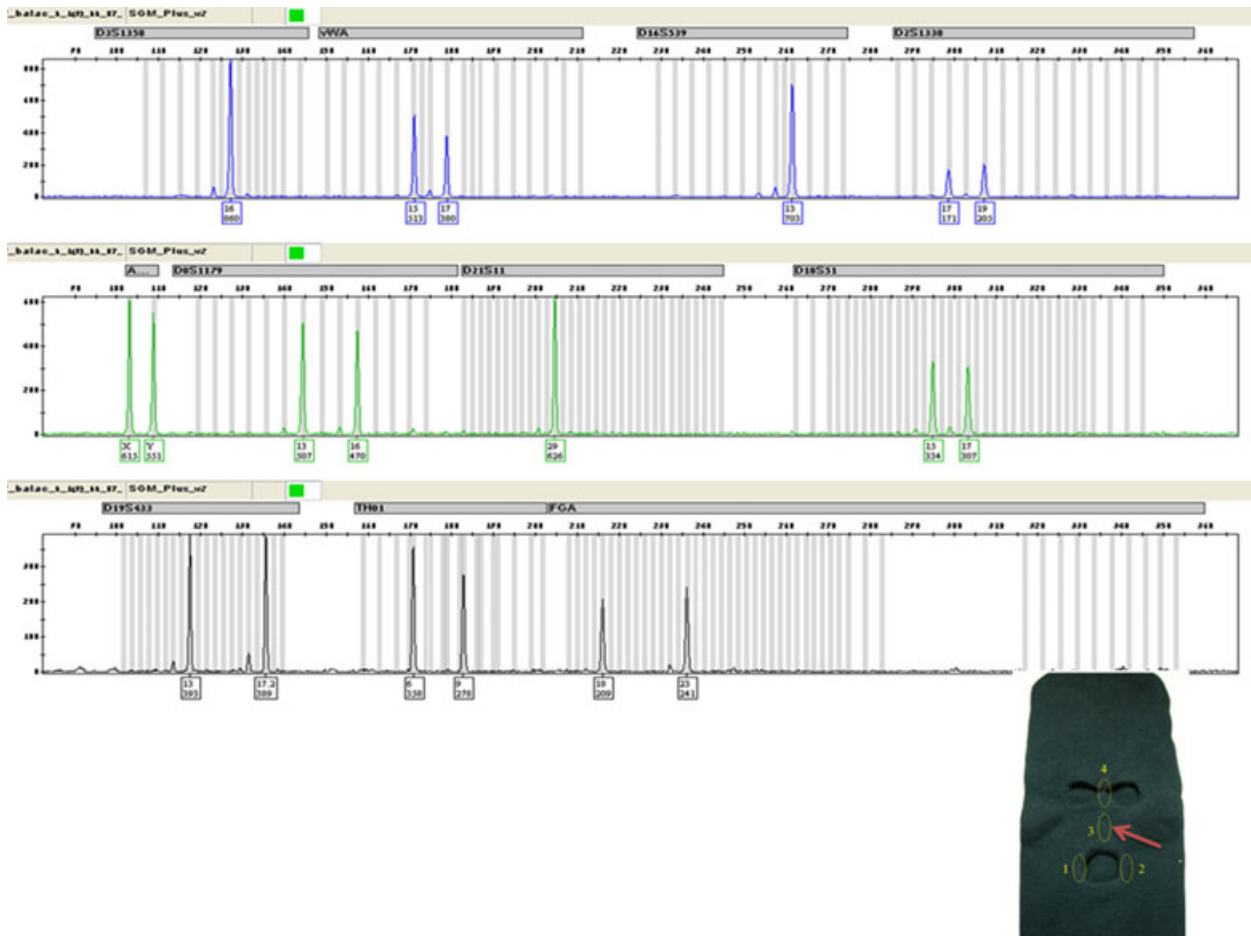
FIG. 2—*Electropherogram of the profile obtained from the tape sampling of the inside nostril area of the ski mask indicated by the arrow.*

sampled by the adhesive tape method. From the results of a DNA database search, a match was found between the profile obtained from the glove and the profile of suspect 2.

Case 2

In a search of the apartment of an individual suspected of stealing 15 electric guitars, police found a rusty, Colt Cobra revolver concealed in a sock. The suspect claimed that he neither owned nor hid the weapon but, rather, that the police must have planted the gun in his apartment.

The *inside* of the sock in which the revolver was concealed was sampled by taping in 4 areas:

- The heel.
- The toe area.
- The fingers area.
- The upper part of the sock.

Particular attention was paid to the collection of material from the stitched seams in these areas; however, the samplings contained insufficient DNA for analysis.

The revolver was sampled by taping on four areas:

- The right-hand side of the handle.
- The left-hand side of the handle.
- The trigger and hammer.
- The cylinder.

From the revolver, the following results were obtained: Taping 2 contained a mixture of at least two individuals. The major profile was of male origin and designated as male 1 (Fig. 3). Taping 3 contained a mixture of at least three individuals. The major profile was of male origin, but differing from the profile from taping 2. This profile was designated as male 2 (Fig. 4).

Taping 1 contained a DNA mixture of at least three individuals. Males 1 and 2 could not be excluded as contributors to this mixture (results not shown). Taping 4 contained a DNA mixture of at least three individuals. Male 1 could not be excluded as contributor to this mixture (results not shown).

Subsequently, in a DNA database search, a cold hit was found between the male 1 profile (Fig. 3) and a DNA profile recovered from plastic restraints in a separate armed robbery case.

Case 3

Two masked men robbed a gas station. One suspect was observed discarding a woolen glove near the crime scene. The police later recovered a glove along the suspect's presumed escape route.

The *inside* of the glove fingers were sampled as follows:

- The tip of the thumb.
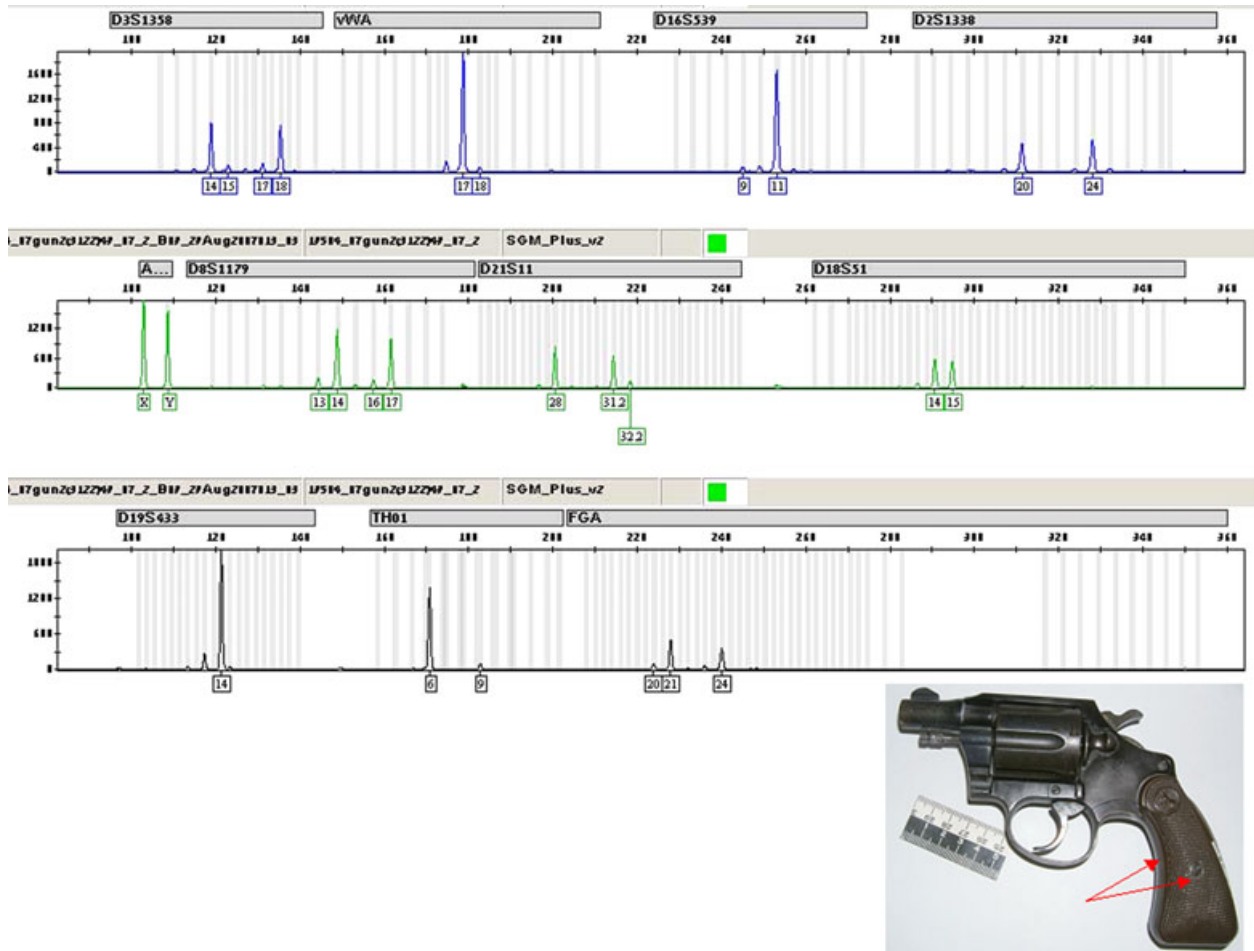- The base of the thumb.
- The tip of the forefinger.

FIG. 3—*Electropherogram of the profile obtained from the tape sampling of the gun handle areas indicated by the arrows.*

The following results were obtained: Taping 1 contained a mixture of at least two individuals with a major profile of male origin designated as male 1 (Fig. 5). Tapings 2 and 3 contained the single male profile differing from male 1 and designated as male 2 (Fig. 6, taping 2).

Reference samples from two suspects were submitted for comparison. The profile designated as male 2 (from tapings 2 and 3) matched the profile of one of the suspects.

Discussion

The three-layer adhesive tape used in our work was originally employed in the Fibers and Polymers Laboratory of the DIFS (Division of Identification and Forensic Science of the Israel Police). That laboratory used A4-sized sheets for the collection of trace evidence, such as hairs, fibers, etc., from crime scenes and crime scene items.

By using the same tape in both the Fibers and Polymer Laboratory and the DNA Laboratory, biological material found on tapes sampled by the Fibers and Polymer Laboratory can be transferred *in situ* to the Biology Laboratory for direct DNA extraction, thereby minimizing contamination.

We have been using these adhesive tapes routinely for the collection of biological material from crime items since 2004. In that period, we have sampled several hundred items. The technique has

been adapted successfully to a wide range of objects, such as clothing, ropes, guns, knife handles, tools, fingernails (usually of victims), etc., recovered from crime scenes.

Furthermore, we have found that the technique is applicable to the objects of a cultural habit indigenous to Israel. Eating sunflower seeds, watermelon seeds, etc., and expectorating the shells is common practice here, and these items are therefore not infrequently encountered at crime scenes. In the past, we were seldom able to obtain a DNA profile by swabbing of the shells or by direct DNA extraction. However, by taping the inside and outside of seed shells, our success rate in profiling these items has markedly increased due possibly to a reduced amount of inhibitors "co-sampled" with DNA.

Using adhesive tapes, we are able to selectively sample specific areas of casework items. Gloves, for example, can be taped on the inside and outside *separately*—thus avoiding surface-to-surface contamination and consequently reducing the chance of obtaining mixtures with possibly irrelevant DNA sources.

We normally sample casework items in two or three areas of presumed highest friction between an individual(s) and the crime scene item. This "screening" may be sufficient to obtain a comparable DNA profile. In some cases, however, where complex mixtures or indication of minute amounts of DNA are encountered, the resampling of areas *adjacent* to the initial taping often yields more informative profiles.
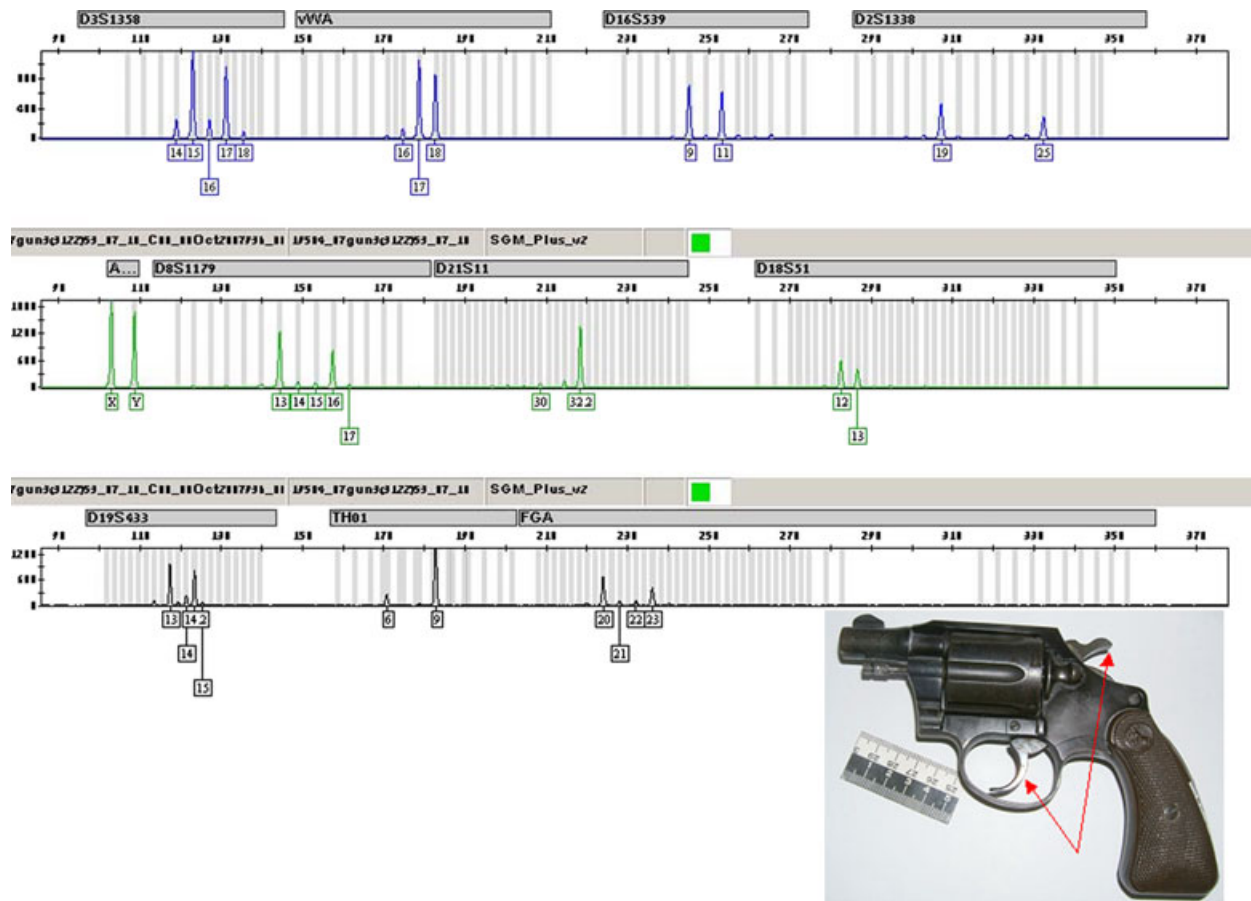
FIG. 4—*Electropherogram of the profile obtained from the tape sampling of the gun trigger and hammer indicated by the arrows.*

In our experience, the preferred areas of sampling are as follows:

- *For robbery or ski masks:* around the inside corners of the mouth hole, the presumed area of the nostrils and below the eyes holes (the eyes shed a large number epithelial cells [3]);
- *For gloves:* the inside tips and the base of the fingers, and at the base of the palm;
- *For socks that have been worn or used as gloves:* the inside heel and toes areas;
- *For socks that have been used to conceal a weapon:* on the upper inner and outer part of the sock—particularly, the seams area (based on the assumption that while handling the sock, both sides were exposed to biological material).
- *For guns:* on the slide (for automatics), the trigger and the trigger guard, the hammer and the handle; *for Fabrique National (FNs):* in particular, the hammer—the arch housing is relatively short in this model and while firing the hammer falls backwards with force onto the skin between the forefinger and the thumb.
- *For tools:* on their handles;
- *For garments that have been worn:* on collars, seams, and areas of stitching (as around pockets or pocket linings).

The conventional cutting method to sample biological material from crime scene items can result in the co-extraction of PCR inhibitors with DNA. On the other hand, sampling with adhesive tape largely avoids the collection of inhibitors, such as dyes. We have successfully amplified DNA from taping samples of biological material on denim, leather, and soiled items—all of which are known to contain PCR inhibitors (10–12).

By extracting DNA from single tape samples in small volumes (200 µL), the difficulty of dealing with dense items (coats, blankets, etc.) is overcome. In the past, areas of these items were cut and placed in one or more tubes for extraction. The need to pool and concentrate these samples increased the chance of both DNA loss and contamination.

In addition to the Chelex method of DNA extraction, we have found that profiles can be obtained from crime scene tape samples by the organic/preferential extraction method in sexual assault cases.

Formerly, samples that did not yield a positive result with presumptive tests for blood, saliva, or sperm were rarely sampled for DNA. However, presumptive tests may give an equivocal result or no result and are not indicative of concealed DNA sources. Therefore, in many cases, we sample relevant areas of crime scene items (ski masks, for example) by tapings as an initial step to DNA extraction without carrying out a presumptive test. These samples are then analyzed routinely using the SGM Plus™ kit and, in rare cases, where a partial profile is obtained, analysis is extended using the Minifiler kit (Applied Biosystems).

In those cases where it may be necessary to carry out a presumptive test (Phadebas, for instance), we first tape an area, to avoid the loss of cellular material and only subsequently carry out the presumptive test. This loss of cellular material has been observed in the acid phosphatase presumptive test in which sperm cells have been shown to diffuse from casework items to filter papers (13,14).

In conclusion, even in the absence of commonly found sources of DNA (blood, semen, saliva, vaginal fluid, etc.), profiles can be
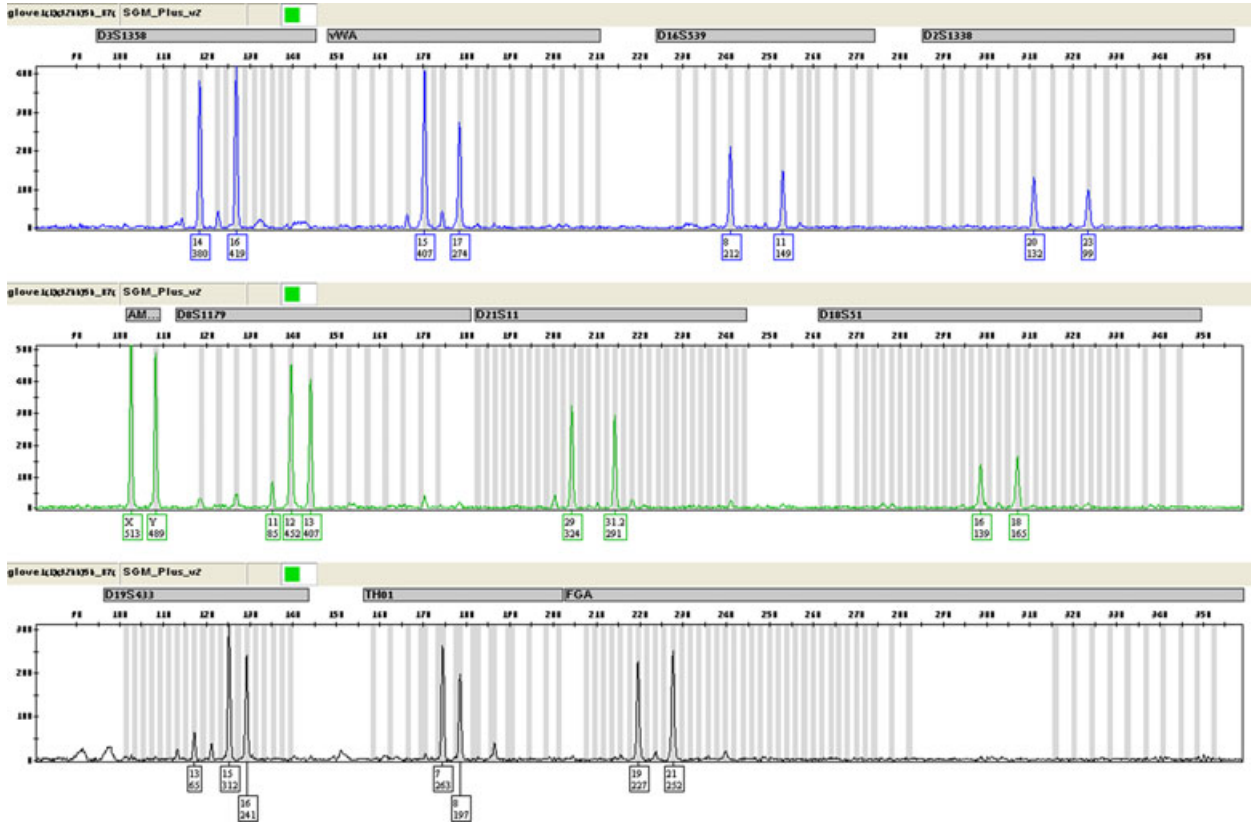
FIG. 5—Electropherogram of the profile obtained from the tape sampling of the inside tip of the glove thumb (the glove was not available for photography).
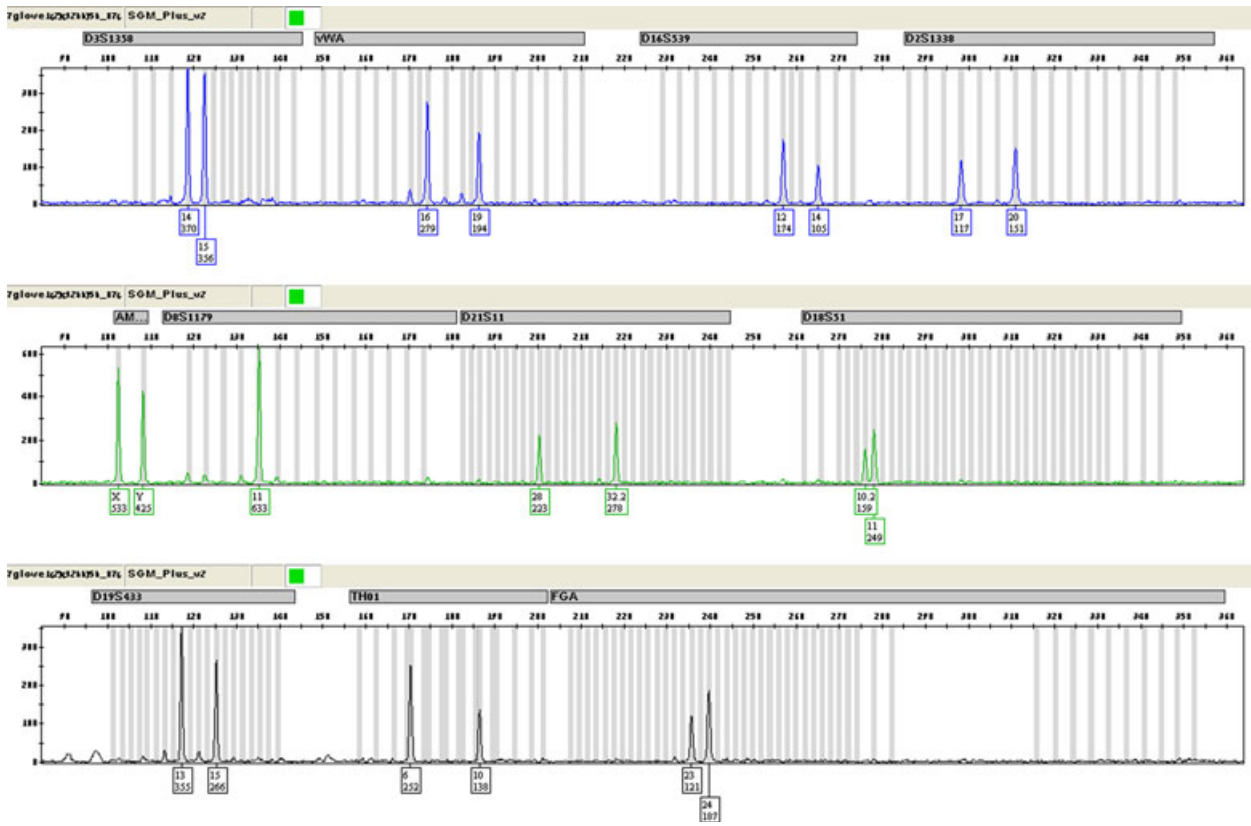

FIG. 6—Electropherogram of the profile obtained from the tape sampling of the inside base of the glove thumb (the glove was not available for photography).

obtained from shed epithelial cells. Humans shed approximately 400,000 skin cells daily from all parts of the body (3). By using adhesive tape, this rich source of untapped DNA can be easily collected for profiling.

We have found that tape collection of biological material from casework items has not only simplified sampling and is non-destructive, but also has a high success rate in genotyping DNA from many previously untested items left at scenes of crimes.

References

1. Rutty GN. An investigation into the transference and survivability of human DNA following simulated manual strangulation with consideration of the problem of third party contamination. Int J Legal Med 2002;116:170–3.
2. Albaz J, Walsh SJ, Curran JM, Moss DS, Cullen J, Bright J-A, et al. Comparison of variables affecting the recovery of DNA from common drinking containers. Forensic Sci Int 2002;126:233–40.
3. Wickenheiser RA. Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact. J Forensic Sci 2002;47:442–50.
4. Bright J-A, Petricevic SF. Recovery of trace DNA and its application to DNA profiling of shoe insoles. Forensic Sci Int 2004;145:7–12.
5. Hall D, Fairley M. A single approach to the recovery of DNA and firearm discharge residue evidence. Sci Justice 2004;44:15–9.
6. Zamir A, Oz C, Wolf E, Vinokurov A, Glattstein B. A possible source of reference DNA from archived treated adhesive lifters. J Forensic Sci 2004;49:68–70.
7. Walsh PS, Metzger DA, Higuchi R. Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 1991;10:506–13.
8. Quantifiler™ kit user's manual. The Perkin-Elmer Corporation. Foster City, CA: Applied Biosystems, 2003.
9. AmpFlSTR® SGM Plus™ kit user's manual. The Perkin-Elmer Corporation. Foster City, CA: Applied Biosystems, 2003.
10. Shutler GG, Gagnon P, Verret G, Kalyn H, Korkosh S, Johnson E, et al. Removal of PCR inhibitor and resolution of DNA STR types in mixed human-canine stains from a five year old case. J Forensic Sci 1999;44:623–6.
11. Montpetit SA, Fitch IT, O'Donnell PT. A simple automated instrument for DNA extraction in forensic casework. J Forensic Sci 2005;50:555–63.
12. Watson RJ, Blackwell B. Purification and characterization of a common soil component which inhibits the polymerase chain reaction. Can J Microbiol 2000;46:633–42.
13. Reshef A, Barash M, Gallili N, Michael A, Brauner P. The use of acid phosphatase papers in DNA profiling. Sci Justice 2005;45:97–102.
14. Brauner P, Barash M, Reshef A, Michael A. Extraction of DNA from 8-year-old acid phosphatase test papers in a gang rape case. J Forensic Ident 2008;58:671–81.

Additional information and reprint requests:
Mark Barash, M.Sc.
Faculty of Health, Science and Medicine
Bond University
Gold Coast, QLD 4229
Australia
E-mail: mbarash@bond.edu.au

TECHNICAL NOTE
CRIMINALISTICS

Bhawna Dubey,[1] *M.Sc.; P. R. Meganathan,*[1] *M.Sc.; and Ikramul Haque,*[1] *Ph.D.*

Molecular Identification of Three Indian Snake Species Using Simple PCR–RFLP Method*

ABSTRACT: Three endangered Indian snake species, *Python molurus, Naja naja*, and *Xenochrophis piscator* are known to be significantly involved in illegal trade. Effective authentication of species is required to curb this illegal trade. In the absence of morphological features, molecular identification techniques hold promise to address the issue of species identification. We present an effective PCR–restriction fragment length polymorphism method for easy identification of the three endangered snake species, *Python molurus, Naja naja*, and *Xenochrophis piscator*. A 431-bp amplicon from cytochrome *b* gene was amplified using novel snake-specific primers following restriction digestion with enzymes *Mbo* II and *Fok* I. The species-specific reference fragment patterns were obtained for the target species, which enabled successful identification of even highly degraded shed skin sample confirming the utility of the technique in case of poor-quality DNA. The assay could be effectively used for forensic authentication of three Indian snake species and would help strengthen conservation efforts.

KEYWORDS: forensic science, polymerase chain reaction, restriction fragment length polymorphism, cytochrome *b,* snake species, molecular identification

Illegal wildlife trade in India has led to the dwindling populations of many snake species. Snakes are used in the leather industry, for traditional medicines, and in the pet market. India has a long history of trade in reptiles; for instance, in 1977, a staggering 4 million snakeskins were exported (1). A significant illegal trade of the three most endangered snake species, Indian Rock Python (*Python molurus*), Indian Cobra (*Naja naja*), and Checkered Keelback (*Xenochrophis piscator*) has exacerbated their predicament and deserves a priority for conservation efforts (2). In view of this, the python (*Python molurus*) and the cobra (*Naja naja*) are listed in Appendix I/II of Convention on International Trades in Endangered Species. Also, the Indian Rock Python is placed in Schedule I, and the Indian Cobra and Checkered Keelback are placed in Schedule II of the Indian Wildlife Protection Act, 1972. Although reptile trade is banned in India, illegal trade still continues covertly. To check illegal trade and to conserve these threatened species, effective species authentication is necessary. Species identification based on morphological characteristics does not prove useful in processed skin/meat products seized as illegal harvests, whereas molecular techniques serve as potential tools to assist the forensic identification. Mitochondrial DNA (mt-DNA) has been utilized as an effective tool for forensic analysis because of their several thousand copy numbers in each cell, a feature of paramount importance while dealing with degraded samples (3). Also previous studies have successfully utilized cytochrome *b* (cyt *b*) gene sequences to differentiate between snake species (4).

Species recognition by DNA analysis can be carried out by methods based on PCR, sequencing, and restriction enzyme analysis (5). The high cost and arduous nature of DNA sequencing limits its use in small laboratories with restricted facilities. Therefore, this study is aimed at developing a simple and effective molecular technique for identification of three endangered snake species of India. We have chosen a restriction fragment length polymorphism (RFLP) technique, because (i) it is less time–consuming, (ii) it is more cost-effective than DNA sequencing, (iii) it requires equipment readily available in most molecular laboratories, and (iv) it has proved its utility in species identification (6–10).

In this study, for easy and effective identification of the three Indian species, we have used novel primers to amplify the cyt *b* gene to generate species-specific PCR–RFLP patterns. Results presented here suggest that this technique may be effective for species-specific identification of three endangered species of Indian snakes. The technique has proved effective even in the case of degraded DNA recovered from a shed skin sample.

Materials and Methods

Sample Collection and DNA Extraction

The tissue samples from eleven Checkered Keelback (*Xenochrophis piscator*), six Indian Rock Python (*Python molurus molurus*), and eight Indian Cobra (*Naja naja naja*), along with some common varieties of Indian snakes (*Ptyas mucosus, Dabioa russelli, Bangarus fasciatus, Eryx johnii, Eryx conicus, Nartix stolata,* and *Ahaetulla nasuta*), were obtained from the Snake Transit House, Jabalpur, Madhya Pradesh. Shed skin samples of Indian Rock Python (*Python molurus molurus*) were obtained from the Calcutta Snake Park, Kolkata, West Bengal. The DNA samples of *Varanus bengalensis, Crocodylus porosus, Gavialis gangeticus,* and *Kachuga tecta* were obtained from the repository of the Central Forensic Science Laboratory, Kolkata, to be included as negative controls. Genomic DNA from tissue samples was isolated using QIAmp tissue DNA extraction kit (Qiagen GmBH, Germany), following

[1]National DNA Analysis Centre, Central Forensic Science Laboratory, 30- Gorachand Road, Kolkata 700 014, West Bengal, India.

the manufacturer's protocol. DNA from shed skin was isolated as per the protocol described elsewhere (11).

PCR and DNA Sequencing

Primer set, RC3F: 5′ TGA GGA CAA ATA TCA TTC TGA G 3′ and RC3R: 5′ TAG GCG AAT AGG AAG TAT CA 3′, was designed based on cyt *b* sequences of snake species available from public databases. The PCR was performed in a 25-μL reaction volume containing 10 ng of DNA template, 5 mM MgCl$_2$ (Invitrogen Life Technologies, Sao Paulo, Brazil), 200 μM dNTPs (MBI Frementas, Glen Burnie, MD), 0.2 μM primers each, 2.5 μL of 10× buffer (Invitrogen Life Technologies, Brazil), and 5 U of Taq polymerase (Invitrogen Life Technologies) in a GeneAmp® 9700 PCR system (Applied Biosystems, Foster City, CA). The PCR conditions were as follows: initial denaturation at 94°C for 4 min followed by 35 cycles of denaturation (94°C for 30 sec), annealing (52°C for 30 sec), and extension (72°C for 30 sec). The reaction terminated with a final extension of 72°C for 5 min followed by 4°C hold. The amplified fragments were detected on 2% agarose gel containing 0.5 μg/mL ethidium bromide stain.

The amplified products were cycle-sequenced using Big Dye Terminator v.3.1 Cycle sequencing kit (Applied Biosystems). Sequencing was performed on 3100 Avant Genetic Analyzer (Applied Biosystems), and the obtained sequences were aligned using Clustal W (12).

Restriction Fragment Length Polymorphism Analysis

Available cyt *b* gene sequences for snake species from public databases along with the sequences obtained in this study were subjected to electronic RFLP using Webcutter 2.0 software (13). Based on these preliminary results, two enzymes (*Mbo* II and *Fok* I) were selected for further experimental analysis. Eight to nine microliters of PCR product (approximately 1 μg) were digested with 1 U of *Mbo* II and *Fok* I (New England Biolabs, Beverley, MA) individually, in a final volume of 25 μL for 4 h at 37°C. The resulting fragments were detected on 3% agarose gel containing 0.5 μg/mL ethidium bromide stain.

Results and Discussion

Novel primers were used for amplifying the partial cyt *b* gene sequence to circumvent the difficulty of obtaining whole sequence in the case of degraded forensic samples. The cyt *b* gene amplification of the three snake species resulted in a fragment of 431 bp. However, DNA from other reptiles like *Varanus bengalensis, Crocodylus porosus, Gavialis gangeticus,* and *Kachuga tecta* did not result in amplification. The sequences of the three snake species were aligned and showed the presence of 167 variable sites, 266 conserved sites, and 165 singleton sites (those in which only a single sequence contains the variant nucleotide) in the cyt *b* fragment. The partial cyt *b* gene sequences obtained in this study as well as from all available snake species in public databases were assembled and subjected to electronic RFLP (data not shown). Following the electronic RFLP, two enzymes *Mbo* II and *Fok* I were then carefully selected to obtain unambiguous differentiation among the three snake species (Table 1). The use of both the enzymes separately allowed the discrimination between target species on an agarose gel (Fig. 1). Electrophoretic analysis of *Fok* I- and *Mbo* II-digested amplicons resulted in the expected fragments as shown in Table 1 except that fragments smaller than 50 bp could not be seen in the gel. The RFLP patterns obtained

TABLE 1—Predicted size of DNA segments (bp) based on electronic restriction fragment length polymorphism analysis with Webcutter 2.0 program.

Species	Fok I	Mbo II
Python molurus	281, 150	209, 128, 82, 12
Naja naja	431	365, 66
Xenochrophis piscator	265, 144, 22	250, 150, 31

were monomorphic for each of the three species studied (data not shown). The technique was tested on certain other common varieties of snakes found in India (*Ptyas mucosus, Dabioa russelli, Bangarus fasciatus, Eryx johnii, Eryx conicus, Nartix stolata,* and *Ahaetulla nasuta*), which yielded different restriction digestion patterns (data not shown) from the three snake species of concern. To test the efficacy of our method in cases of degraded samples, we used a highly degraded shed skin sample of *Python molurus molurus* (*P. molurus*) yielding very low-quality DNA. Complete concurrence was obtained in the profiles generated from reference samples and the tested shed skin sample (Fig. 2).

The over exploitation of snakes has drawn alarms from conservation agencies. To protect these threatened species, effective identification is indispensable to monitor illegal trade. In recent years, various molecular approaches like mitochondrial DNA sequencing (14,15), AFLPs (16), arbitrarily primed PCR (AP-PCR), and RAPD techniques have been introduced to identify snake species (17,18). Problems with DNA sequencing techniques (19) reflect the vital need for augmenting an easy, rapid, and cost-effective method, which could enable the authentication of snake products. This study provides a simple and effective PCR–RFLP technique for rapid identification of three endangered snake species of India, *Python molurus molurus* (*P. molurus*), *Naja naja naja* (*N. naja*), and *Xenochrophis piscator* (*X. piscator*).

The method is intended to be performed with minimal laboratory equipments, cost, and time. The results support the worthiness of the method in cases of degraded noninvasive (shed skin) samples. Hence, we recommend the use of this PCR–RFLP technique for quick identification of the three endangered snake species discussed
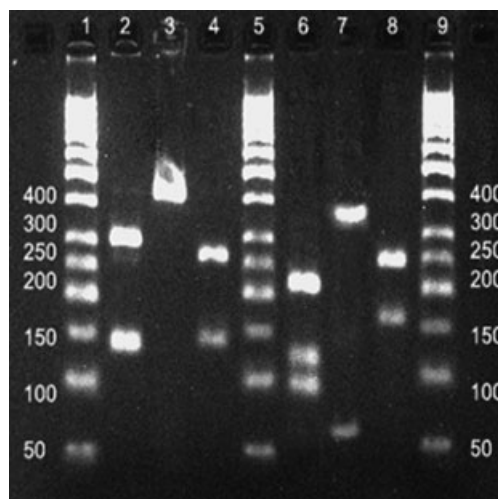

FIG. 1—PCR-restriction fragment length polymorphism patterns for cyt b gene obtained by digestion with enzymes: Fok I—Lanes 2–4; Mbo II—Lanes 6–8; Lanes 2, 6: P. molurus; Lanes 3, 7: N. naja; Lanes 4, 8: X. piscator. Lanes 1, 5, 9—Molecular size marker. (Fragments less than 50 bp are not visible.)
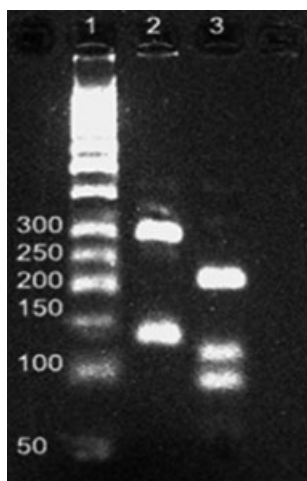
FIG. 2—*PCR-restriction fragment length polymorphism patterns produced in shed skin sample of* P. molurus *using enzymes: Lane 2—Fok I; Lane 3— Mbo II; Lane 1—Molecular size marker.*

earlier. It is expected not only to greatly improve species identification precision but strengthen the law enforcement agencies for an effective check on illegal trade and poaching and thereby will prove useful in the conservation efforts to save these species from extinction.

References

1. http://www.wpsi-india.org/news/27022004.php (assessed November 12, 2008).
2. Whitaker R. Common Indian snakes: a field guide. New Delhi: Macmillan India, 2006.
3. Bravi CM, Lirón JP, Mirol PM, Ripoli MV, Peral-García P, Giovambattista G. A simple method for domestic animal identification in Argentina using PCR-RFLP analysis of cytochrome b gene. Leg Med 2004;6:246–51.
4. Wong K-L, But P, Shaw P-C. Application of cytochrome b DNA sequences for the authentication of endangered snake species. Forensic Sci Int 2004;39:49–55.
5. Teletchea F, Maudet C, Hanni C. Food and forensic molecular identification: update and challenges. Trends Biotechnol 2005;23:359–66.
6. Palo JU, Merila J. A simple RFLP method for identification of two ranid frogs. Conserv Genet 2003;4:801–3.
7. Fajardo V, Gonzalez I, Lopez-Calleja I, Martin I, Hernandez PE, Garcia T, et al. PCR-RFLP authentication of meats from red deer (*Cervus elaphus*), fallow deer (*Dama dama*), roe deer (*Capreolus capreolus*), cattle (*Bos taurus*), sheep (*Ovis aries*) and goat (*Capra hircus*). J Agric Food Chem 2006;54:1144–50.
8. Cossios D, Angers B. Identification of Andean felid feces using PCR-RFLP. Mastozool. Neotrop 2006;3:239–44.
9. Fagundes V, Nogueira CDA. The use of PCR-RFLP as an identification tool for three closely related species of rodents of genus Akodon (Sigmodontinae, Akodontini). Genet Mol Biol 2007;30:698–701.
10. Rohilla MS, Tiwari PK. Restriction fragment length polymorphism of mitochondrial DNA and phylogenetic relationships among five species of Indian fresh water turtles. J Appl Genet 2008;49:167–82.
11. Fetzner JW. Extracting high quality DNA from shed reptile skin: a simplified method. BioTechniques 1999;26:1052–4.
12. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80.
13. http://www.firstmarket.com/cutter/cut2.html (assessed November 24, 2008).
14. McEwing R, Pook CE. Mitochondrial DNA sequences from dried snake venom: a DNA barcoding approach to the identification of venom samples. Toxicon 2005;46:711–15.
15. Wuster W, Thorpe RS. Naja siamensis, a cryptic species of venomous snake revealed by mitochondrial sequencing. Experientia 1994;50:75–9.
16. Creer S, Thorpe RS, Malhotra A, Chou W-H, Stenson AG. The utility of AFLPs for supporting mitochondrial DNA phylogeographical analyses in Taiwanese bamboo viper, *Trimeresurus stejnegeri*. J Evol Biol 2004;17:100–7.
17. Wang YQ, Zhou KY. Study on randomly amplified polymorphic DNA of 10 species of snakes in Colubridae. Chin J Appl Env Biol 1996;2:273–9.
18. Wang YQ, Zhou KY. A preliminary study on the identification of crude snake drugs by molecular genetic markers. Acta Pharm Sinica 1997;32:384–7.
19. Purcell M, Huber H, Park L. Molecular methods for genetic identification of salmonid prey from Pacific harbor seal (*Phoca vitulina richardsi*) scat. Fish Bull 2004;102:213–20.

Additional information and reprint requests:
Ikramul Haque, Ph.D.
National DNA Analysis Centre
Central Forensic Science Laboratory
30 Gorachand Road
Kolkata 700 014
West Bengal
India
E-mail: haque_cfslk@yahoo.co.in

TECHNICAL NOTE
CRIMINALISTICS

Robert G. Weston,[1] *M.Sc.*

Quick Screening of Crystal Methamphetamine∕Methyl Sulfone Exhibits by Raman Spectroscopy

ABSTRACT: The analysis of mixtures of "crystal meth" (usually comprised of methyl sulfone [MS] and methamphetamine [MA]) by gas chromatography-mass spectrometry (GCMS) is routine in many forensic drug laboratories. The utilization of Raman spectroscopy for the identification of such mixtures quickly and without the need for a separation technique is discussed. Samples were dissolved in water and Raman spectra of the resulting aqueous solutions were collected. By comparing these spectra to spectra of methylsulfone and MA mixtures of known composition, an indication of the composition of the sample can be obtained in only a few minutes. This spectral comparison also can be used as a semi-quantitative analysis of MA concentrations in such exhibits.

KEYWORDS: forensic science, controlled substances, crystal meth, methamphetamine, methyl sulfone, Raman spectroscopy, drug mixtures, isopropylbenzylamine

The use of Raman spectroscopy as a drug analysis technique has been reported often in the past decade. Numerous articles have been published during that time describing the use of Raman for the analysis of cocaine (1–5), benzodiazepines (6), amphetamines (7–9), and designer drugs such as ecstasy tablets (10,11). The current work describes the use of Raman spectroscopy for the qualitative and semi-quantitative analysis of crystal meth samples, with the only sample prep being the dissolution of the samples in deionized water (7,8). After dissolution, Raman spectra were then collected and compared to the spectra of mixtures of methamphetamine (MA) and methyl sulfone (MS), to allow for identification and approximate quantification of the samples. Recently, instances of *N*-isopropylbenzylamine hydrochloride (IBA) being present in ecstasy and crystal meth mimics have been reported (12–14). This technical note will also address the potential of analyzing crystal meth samples that may contain IBA.

Methods

Standards of MA hydrochloride, isopropylbenzylamine, and MS were purchased from Fisher Scientific (Hampton, NH). Isopropylbenzylamine HCl was manufactured in-house by passing hydrogen chloride gas through the free base. Mixtures of MA and MS were prepared by dissolving each standard in deionized water (water does not have an appreciable Raman spectrum in the wavenumber range utilized in this work [7]), and then mixing appropriate volumes of each standard. These mixtures were prepared in 10% increments (e.g. 90% MA∕10% MS, 80% MA∕20% MS, etc. where the percentages indicate the relative weight proportion of

each solute) with a total solute concentration of 50 mg∕mL. This concentration of solute is much greater than concentrations typically used in gas chromatography-mass spectrometry (GCMS) or high-performance liquid chromatography (HPLC) analysis because of the relatively weak Raman scattering effect. This concentration allows for easy detection of the spectral peaks of both MS and MA. Raman spectra of the pure MS and MA standards, as well as spectra of the mixtures of those standards, were obtained and entered into a spectral library for later searching and spectral analysis. For each standard, the integration time was 15 sec with five scans being averaged to compile the final spectrum. The spectral software utilized was OMNIC driver version 7.3, firmware version 2.08. Raman spectra were acquired with a DeltaNu (Laramie, WY) Advantage NIR spectrometer. The source wavelength is 785 nm, and the spectral range is 100–2000∕cm.

The Raman spectrometer was calibrated each day using a polystyrene standard and the calibration function of the software program. No further optimization or adjustment of the instrumental parameters was performed.

Forensic exhibits were obtained from routine casework items submitted to Oklahoma State Bureau of Investigation for analysis. The only criteria for the items selected for this study were that the items had to have the macroscopic physical appearance of a crystalline substance. The exhibits were numbered as they were collected, and are named throughout this work as FE1, FE2, etc. (for Forensic Exhibit 1, Forensic Exhibit 2). One hundred forensic exhibits were analyzed for this study. Samples of all forensic exhibits were dissolved in deionized water to a concentration of 50 mg∕mL and then analyzed by Raman spectroscopy.

All forensic exhibits were also analyzed by GCMS and HPLC. For GCMS analysis, a sample of each forensic exhibit was dissolved in methanol to an approximate 1–2 mg∕mL concentration. This analysis was used to determine if compounds other than MA and MS were present. If a third component was a part of the FE,

[1]Criminalist, Oklahoma State Bureau of Investigation, 800 E. 2nd Street, Edmond, OK 73034.

the Raman spectra were evaluated to see if there was a spectral difference allowing for the detection of that third component. HPLC was used to quantify the percentage MA in the forensic exhibits so that the merits of Raman as a semi-quantitative tool in crystal meth analysis could be evaluated. Samples for HPLC analysis were dissolved in water to a concentration of 1 mg/mL and quantitated using an ultraviolet diode array with 210 nm as the detection wavelength.

Results and Discussion

Qualitative Analysis

In the following discussion, a strong peak is described as a peak that has a peak height of >50% the height of the most intense peak in the spectrum. Moderate peaks are *c.* 25–50% the height of the most intense peak, and weak peaks are those with peak heights <25% of the most intense spectral peak.
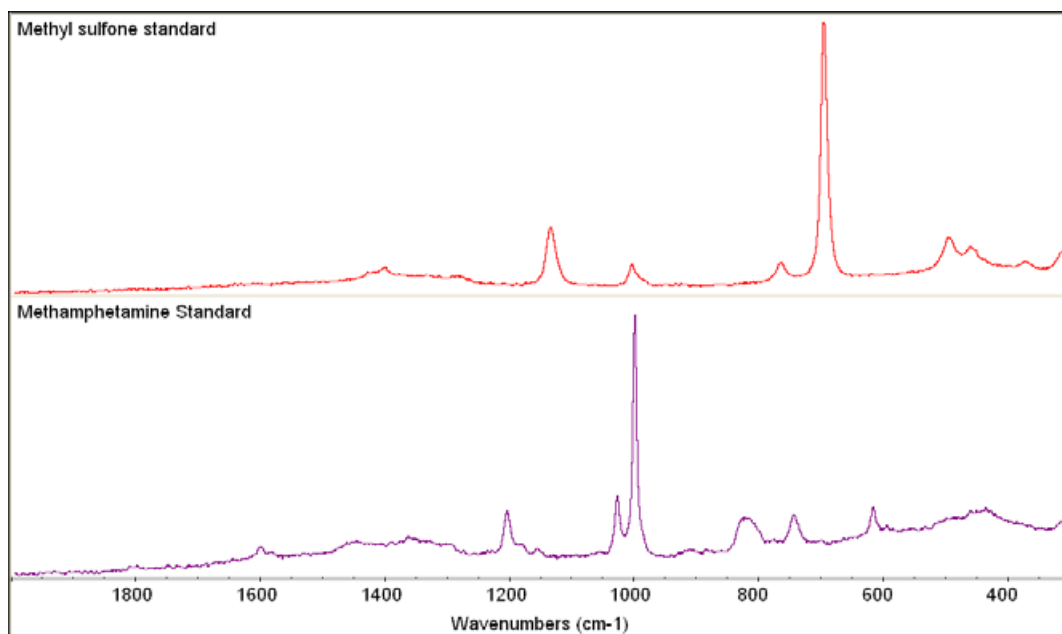

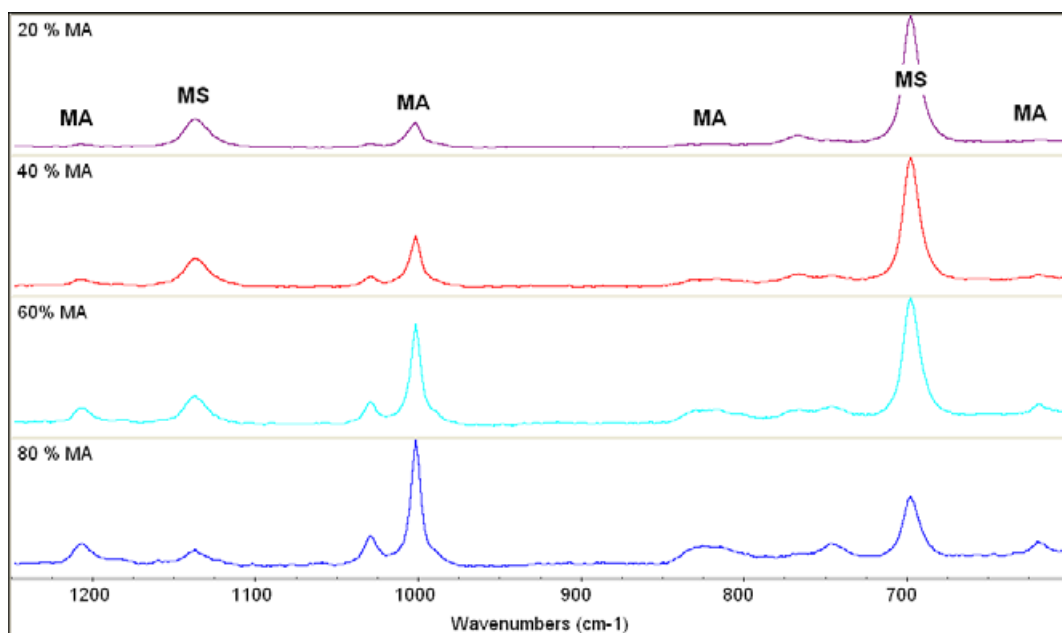FIG. 1—*Raman spectra of methyl sulfone (top) and methamphetamine (bottom).*

FIG. 2—*Mixtures of methamphetamine (MA) and methyl sulfone (MS). From top to bottom the ratios of MA/MS are 20/80, 40/60, 60/40 and 80/20. The compound responsible for each of the major spectral peaks is also indicated above the peak location in the top spectrum.*
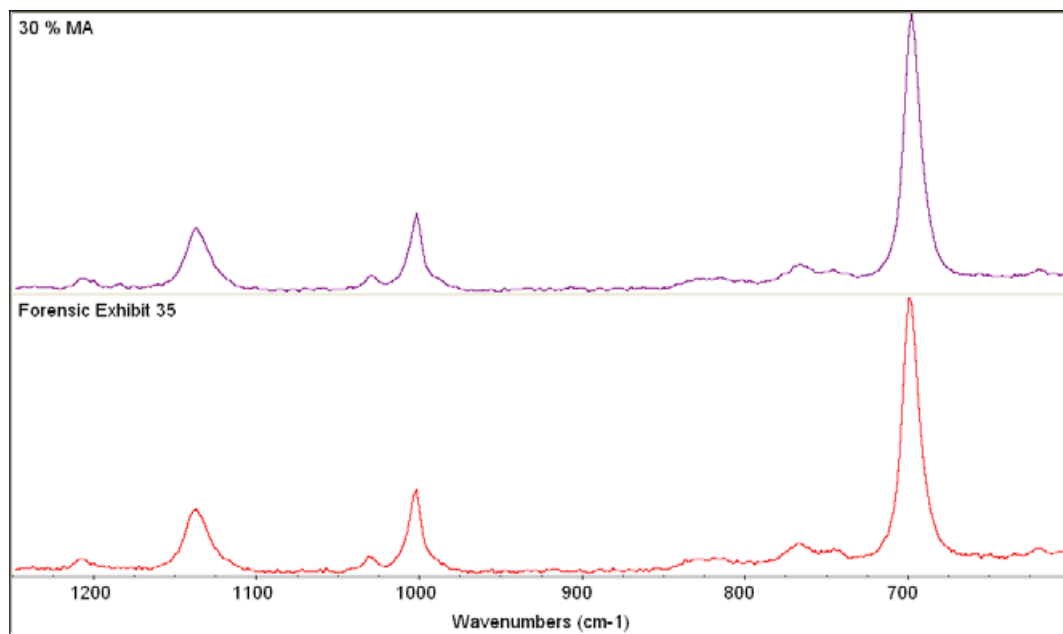
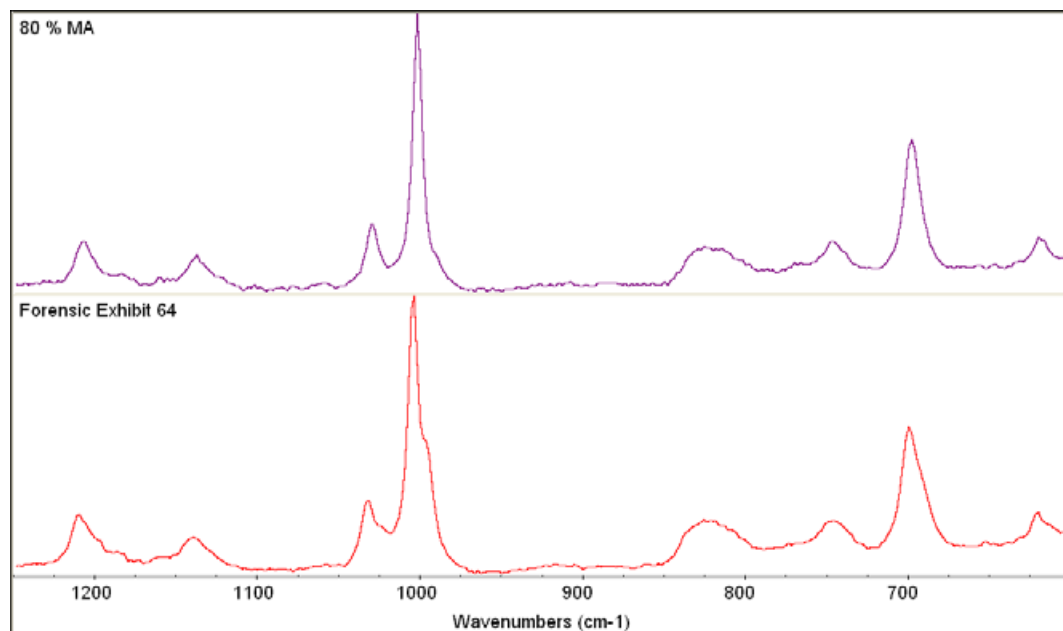FIG. 3—*Spectra of 30% methamphetamine standard (top) and FE35.*

FIG. 4—*Spectra of 80% methamphetamine standard (top) and FE64.*

Analysis of the Raman spectra of aqueous MS and MA reveals that each compound gives a relatively simple Raman spectrum (see Fig. 1). The most intense peak of MS is found at 703/cm, a moderate peak at 1136/cm, and weak but discernable Raman bands at 1407, 1008, 768, 499, and 315/cm. MA has its most intense Raman band at 1004/cm (attributed to the monosubstituted benzene ring), a moderate peak at 1032/cm, and weak bands at 1209, 831 (broad), 748, and 621/cm. The weak 1008/cm band of MS does coincide with the 1004/cm band of MA, but the determination of the presence of MA in crystal meth is not solely dependent upon this peak. As all discernable MA peaks are in the 1200–600/cm range, the remainder of this work focuses on that region of the spectrum.

A requirement for utilizing Raman spectroscopy for substance identification is that the spectra of the drug of interest and the diluents must not have significant peak co-incidence. Raman spectra of the mixtures of standards are presented in Fig. 2, with an indication of which compound is responsible for the peaks found in each spectral region. By analyzing Raman spectra of standards and of mixtures of standards, it has been determined that there are sufficient differences in the spectra which allow for the identification of both MA and MS in mixtures of these substances, provided the concentration of MA is sufficient (see Detection Limit).

All forensic exhibit spectra were evaluated in the 1200–600/cm region by comparison with the spectrum of MA, MS,
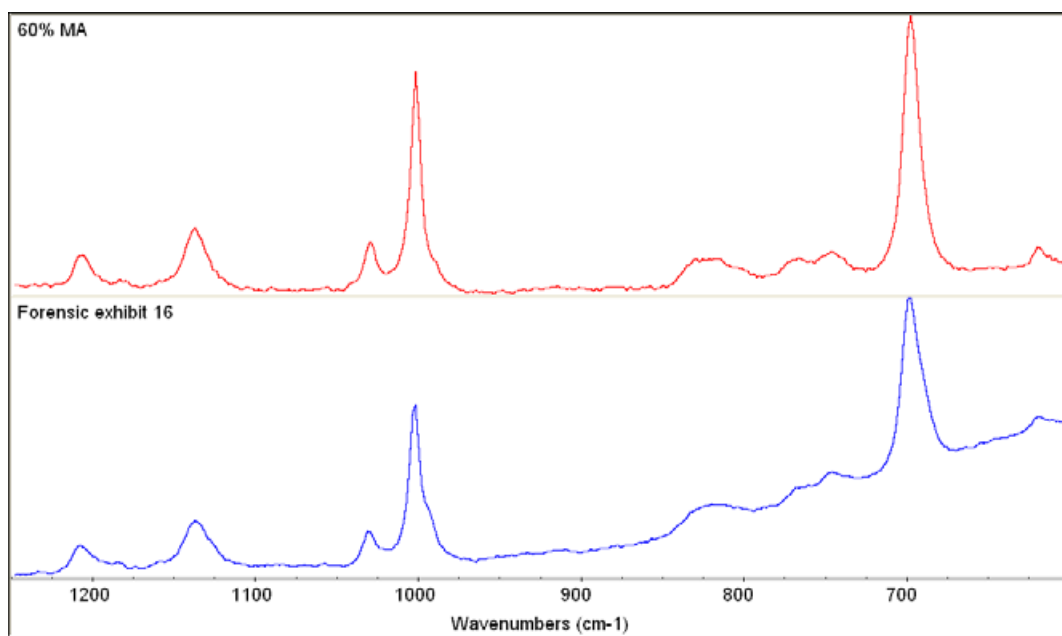
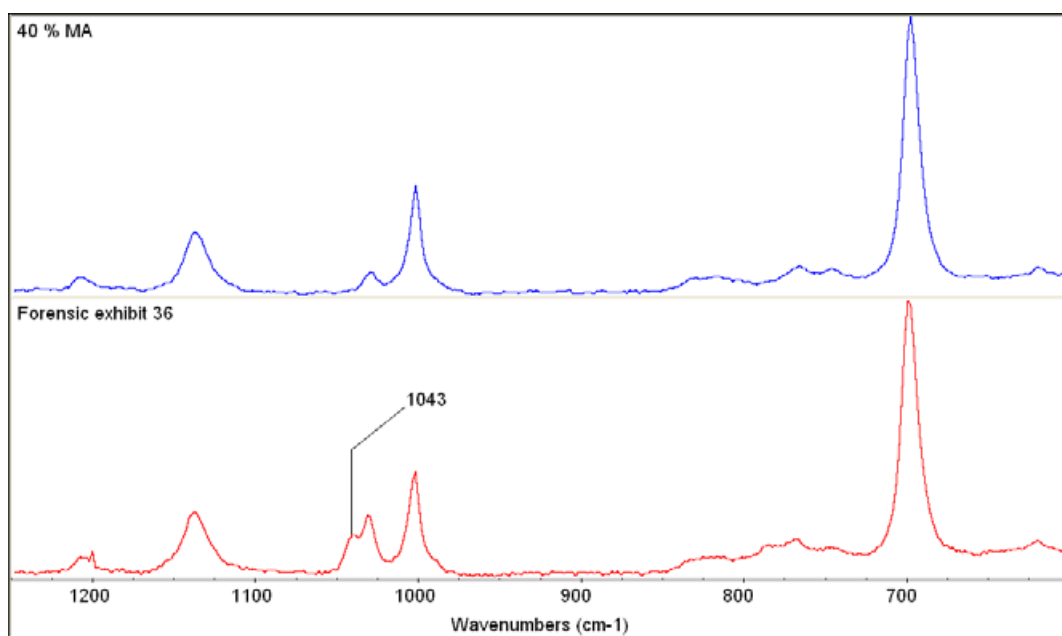FIG. 5—*Spectra of 60% methamphetamine standard (top) and FE16.*

FIG. 6—*Spectra of 40% methamphetamine standard (top) and FE36. Note the peak at 1043/cm in the spectrum of FE36 not present in the standard.*

and MA/MS mixed standards. The spectra of FE35 and FE64 are presented in Figs. 3 and 4, respectively, as examples of the high degree of spectral similarity between the forensic exhibits and a MS/MA mixed standard. This evaluation showed no remarkable differences between the Raman spectra of forensic exhibits and a standard other than in FE16, FE36, FE81, and FE84. The Raman spectrum of FE16 (Fig. 5) did not appear to have any significant differences from the standard other than a strong fluorescence signal. The fluorescence is the increased baseline observed in the lower wavenumber region of the spectrum. Many organic compounds with extended conjugation (greater than a benzene ring) are known to have a strong fluorescence signal in Raman spectroscopy. This fluorescence may be from a compound present in the mixture in a very low concentration, as no compounds other than MS and MA were detected by GCMS. For FE36 (Fig. 6), FE81, and FE84 (not shown), there is one Raman peak (1043/cm) that does not appear in the spectra of the standards, indicating the presence of a third substance. In instances such as this, further instrumental analysis should be performed to determine the identity of the third substance. For these particular items, GCMS analysis revealed this third component as nicotinamide.
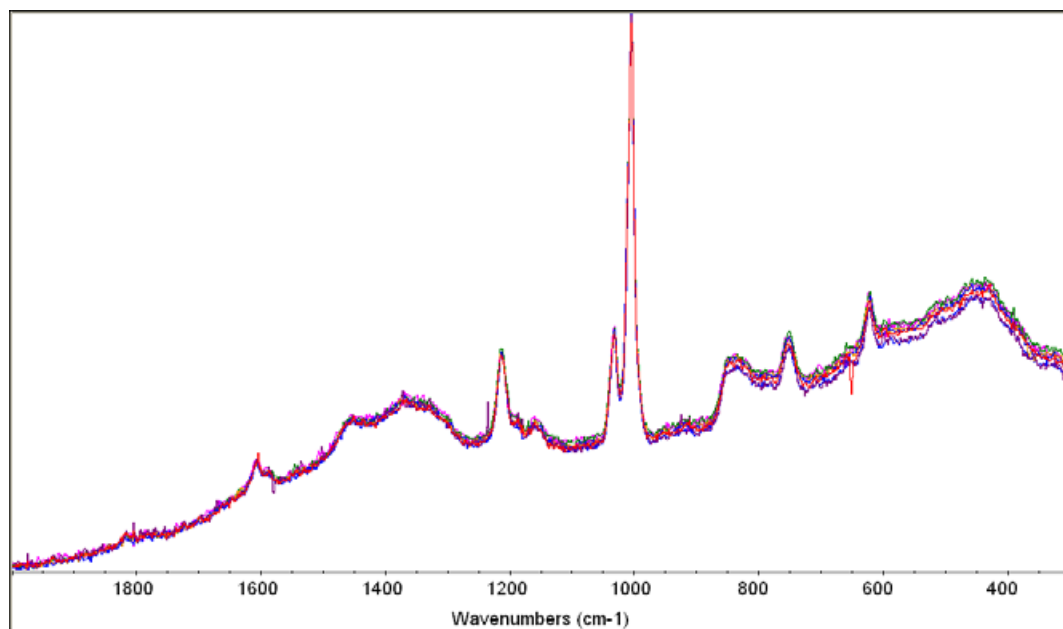
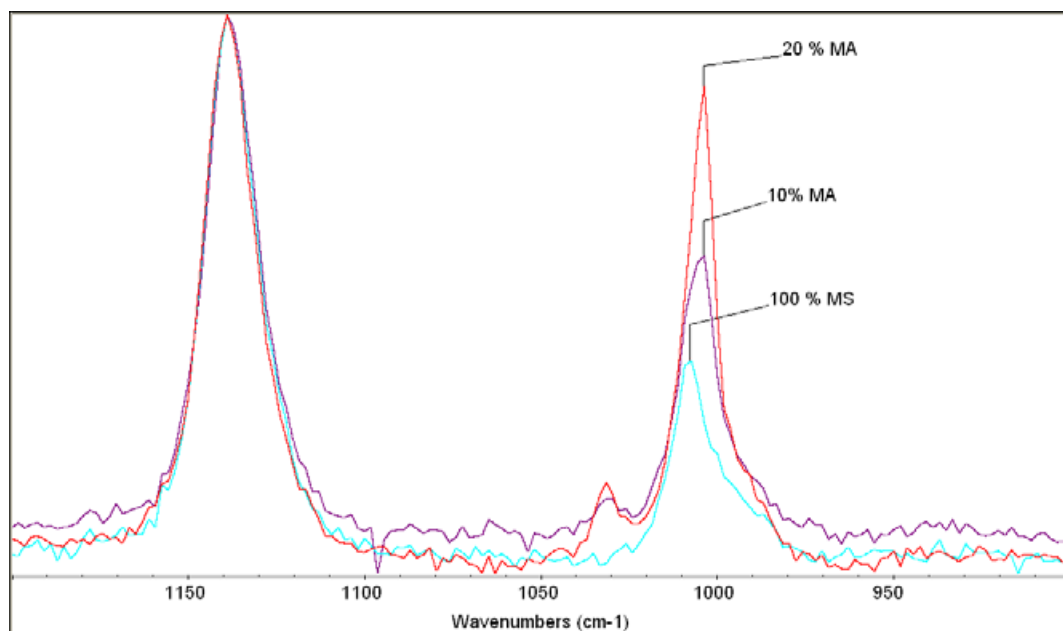FIG. 7—*Ten replicate spectra of a 50/50 mixture of methamphetamine and isopropylbenzylamine hydrochloride.*

FIG. 8—*Spectra of 20% methamphetamine (MA), 10% MA, and 100% methyl sulfone (MS) in the 1200–950/cm spectral range.*

The reproducibility of this technique was investigated by performing 10 replicate analyses on a mixture of IBA and MA. These spectra are collected in Fig. 7.

Detection Limit

Another area of interest is the limit of detection of MA in a mixture with MS using Raman spectroscopy. MA gives a much weaker Raman spectrum than an equal concentration of MS, so MA can be hard to detect at low concentrations. The only discernable difference in the spectra of MS and a 10% MA/90% MS mixture is the height of the peak near 1004/cm. By referencing Fig. 8, one can see that at 10% MA there is little spectral difference with the MS standard. The 1008/cm peak of MS is shifted to 1004/cm in the 10% MA mixture, and the intensity of the peak at 1004/cm does vary relative to the MS peak at 1136/cm, allowing for the detection of MA. However, the concentration must be 20% MA in order to detect a second MA peak at 1032/cm. Based on this analysis, concentrations of MA lower than c. 20% may not be readily detected using Raman spectroscopy. Thirteen of the 100 forensic exhibits in this study were quantified by HPLC as having concentrations of <20% MA, with four of those having no MA detected. Clearly, samples in which no MA is detected by Raman may in fact contain MA, and thus additional instrumental techniques
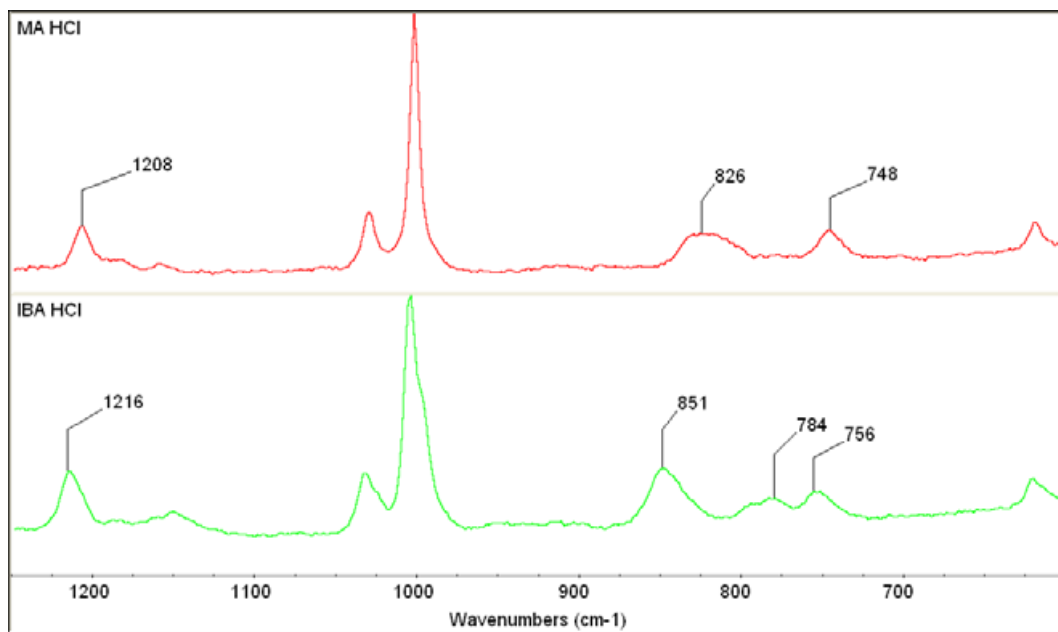
FIG. 9—*Spectra of methamphetamine (top) and isopropylbenzylamine hydrochloride. The significant spectral differences are in the 900–700/cm region (see Fig. 10).*
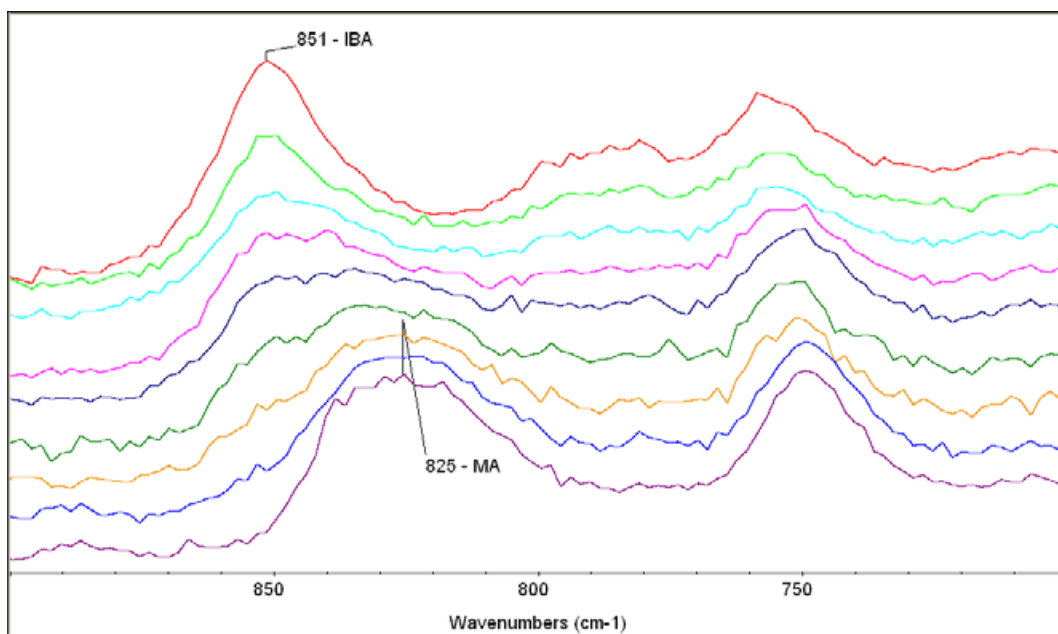

FIG. 10—*Spectra of mixtures of methamphetamine (MA)/isopropylbenzylamine hydrochloride (IBA) in the 900–700/cm region. The relative percentage of MA increases from 0 (top) to 100 (bottom) in 12.5% increments.*

should be employed when analyzing low-concentration MA samples.

Isopropylbenzylamine

Seizures of IBA as ecstasy and crystal meth mimics have been recently reported in the western United States (12–14). The Raman spectrum of IBA is presented with the spectrum of MA for comparison purposes in Fig. 9. While the most intense peak of each compound is at 1004/cm, there are spectral differences at both higher

and lower wavenumbers. MA has peaks at 1208, 826, and 748/cm compared to 1216, 851, 748, and an additional peak at 784/cm in IBA. Spectra of mixtures of IBA and MA were analyzed and are presented in Fig. 10. The mixtures show peak positions intermediate to those of individual compounds. As the relative concentration of IBA to MA increases, there is a shift observed in the peak near 825/cm toward 851/cm. While these spectral shifts are not large, they may be used to indicate the presence of IBA.

Analysis of the forensic exhibits in this study by GCMS revealed IBA present in five FEs. The amount of IBA in four of these items
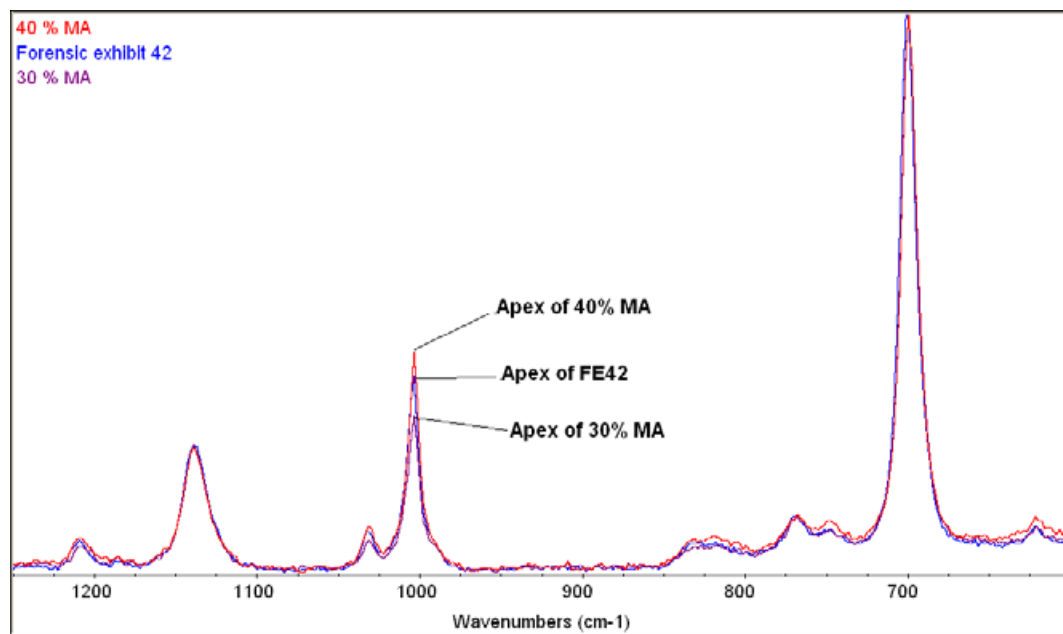
FIG. 11—*Spectrum of FE42 bracketed by spectra of 30% and 40% methamphetamine (MA) standards. The concentration of MA in the FE was determined by high-performance liquid chromatography analysis to be 35.9%.*
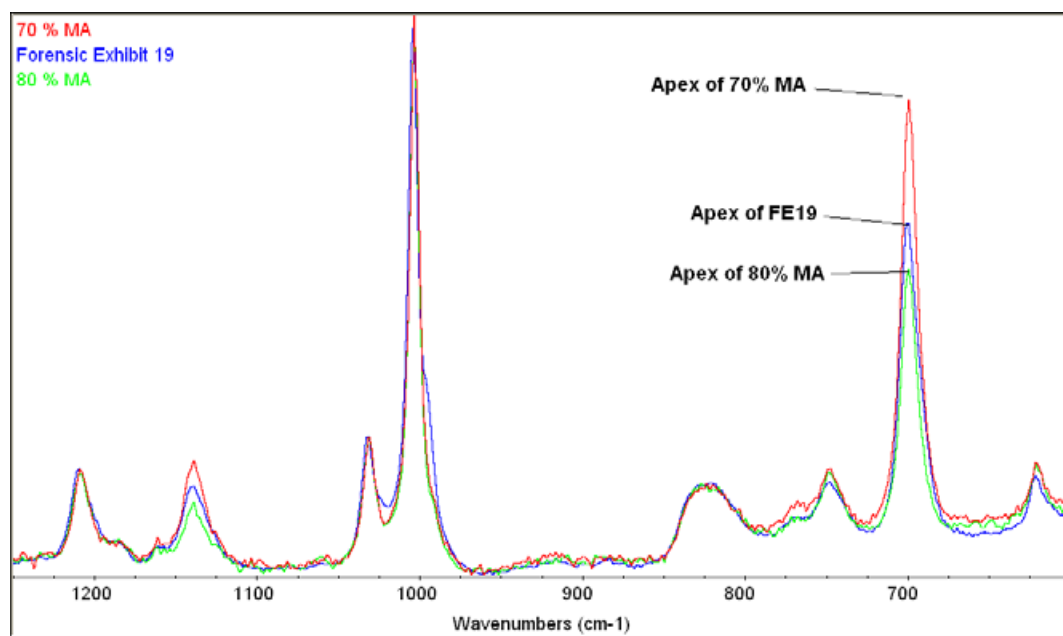

FIG. 12—*Spectrum of FE19 bracketed by spectra of 70% and 80% methamphetamine (MA) standards. The concentration of MA in the FE was determined by high-performance liquid chromatography analysis to be 73.7%.*

was small (*c.* 10% based on chromatographic peak height) relative to the amount of MA present, so the detection of IBA was not possible in the Raman spectra. For the fifth FE, the ratio of MA to IBA was *c.* 1:1 based on chromatographic peak height, but the percentage of MA in the sample was low (8.5%). Therefore, the spectral peaks which would allow for the detection of both IBA and MA were too small to be analyzed. Consistent with the discussion on detection limits, this Raman methodology does not appear well-suited for the analysis of samples containing <20% of either IBA or MA in MS. Also, if there is a small amount of IBA relative to

the amount of MA (or vice versa), Raman is not likely to clearly indicate the presence of the less-abundant component. These limitations need to be assessed when using this technique for crystal meth analysis.

Semi-quantitative Analysis

Because the ratios of the heights of the main spectroscopic peaks of MA (1004/cm) and MS (703/cm) are dependent upon the relative concentrations of the components, "bracketing" the unknown

spectrum to the spectra of known percentages of MA∕MS mixtures allows for a very quick semi-quantitative assessment of crystal meth samples. (The ratio of MA:MS at which the heights of these peaks are the same is *c.* 68:32.) Using the OMNIC program, the most intense peak in the spectra being analyzed is normalized (based on peak height) to the spectra of the known standards. The spectra are then overlaid. The peak height of the second most intense peak then falls in between two of the mixed standard spectra. The concentrations of the two most closely matching spectra then give the approximate concentration range of MA in the FE. In Figs. 11 and 12, this is demonstrated for two FEs, showing the ease with which this can be performed. (See figure legends for HPLC-determined concentrations of MA in these FEs.) In each figure, the MA peak height is intermediate to the height of the standards by which it is bracketed.

Comparison of quantification values by HPLC and Raman analysis was done to see if the percentage of MA determined by HPLC fell in the range indicated by the Raman analysis. In 68 of the 100 FEs analyzed in this study, the HPLC value was included in the Raman range. If the range was expanded by 5% on each side (i.e., instead of 30–40%, the range is increased to 25–45%) an additional 25 FEs were included. When the range was even further expanded to +/−10% of the original range (i.e., from 30–40% to 20–50%), 99 of the 100 samples were included. Only one FE fell outside of the expanded range. These results demonstrate that for the simple MA∕MS mixtures commonly encountered, Raman is a quick means of identification with approximate quantification. Further work may allow for a more precise method for quantitation of MA in simple crystal meth mixtures by finding a mathematical relationship between the peak intensities of the MA and MS peaks and the percent composition of the mixture.

Acknowledgments

References

1. Ryder AG. Classification of narcotics in solid mixtures using principal component analysis and Raman spectroscopy. J Forensic Sci 2002;47(2):275–84.
2. Ryder AG, O'Connor GM, Glynn TJ. Identifications and quantitative measurements of narcotics in solid mixtures using near-IR Raman spectroscopy and multivariate analysis. J Forensic Sci 1999;44(5):1013–9.
3. Ryder AG, O'Connor GM, Glynn TJ. Quantitative analysis of cocaine in solid mixtures using Raman spectroscopy and chemometric methods. J Raman Spectrosc 2000;31:221–7.
4. Carter JC, Brewer WE, Angel SM. Raman spectroscopy for the in situ identification of cocaine and selected adulterants. Appl Spectrosc 2000;54:1876–81.
5. Hodges CM, Hendra PJ, Willis HA. Fourier transform Raman spectroscopy of illicit drugs. J Raman Spectrosc 1989;20:745–9.
6. Neville GA, Shurvell HF. Fourier transform Raman and infrared vibrational study of diazepam and four closely related 1,4-benzodiazepines. J Raman Spectrosc 1990;21:9–19.
7. Tsuchihashi H, Katagi M, Nishikawa M, Tatsuno M, Nishioka H, Nara A, et al. Determination of methamphetamine and its related compounds using Fourier transform Raman spectroscopy. Appl Spectrosc 1997;51:1796–9.
8. Drug detection using Raman spectroscopy—simple, rapid identification and characterisation of illicit substances, http://www.incom.co.jp/topics/copro_pdf/renishaw/Drug%20detection%20using20%Raman%20spectroscopy.pdf
9. Katainen E, Elomaa M, Laakkonen U, Sippola E, Niemela P, Suhonen J, et al. Quantification of the amphetamine content in seized street samples by Raman spectroscopy. J Forensic Sci 2007;52(1):88–92.
10. Bell SEJ, Beattie JR, McGarvey JJ, Peters KL, Sirimuthu NMS, Speers SJ. Development of sampling methods for Raman analysis of solid dosage forms of therapeutic and illicit drugs. J Raman Spectrosc 2004;35:409–17.
11. Bell SEJ, Burns DT, Dennis AC, Speers JS. Rapid analysis of ecstasy and related phenethylamines in seized tablets by Raman spectroscopy. Analyst 2000;125:541–4.
12. N-Isopropylbenzylamine hydrochloride (as "ice" methamphetamine mimics) on the west coast. Microgram Bull 2008;41(3):31–2.
13. Very large seizure of N-isopropylbenzylamine hydrochloride in Bakersfield, California. Microgram Bull 2008;41(4):38–9.
14. Ecstasy mimic tablets (actually containing isopropylbenzylamine, cocaine, and caffeine) in Los Angeles, California. Microgram Bull 2008;41(8):71.

Additional information and reprint requests:
Robert G. Weston, M.Sc.
Oklahoma State Bureau of Investigation
800 E. 2nd Street
Edmond, OK 73034
E-mail: robert.weston@osbi.ok.gov

TECHNICAL NOTE
CRIMINALISTICS

Lothar Schwarz,[1] *Dr. Phil. Nat. and Inga Klenke,*[1] *C.T.A.*

Improvement in Latent Fingerprint Detection on Thermal Paper Using a One-Step Ninhydrin Treatment with Polyvinylpyrrolidones (PVP)

ABSTRACT: Most thermosensitive surfaces of thermal paper turn black when they come into contact with polar organic solvents such as are used in ninhydrin petroleum benzin solution. This dark staining reduces the contrast between the developed fingerprint and the background to such an extent that the identification process becomes very difficult. Integrating polyvinylpyrrolidones (PVP) into a ninhydrin solution prevents the black staining, and the developed fingerprints appear in clear contrast to the background. The new ninhydrin solution containing PVP is successful compared to the two-step ninhydrin–acetone washing method for thermal paper which is popular in Germany.

KEYWORDS: forensic science, fingerprint, detection, thermal paper, ninhydrin, polyvinylpyrrolidones, blackening

Thermal paper is a problematic item because most thermosensitive layers become dark when they come into contact with the conventional ninhydrin petroleum benzin (NPB) (petroleum ether) solution. So the contrast between the background and the fingerprint which is developed is greatly reduced resulting in the identification process becoming considerably more difficult.

The dark staining is caused by the polar organic solvents contained in the NPB solution. Literature is available on some alternative methods for developing latent fingerprints on thermal paper (1–9).

However, besides the iodine vapor process, only chemical processes containing ninhydrin play an important role in Germany. At the moment, there are primarily three practice-orientated methods in use:

- removing the thermosensitive layer by washing the item with acetone prior to or following the ninhydrin treatment (10,11),
- using hemiketals of ninhydrin and alcohols boiled at higher temperatures (o-alkyl ninhydrin derivates) such as ThermaNin (BVDA int. b.v., the Netherlands) or iso nonyl ninhydrin (INON) (12), and
- decolorizing the thermosensitive layer after the ninhydrin treatment by using a G3 solution.

The G3 solution, which we dealt with in a report in 2007, produces good results (13). It involves the ninhydrin-treated items being dipped in the G3 solution for some seconds. This causes the thermosensitive layer to become decolorized without being removed from the paper base as is the case in the acetone washing process. Therefore, fingerprint compounds in this layer are not removed but are preserved for the identification process. The developed fingerprints become clearly visible and appear blue–violet in

color on many thermal papers which also augments the contrast to the background.

A disadvantage of this method lies in the two-step application because the item must first be treated with ninhydrin and subsequently decolorized with G3 which results in the treatment taking twice as long. Furthermore, the decolorization can fade again over time, and also the contrast between the print and the background decreases. This is probably based on the volatility of the compounds used in the G3 solution.

Therefore, the aim of further experiments was to develop a one-step application that would reduce the application time considerably. The first experiments were already described in 2007.

The integration of ninhydrin into the G3 solution (G3N) resulted in a second phase which is dark blue in color. The result of test strips dipped into the other phase appears more blurred and weaker than in the two-step application.

Reports indicating that glue sticks containing polyvinylpyrrolidones (PVP) can also decolorize ninhydrin-treated thermal paper led to a closer consideration of PVP. PVP has a higher molar mass than the pyrrolidones used in a G3 solution, and it is expected that the decolorization will become more stable in the course of time because of its reduced volatility. In addition, PVP is nontoxic and reasonably priced.

At first, experiments were conducted with PVP K15 ($M_w \sim$ 15,000; Fluka, Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany).

Alcoholic solutions of PVP K15 decolorize thermal paper permanently but can only be diluted with petroleum benzin to a negligible extent. Only small amounts of PVP K15 can be integrated into the ninhydrin solution by using a higher amount of ethanol. The resulting solution is dark blue in color, and the effect on thermal paper is less than desired.

To increase the solubility, a search began to find a PVP with a lower molar mass and Kolloidon® 12 PF (M_w = 2000–3000; BASF ChemTrade GmbH, Burgbernheim, Germany) was found, which is used in the pharmaceutical business.

Initial tests showed than it is possible to incorporate more Kolloidon® 12 PF than PVP K15 into the ninhydrin solution, and the resulting solution does not turn blue, which is explained by the higher purity.

This was the starting point to produce a solution containing both PVP and ninhydrin for a one-step ninhydrin application to thermal paper on a nonpolar solvent base with a balanced ratio of ninhydrin, PVP, and a solubility promoter. This means the concentration of ninhydrin must be sufficient to develop prints, the concentration of PVP must be sufficient to prevent the dark staining, and the amount of polar solvent must be kept as low as possible so as not to affect the treated documents and fingerprints. In comparison, the NPB solution has a ninhydrin concentration of 6.0 g/L and a ratio of polar to nonpolar solvent of 1:24, the G3 solution of c. 1:50. The INON solution has a ninhydrin concentration of 2.3 g/L, and polar solvent is not needed.

Various types of alcohol, nonpolar solvents, and concentrations of ninhydrin and Kolloidon® 12 PF were tested to develop a usable ninhydrin solution containing PVP. Initially, solubility tests involving different nonpolar solvents in different alcoholic Kolloidon® 12 PF solutions were conducted, these were followed by numerous tests aimed at optimizing the composition, and finally a comparison was made between the resulting new ninhydrin solution containing PVP and the acetone washing procedure.

For this comparison, 1-day-old, 1-week-old, and 1-month-old fingerprints on thermal paper are split in two. One half is treated in one step with the new ninhydrin solution containing PVP and the other in two steps with ninhydrin and acetone.

Materials and Methods

Solubility Test of Petroleum Benzin in Alcoholic Kolloidon® 12 PF Solutions

For this test, 0.5, 1.0, or 2.0 g Kolloidon® 12 PF, respectively, are completely dissolved in 10 mL alcohol (methanol pro analysis (p.a.), ethanol abs. p.a., 2-propanol p.a., 1-propanol p.a., 1-butanol p.a.). Then while being stirred, petroleum benzin 40–60°C fraction (p.a.) is added until a clouding begins to form. A record is kept of the amount of petroleum benzin (petroleum ether) used. Afterward, ninhydrin-treated thermal paper test strips are dipped into each resulting solution to decolorize them.

Solubility Test of Nonpolar Solvent in Alcoholic Kolloidon® 12 PF Solutions

For this test, 1.0 g Kolloidon® 12 PF is dissolved completely in 10 mL alcohol (methanol p.a., ethanol abs. p.a., 2-propanol p.a., 1-propanol p.a., 1-butanol p.a.). Then under stirring conditions, the nonpolar solvent (cyclohexane p.a., petroleum benzin 40–60°C fraction p.a., n-hexane p.a., pentane purum, heptane p.a.) is added until such time as a clouding forms. A record is kept of the amount of the nonpolar solvent used. Afterward, ninhydrin-treated thermal paper test strips are dipped into each resulting solution for the purpose of decolorization.

Preparation and Application of the Ninhydrin Solution Containing PVP (NinK12 Solution)

One gram ninhydrin (p.a.) and 2.0 g Kolloidon® 12 PF are dissolved in 50 mL 1-propanol (p.a.) at 30–40°C until a clear solution results. A dilution made up of 450 mL of a mixture of two parts pentane and one part cyclohexane is added to this alcoholic

solution while it is being stirred. The light clouding that forms does not impact the application and disappears overnight.

The application of the resulting solution is similar to the NPB solution: Items are treated by being dipped into the ninhydrin solution containing PVP.

Preparation of the Thermal Paper Test Strips

The thermal paper test strips are produced by printing solution of amino acids on to thermal paper by a modified office bubble jet printer as described in Ref. (13).

Preparation of the NPB Solution

Fifteen grams of ninhydrin (p.a.) is fully dissolved in ethanol abs. (p.a.) and then filled up with petroleum benzin 40–60°C fraction to 2.5 L.

Comparison of the Two-step Ninhydrin-acetone and the One-step Ninhydrin Treatment Containing PVP

For comparative purposes, 1-day-old, 1-week-old, and 1-month-old fingerprints of four different people on 60 receipts of thermal paper from 10 different department stores are cut in the middle.

One half is first treated with NPB solution and after drying developed at c. 65% relative humidity for 2 days at room temperature. Afterward, the half is washed with acetone until the black staining has completely disappeared.

The other half is first treated with ninhydrin solution containing PVP and after drying developed at c. 65% relative humidity for 2 days.

Subsequently, both halves are fitted together accurately again and compared. This comparison is repeated once again 6 months later.

TABLE 1—*Solubility of petroleum benzin in different concentrated alcoholic Kolloidon® 12 PF solutions.*

	10 mL Methanol	10 ml Ethanol	10 mL 2-Propanol	10 mL 1-Propanol	10 mL 1-Butanol
0.5 g PVP, petroleum benzin	10 mL	73 mL	82 mL	130 mL	154 mL
1.0 g PVP, petroleum benzin	8 mL	57 mL	65 mL	102 mL	128 mL
2.0 g PVP, petroleum benzin	6 mL	39 mL	49 mL	72 mL	0 mL*

PVP, polyvinylpyrrolidones.
*PVP does not dissolve completely in the amount of alcohol.

TABLE 2—*Solubility of nonpolar solvent in alcoholic Kolloidon® 12 PF solutions.*

	10 mL Ethanol	10 mL 2-Propanol	10 mL 1-Propanol	10 mL 1-Butanol
1.0 g PVP, cyclohexane	88 mL	121 mL	204 mL	238 mL
1.0 g PVP, petroleum benzin	57 mL	65 mL	102 mL	128 mL
1.0 g PVP, n-hexane	55 mL	59 mL	98 mL	124 mL
1.0 g PVP, pentane	53 mL	57 mL	92 mL	115 mL
1.0 g PVP, heptane	40 mL	52 mL	81 mL	103 mL

PVP, polyvinylpyrrolidones.

FIG. 1—(a) Left: Development by ninhydrin solution containing polyvinylpyrrolidones (PVP) (NinK12); Right: Development by ninhydrin petroleum benzin (NPB) solution. (b) Left: Development by ninhydrin solution containing PVP (NinK12); Right: Development by NPB solution and afterward acetone washing.
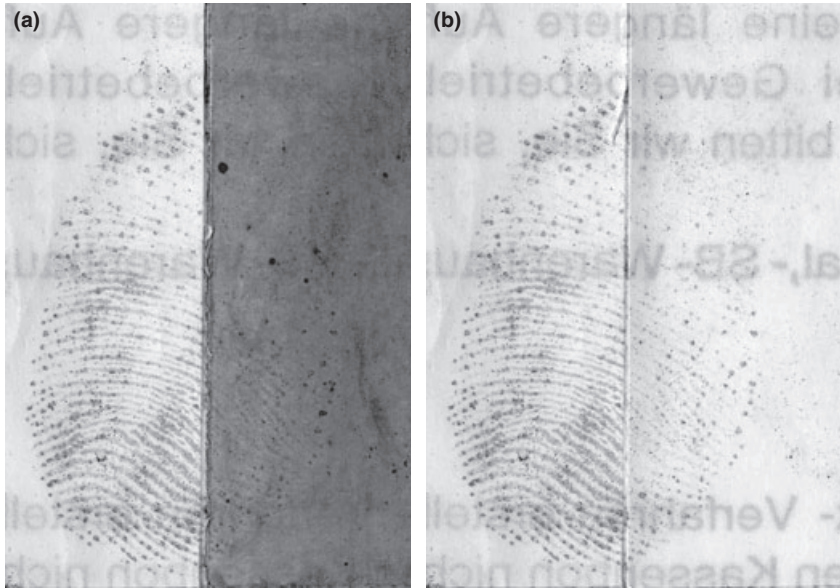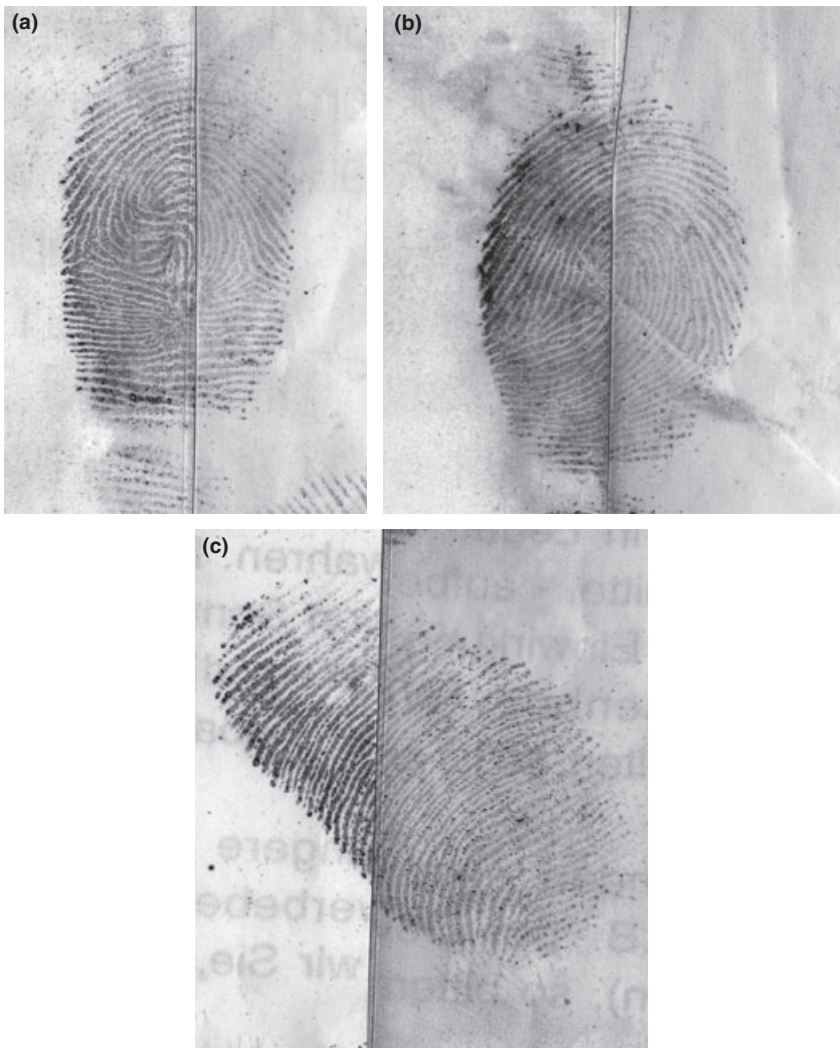
FIG. 2—(a–c) Left: Development by ninhydrin solution containing polyvinylpyrrolidones (NinK12); Right: Development by ninhydrin petroleum benzin solution and afterward acetone washing.

Results and Discussion

As can be seen in Table 1, the solubility of petroleum benzin in an alcoholic PVP solution increases sharply in a sequence ranging from methanol to ethanol, 2-propanol and 1-propanol as far as 1-butanol. At the same time, the ratio of nonpolar to polar solvent decreases with the amount of PVP used. Dipping the ninhydrin-treated thermal paper test strips into the different PVP solutions results in different decolorizations. Excellent results are achieved with 2.0 g PVP in 1-propanol, 2.0 g PVP in 2-propanol, and 1.0 g PVP in 1-butanol.

Extended solubility tests with different nonpolar solvents show similar results, the solubility of the nonpolar solvent increases in the sequence heptane < pentane < n-hexane < petroleum benzin < cyclohexane. The solubility of cyclohexane is about twice as high as that of heptane (see Table 2). Dipping the ninhydrin-treated thermal paper test strips into the different PVP solutions also results in different decolorizations. Excellent results are achieved with 2-propanol, 1-propanol, and 1-butanol containing pentane solution. The cyclohexane solution produces the worst result, probably mainly attributed to the fact that the PVP concentration in this solution is much lower.

1-propanol is chosen for the final ninhydrin solution containing PVP because, following 1-butanol, this can dissolve most nonpolar solvents, and 1-butanol needs a longer time to dry on the dipped test strips. The choice of the nonpolar solvent is a mixture of two parts pentane and one part cyclohexane. This is a compromise between a better print image that results from the use of pentane as distinct from cyclohexane and enhanced PVP solubility which is achieved by using cyclohexane.

The final ninhydrin solution containing PVP (we call it "NinK12") consists of 3.0 g/L ninhydrin and 4.0 g/L PVP and has a ratio of 1:9 polar to nonpolar solvent. The concentration of ninhydrin is therefore only half as high as that in NPB solutions (6.0 g/L) but higher than in INON solutions (2.3 g/L). The ratio of polar to nonpolar solvent is not as favorable as in NPB solution (1:24), and some inks run.

In the direct comparison between ninhydrin solution containing PVP (NinK12) and ninhydrin + acetone washing with 300 prints on 60 receipts, the ninhydrin solution containing PVP (Figs. 1a,b, 2a–c) shows equal or better results on 58 receipts. Ninhydrin and acetone washing is slightly better on two receipts only.

Prints developed by the ninhydrin solution containing PVP (NinK12) are a bluish-violet color, prints developed by ninhydrin + acetone washing are a reddish color. These results remained unchanged during the 6-month observation period.

Conclusions

Overall, the new method, as a one-step method, has the advantage of being time-saving. PVP is nontoxic and can be acquired at a reasonable price. Developed prints are of a bluish-violet color with a high degree of contrast. The prevention of the staining and the bluish-violet color of the prints remain stable over the 6-month observation period. The "NinK12" solution is as easily prepared as the established NPB solution and is also stable. A disadvantage of the new method is the relatively unfavorable ratio from polar to nonpolar solvent of 1:9, which allows for the possibility of particular inks being dissolved on the thermal paper.

If an even lower molar mass PVP can be found, there is a good chance to push the polar to nonpolar solvent ratio to above 1:10, which will largely prevent inks running.

Further tests are intended to show whether it is also possible to integrate PVP in DFO or indandione solutions to open up the possibility of applying these methods as one-step DFO and indandione treatments of thermal paper.

References

1. McMahon P. Procedure to develop latent prints on thermal paper. The Print 1998;14(2):4.
2. Brennan JS, Bramble S, Crabtree S, Wright S. Fuming of latent fingerprints using DMAC. J Forensic Ident 1995;45:373–80.
3. Sasson Y, Almog J. Chemical reagents for the development of latent fingerprints. I. Scope and limitations of the reagent 4-Dimethylamino-Cinnamaldehyde. J Forensic Sci 1978;23:852–5.
4. Broniek B, Knaap W. Latent fingerprint development on thermal paper using muriatic (hydrochloric) acid. J Forensic Ident 2002;52:427–32.
5. Schwarz L, Frerichs I. Advanced solvent-free application of ninhydrin for detection of latent fingerprints on thermal paper and other surfaces. J Forensic Sci 2002;47:1274–7.
6. Sears V re: Latent fingerprint development on thermal paper using muriatic (hydrochloric) acid. J Forensic Ident 2002;52:678.
7. Ma R. Chemical fuming: a practical method for fingerprint development on thermal paper. J Forensic Ident 2006;56:364–73.
8. Wakefield M. The development of latent fingerprints on thermal paper using a novel, solvent-free method. J Forensic Ident 2005;55:202–13.
9. Scott M. Improved results in the development of latent fingerprints on thermal paper. J Forensic Ident 2008;58:424–7.
10. Bartko S. Daktyloskopische spuren auf papier. Kriminalistik 2000;54: 405–10.
11. Bartko S. Das ninhydrinverfahren. Kriminalistik 2002;56:193–5.
12. Takatsu M, Kageyama H, Hirata K, Akashi S, Yoko Ta T, Shiitani M, et al. Development of a new method to detect latent fingerprint on thermal paper with o-alkyl derivative of ninhydrin. Rep Nat Res Inst Police Sci 1991;44:1–6.
13. Schwarz L, Klenke I. Enhancement of ninhydrin- or DFO-treated latent fingerprints on thermal paper. J Forensic Sci 2007;52:649–55.

Additional information—reprints not available from author:
Lothar Schwarz, Dr. Phil. Nat.
Bundeskriminalamt, KT 41
65173 Wiesbaden
Germany
E-mail: Lothar.Schwarz@bka.bund.de

JOURNAL OF FORENSIC SCIENCES

TECHNICAL NOTE

DIGITAL & MULTIMEDIA SCIENCES

Marcos Faundez-Zanuy,[1] Ph.D; Jose J. Lucena-Molina,[2] M.Sc.; and Martin Hagmüller,[3] Ph.D.

Speech Watermarking: An Approach for the Forensic Analysis of Digital Telephonic Recordings*

ABSTRACT: In this article, the authors discuss the problem of forensic authentication of digital audio recordings. Although forensic audio has been addressed in several articles, the existing approaches are focused on analog magnetic recordings, which are less prevalent because of the large amount of digital recorders available on the market (optical, solid state, hard disks, etc.). An approach based on digital signal processing that consists of spread spectrum techniques for speech watermarking is presented. This approach presents the advantage that the authentication is based on the signal itself rather than the recording format. Thus, it is valid for usual recording devices in police-controlled telephone intercepts. In addition, our proposal allows for the introduction of relevant information such as the recording date and time and all the relevant data (this is not always possible with classical systems). Our experimental results reveal that the speech watermarking procedure does not interfere in a significant way with the posterior forensic speaker identification.

KEYWORDS: forensic science, watermarking, digital recording, forensic speaker identification

According to Broeders (1), rapid technical developments in the world of telecommunications, in which speech and data are increasingly transmitted through the same communication channels, may soon blunt the efficacy of traditional telephone interception as an investigative and evidential tool. The gradual shift from analog to digital recording media and the increasingly widespread availability of digital sound processing equipment, as well as its ease of operation, make certain types of manipulation of audio recordings comparatively easy to perform. If done competently, such manipulation may leave no traces and may therefore be impossible to detect. A development in the field of authenticity and integrity examinations of audio recordings in the analog domain are the use of Faraday crystals as pioneered by a number of Russian scientists (2), ferrofluid development and magneto-resistive technology. This potential gain is offset by the widespread availability of relatively inexpensive digital sound processing equipment. Being part of the chain-of-custody process, audio recordings, like all digital data, are therefore increasingly required to be authenticated by means of checksums and hash codes or other methods to ensure their integrity. The formulation of standards for the forensic examination of audio recordings as undertaken by the Scientific Working Group on Digital Evidence (3) is a useful initiative, which may serve to improve standards across the whole field of forensic audio examination. However, it does not provide any technical solution.

[1]Escola Universitària Politécnica de Mataró, Avda. Puig i Cadafalch 101-111, 08303 Mataró, Barcelona, Spain.
[2]Dirección General de la Guardia Civil 28003, Spain.
[3]Graz University of Technology Graz, 8010, Austria.

Digital technology has provided a number of benefits such as the potential for higher audio fidelity, the possibility to make exact copies of an original recording, and the minimization of negative effects such as scratches and other physical defects by means of error correction codes (such as Reed-Solomon code used on CD-audio). However, some drawbacks have also appeared such as the possibility to modify (insert and/or delete) some portions of a given recording in a transparent way that makes this alteration process inappreciable. Forensic applications require the study of the integrity and authenticity of recordings, and this problem is not trivial for digital recordings. The authors believe that the solution to this problem may be found if we change the perspective: to study the signal instead of studying the recording format. For this purpose, some relevant information must be added to the audio signal, with the following desired properties:

• This added information must be imperceptible by the human ear.
• It must be impossible to remove or change without destroying the signal itself. Thus, it cannot be a file header or a different associated file.
• It must be possible to determine the presence of this added information and its content in an easy and unequivocal way (without ambiguities).

This process does not apply to typical forensic audio recordings made by body recorders or intercepted communications that do not have the benefit of the initial embedded watermark.

This article deals with the authentication of monoaural digital speech recording in A-law format (8 bit per sample) for forensic purposes. In addition, there is an additional interest on the feasibility to apply automatic speaker recognition tools. Both issues,

authentication of digital recording and speaker recognition, are fundamental to courts of law.

Approaches to Digital Audio Authentication

In the past few years, several approaches have been presented to complement the traditional techniques for the authentication of digital recordings that are to be presented in court as evidence.

Recently, the use of an electrical network frequency (ENF) has been proposed as a criterion for determining the authenticity of a recording by Grigoras (4). This approach exploits the fact that even though the electrical network frequency is fixed at 50/60 Hz, a certain amount of variation is permitted. The pattern of those variations is unique for every given time instance. Furthermore, the pattern is the same over a connected electrical supply net, e.g., the European power supply network. If a recording with a device connected to the power supply net is done—besides the intended signal—it also may record the ENF signal. Even a battery powered appliance used in the range of the electromagnetic field of the power supply network records the ENF signal. So if one has reference data of the ENF at a certain instance of time, the ENF signal of a questioned recording can be compared to this reference.

The advantage of this method is that the recording in question does not have to be prepared for this type of investigation. On the other hand, because the ENF information is recorded in a very narrow spectral band, which is usually not even used in human speech production, it can easily be removed by simple high-pass filtering. Additionally, patterns of 50/60 Hz variations in the same grid may or may not be proven to be unique at the resolution of current ENF recordings, and it should be considered a proposal.

Another proposal by Cooper (5) was to use the serial copy management system (SCMS) that is implemented in the digital transmission standard of professional and consumer digital audio equipment. The SCMS flags can be used to determine whether the recording presented is an original recording or whether it is possible that a forging attempt has been made. While this approach may work if the potential forger only has access to a standard consumer type equipment, it does not help in case of a well-equipped professional equipment. Some professional sound cards have full control over the SCMS settings, so any SCMS setting can be fabricated. In addition to this, using high-quality D/A and A/D converters may only result in unnoticeable degradation, but would still bypass the SCMS.

The addition of metadata to the sound signal data is also proposed. The metadata can be a hash code, which can be derived from the digital representation of the sound, or content-based information, i.e., the linear prediction coefficients (LPC) of the speech signal. It has been suggested by Alexander and McElveen (6) to incorporate the metainformation with the signal (i.e., the wav audio format allows the embedding of metadata in the file), or transmit the metadata via an additional channel. These approaches introduce the problem of authentication of the metadata because this data can be easily edited with the appropriate tools and specific knowledge of the metadata.

Watermarking

Fortunately, forensics is not the only field where the problem described in section 1 has to be faced. Digital recording technology also permits facile data access and an increased opportunity for violation of copyright and tampering with or modification of content. This problem has been addressed by means of data hiding by Bender et al. (7) and Steinebach and Dittmann (8). Here, the data of interest is directly encoded into the signal, rather than into a file

header, so that it is virtually impossible or very difficult to replace this information by an altered one. Another advantage is that it is not dependent on the type of encoding of the signal.

Introduction

Data hiding and watermarking have received a lot of attention by different research communities during the past two decades. To avoid confusion about terminology, a brief definition of terms and introduction will be provided. An extensive introduction would go beyond the scope of this article. Introductory articles can be pointed out, which cover both the historical methods and current watermarking approaches, such as Petitcolas et al. (9), Hartung and Kutter (10).

Data hiding is the general term for hiding information on either a channel or in a medium. Often, it is also used synonymously with the term "information hiding."

The former is used for example by the military for covert communication or steganography (Greek, meaning "covered writing"). In contrast with encryption, where an unauthorized listener should not be able to decode the information transmitted, the intruder should not even be aware of a message being sent.

The latter would include information embedded in an image or an audio file, which an uninformed user should not be aware of. This is also a form of watermarking.

The important feature for both data and information hiding is that an unauthorized person should not be aware of the transmission of a message.

Watermarking is the embedding of information in a medium, which functions as a host for this additional information. It can either be visible or invisible for the general user.

Watermarking originates in the marking of paper by mills as a means of identifying origin and quality of a sheet of paper. This is still used today sometimes for high-quality paper products. A very sophisticated form of paper watermarking is used as a security feature of money bills. A form of visible digital watermarking is sometimes found in images for copyright protection, where information about the copyright status is visibly embedded in the picture (see Fig. 1).

In contrast, invisible or transparent digital watermarks are not to interfere with the perception of the original multimedia content.

FIG. 1—*Visible watermarking. The copyright information is directly embedded in the image and visibly for the viewer of the image.*

This is the main focus of research for digital media watermarking. Among others, they are used for the purpose of copyright protection, identification, annotation, and authentication of digital media contents.

A watermark is always some additional information connected to a host signal. When the watermark is invisible for the unauthorized person, it is used for data hiding.

There are several, and often conflicting characteristics that are of interest for watermarking.

- *Quantity:* A user typically wants to embed as much data as possible in a given media content. One important factor for the embedding capacity is the bandwidth of the host signal, i.e., there is much more additional data that can be put into a video stream than on a telephone speech channel.
- *Robustness:* The hidden data needs to be invariant when a "host" signal is subject to a certain amount of distortion and transformation, e.g., lossy compression, channel noise, filtering, resampling, cropping, encoding, digital-to-analog (D/A) conversion, and analog-to-digital (A/D) conversion, etc.
- *Security:* The embedded data should be immune to intentional modification attempts to remove or manipulate the embedded data. For authentication purposes, this would be of high importance.
- *Transparency:* The perceptual impact of invisible watermarks on the host content should be minimized. The host signal should not be objectionably degraded, and the embedded data should be minimally perceptible.
- *Recovery:* The embedded data should be self-clocking or arbitrarily reentrant. This ensures that the embedded data can be recovered when only fragments of the host signal are available, e.g., if a sound bite is extracted from an interview, data embedded in the audio segment can be recovered. This feature also facilitates automatic decoding of the hidden data, because there is no need to refer to the original host signal.

As mentioned before, some of these properties are competitive. For example, to maximize security one has to reduce the inserted data quantity. Depending on the application, for some characteristics, a certain trade off has to be made in favor of more important properties.

There has been a lot of research effort in image watermarking and to a lesser extent in music. The main commercial application to date is copyright protection. For instance, commercial products such as Corel Draw and Photoshop include an option for "watermarking" based on Digimarc technology.

It is important to emphasize that this technique is not related to encryption, because watermarking cannot restrict or regulate access to the host signal. Rather, it must ensure that embedded data remains inviolate and recoverable.

In addition, the encryption process is reversible. This means that once a hacker is able to decrypt a file, this file is not protected anymore in the future. On the other hand, watermarking is an irreversible procedure, and the watermark cannot be removed without altering the quality of the host signal to the extent that it is worthless for any other future usage. Watermarking can provide protection of intellectual property rights, give an indication of content manipulation, and provide a means of annotation.

In summary, a general watermarking system (see Fig. 2) consists of an embedding block, which overlays the host data with the additional information. The watermarked signal will then be transmitted on an arbitrary channel to the receiver of the message. On this channel, both harmless modification (such as compression) and malicious attacks aim at modifying or removing the watermark. On the receiver side, the embedded data can be extracted again and, in the case of an authentication application, should give information of whether the content of the media is an original or a modified version.

In our case, the unknown channel depicted in Fig. 2 is a telephone channel, and the watermark embedding and extraction is performed by spread spectrum approach with simplified frequency masking using Linear Prediction (LP) coefficients (11).

While these properties are common for all digital contents, some particularities apply depending on the media type (image, video, text, music, voice, etc.). This article is focused on speech signals.

Spread spectrum radio communications, long a favorite technology of the military because it resists jamming and is hard for an enemy to intercept, is now on the verge of potentially explosive commercial development, according to Schilling et al. (12). The reason is that spread spectrum signals, which are distributed over a wide range of frequencies and then collected onto their original frequency at the receiver, are so inconspicuous as to be "transparent." Just as they are unlikely to be intercepted by a military opponent, so are they unlikely to interfere with other signals intended for business and consumer users.

Spread spectrum uses wideband, noise-like signals. Because spread spectrum signals are noise-like, they are hard to detect.
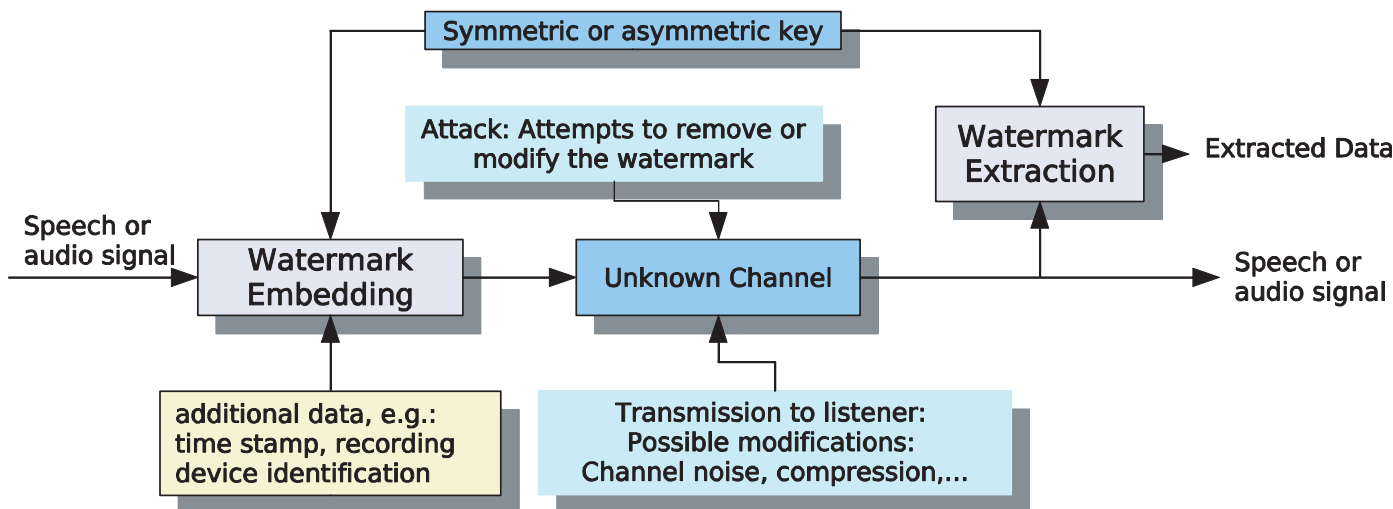

FIG. 2—*Basic watermarking system. Additional data is embedded into the audio stream. Transmitted over an unknown channel and extracted at the receiver side.*

TABLE 1—*Differences between CD-audio and speech watermarking.*

	CD-Audio Watermarking	Speech Watermarking
Channel noise	Very low	Can be high
Bandwidth	Wideband (20 kHz)	Narrowband (<4 kHz)
Allowed distortion	Not perceivable	Low
Processing delay	No issue	Very low (for real-time communication)

Spread spectrum signals are also hard to intercept or demodulate. Further, spread spectrum signals are harder to jam (interfere with) than narrowband signals. These low probabilities of intercept and antijam features are why the military has used spread spectrum for so many years. The spread spectrum techniques, originally developed for military communications, have applications in speech watermarking, as described in the next section.

Watermarking of Speech Signals

Audio watermarking research has focused on copyright protection for digital audio recordings. However, the lack of commercial applications for speech has slowed down the number of publications related to this topic. Watermarking for speech signals is different in some aspects from audio watermarking because of the much narrower signal bandwidth and the special properties of a speech signal. Compared to the 44.1 kilohertz (kHz) sampling rate for CD-audio, telephony speech is usually sampled at 8 kHz. Therefore, less information can be embedded in the signal. For perceptual hiding, the masking levels usually have to be calculated. The commonly used algorithms have their origin in audio compression, such as the widespread Motion Picture Experts Group (MPEG) audio coder. They are optimized for CD-audio bandwidth and computationally very expensive, because compression usually does not need to be performed in real time. Another difference is the expected channel noise. CD-audio recordings usually have very low noise. The audio signal loses its commercial value if the background noise rises beyond a certain threshold. The amount of distortion added to the signal (such as a CD recording) by a watermark should be very low or else the artist and consumer will find it unacceptable. Speech, on the other hand, is very often transmitted over noisy channels, particularly in air traffic control voice communication and telephone speech transmission. On the one hand, the channel noise is a disadvantage; on the other hand, this allows much more power for the watermark signal, because the channel noise will typically cover it. In some applications, the
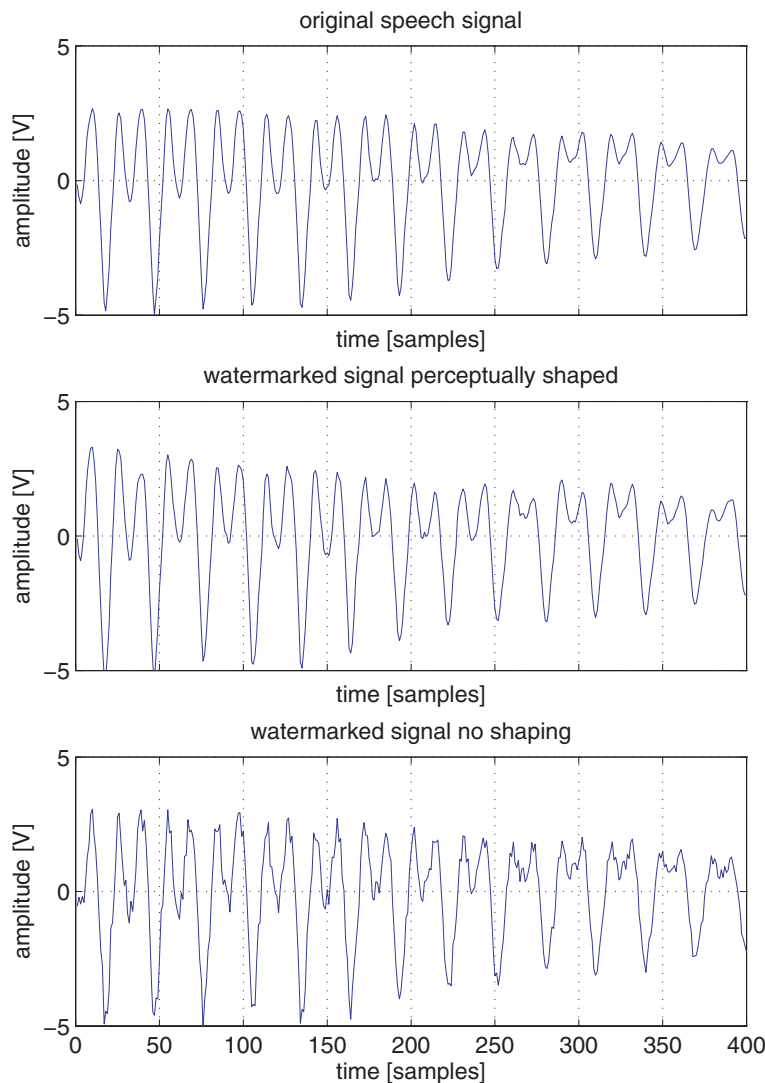

FIG. 3—*Waveform plot of an example of a speech fragment, with (middle) and without (bottom) perceptual weighting compared with the original (top).*

listener expects a certain amount of noise in the signal. A summary of the differences can be seen in Table 1.

Hagmüller and Kubin (13) developed a watermarking system for air traffic control (speech communication among the airplane pilot and the air traffic controller at the airport), with the goal of improving air traffic security. It is based on a spread spectrum approach, which means that low bandwidth information is modulated with a broadband pseudorandom signal, so it is spread all over the available bandwidth. Then Faundez-Zanuy et al. (11) proposed an enhanced biometric security system by combining watermarking with biometric speaker recognition (14). This proposal solves the problem of replay attacks that a biometric recognizer can suffer performing over a remote channel. In this case, the authors used the watermark for introducing an expiry date that makes the use of a prerecording of the genuine user worthless. The watermarking system description can be found in (11). An advantage of the proposed scheme is that the perceptual weighting hides the watermark below the formant peaks of the speech spectrum. Thus, the "noise"

of the watermark is introduced on those portions of the spectrum where the signal has considerable amplitude. This has two important effects:

- The perceptual quality is high for a human listener.
- The LPC envelope is well-preserved, so the posterior LPCC parameters (Cepstral parameters derived from LPC) used by the speech recognizer will be less corrupted than their counterpart without perceptual weighting.

Figures 3, 4, and 5 illustrate this behavior. Figure 3 shows the waveform representation of an original vowel fragment, and the same frame with and without perceptual weighting; it is clearly apparent that the weighted watermark introduces fewer artifacts. Looking at the LPC envelope (Fig. 4) and the periodogram (Fig. 5), it can be seen that the weighting procedure yields a signal closer to the original one. Thus, the LPCC parameters obtained by a recursion from the LPC coefficients (15) will be less corrupted.
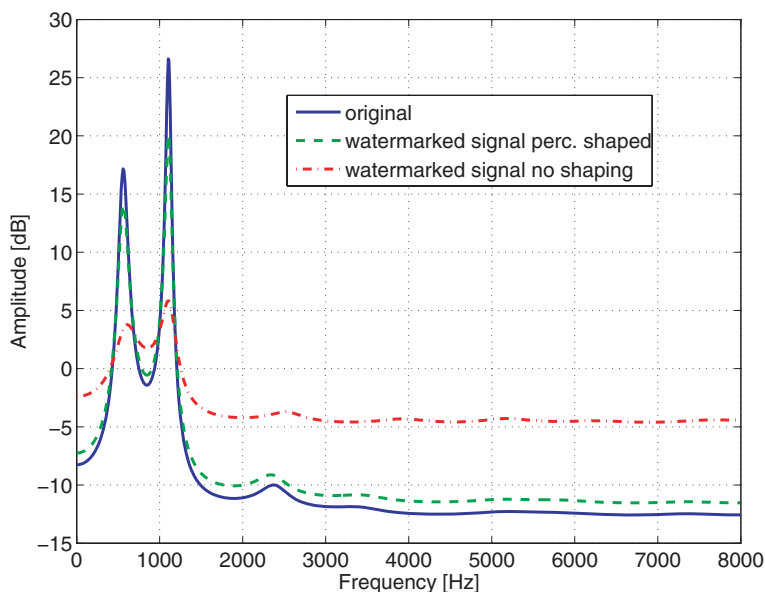

FIG. 4—*Example of LPC spectrum envelope of a speech segment, with and without perceptual weighting, compared with the original.*
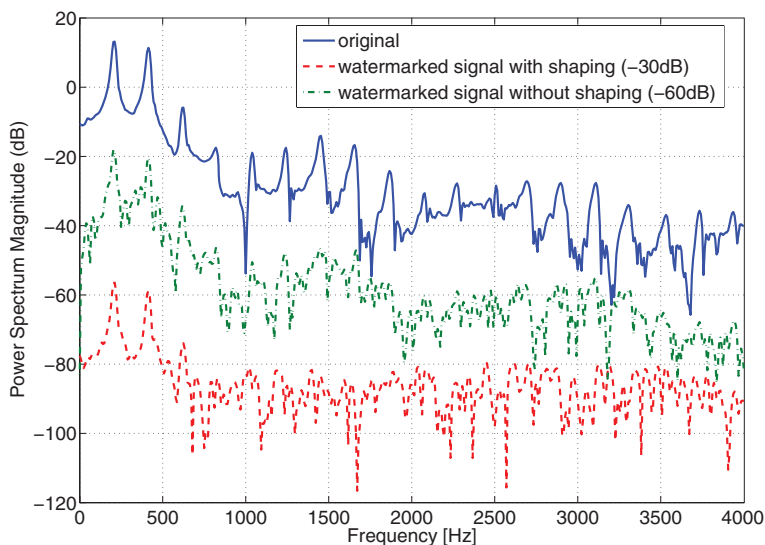

FIG. 5—*Example of power spectrum density of a speech fragment, with and without perceptual weighting compared with the original.*

Watermarking of Speech Signals for Forensic Applications

Recordings made by law enforcement agencies in the exercise of their legal competences to prevent crimes or during crime investigations are mainly carried out at the present time using digital technology. This technology allows us to have digital audio files, and there is a wide and varied collection of equipment, interfaces, protocols, and formats at hand in the market. The available information technology (IT) security systems based on digital technology make the detection of any fraudulent alteration of the original content of a recording almost impossible, reaching information protection levels never seen before with analog technology.

Public document authenticity has always been a concern for public administration regarding the management of official documents. Credibility or authenticity of those documents has been protected using complex physical or chemical procedures.

By applying watermarking to digital technology, i.e., inserting it into digital files, two kinds of information are linked. A speech or image file will contain additional data (watermark) such as the owner name, date, data status (public, restricted), or any other relevant information in compliance with the criteria established by the author who marks the files. Watermarking is performed in an effort to guarantee that the information stored in a file faithfully matches the purpose intended by their authors: to provide data reliability by including a signal certifying its origin.

In particular, watermarking for forensic audio purposes should comply with the following features:

• Inaudibility to reveal its presence only to those individuals legally entitled to know it;
• Ability to be inserted into very short speech segments that could be used in any kind of situations;
• Possibility of using automatic speaker recognition systems without a relevant efficiency loss;
• Enough flexibility to insert information with informative or identity value: time stamping, anagrams, encrypted information, etc.
• Irreversibility to make impossible watermark removal.

Owing to the fact that information gathered by law enforcement agencies is submitted to other legally entitled partners (civil, military, and judicial authorities, etc.), when this information involves digital data it has to be protected. The protection should take into account that a part of the information contained in those protected files could be used in later processes such as copies aimed at authorized listeners. It should be possible to authenticate the origin of any digital information from law enforcement agencies, regardless whether or not it was directly provided by these agencies or when it involves partial or full copies of it. This goal can be fulfilled by means of watermarking, which makes possible the detection of any insertion or erase processes implemented in the original recording. Therefore, watermarking improves the credibility of any copy made from the original recorded information.

For the aforementioned reasons, digital signature and watermarking are two IT security systems that are very well matched for audio police recordings. They complement one another. They can be considered equally useful to be applied in any kind of digital law enforcement audio applications.

High Court Jurisprudence (it depends on the countries, but this is the case in Spain) requires police forces to submit the original recordings as a requirement to determine the unquestionable credibility of their content. This is a consequence of the presumption of innocence principle which establishes that the defendant is presumed innocent until proven guilty by the prosecution. Therefore, the authors consider it strongly advisable to take IT security measures:

• To perform an effort to guarantee the integrity of original recordings;
• To make possible the detection of nonauthorized alterations; and
• To perform an effort to guarantee the origin of the information when segmentation of the original recording is needed and those segments are disclosed to third parties.

Proposed Digital Audio Watermarking System

Description

The proposed system (see Fig. 6) uses watermarking technology, such as briefly introduced in section 2.2, to embed a time-stamp information into the voice signal. This should be done directly in the recording device. The system proposal embeds a watermark, which contains the time of the current recording at every second,
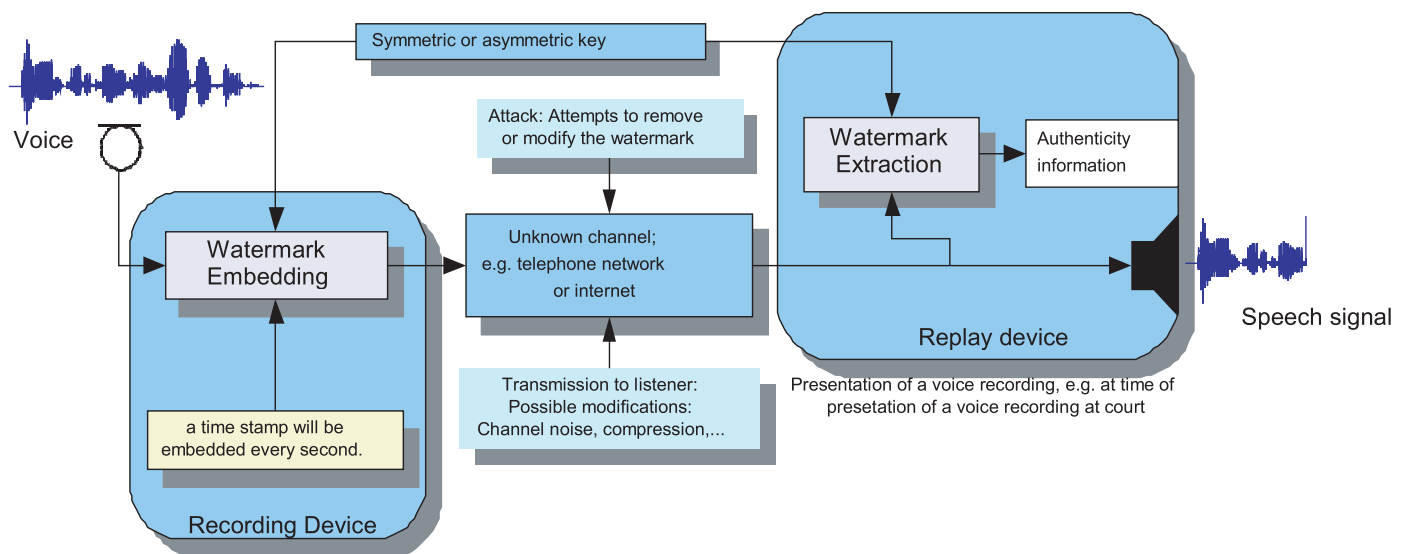

FIG. 6—*Watermarking system for speech authentication.*

TABLE 2—*Time encoding in 36 bit (year, month, day, hour, minute, second).*

	Year	Month	Day	Hour		
Format	y	m	d	h	m	s
Range	[2000, 3023]	[1, 12]	[1, 31]	[0, 23]	[0, 59]	[0, 59]
Bits	10	4	5	5	6	6

based on the device's internal clock. For this purpose, 36 bits/second are inserted, which is enough to embed the exact time for every second of the speech stream (see Table 2). Any insertion or deletion would destroy the synchronization and therefore show the place of the modification at accuracy of 1 sec. The embedding and reading of the time stamp can only be performed if a key, which determines the modulation sequence, is available.

The watermarked signal can then be stored or distributed, in the case of, for example, the use of a voice sample in court so that the authenticity of the recording can be determined. If the recording has been tampered with, e.g., a word has been removed, replaced, or inserted, the integrity of watermark at this time instance will be lost. Because of the time resolution of the embedded time stamp, not only can the general authenticity of the recording be investigated, but also the exact second at which a possible tampering

attack has been made will be apparent as a break in the continuous time stamp (see Table 3).

The next section is devoted to experiments about the interaction between speech watermarking and speaker identification for forensic applications. It is important to emphasize that the use of automated speaker identification/verification systems is not presently widely accepted in the field, nor in courts of law worldwide. For instance, U.S. courts do not presently accept this method of speaker identification/voice comparison. Nevertheless, it is accepted in some courts, including the following countries: Spain, France, Netherlands, Chile, Colombia, South Korea.

Database Description and Experimental Results

All the speech material used in the experiment belongs to the B.D.R.A. (Base de Datos de Registros Acústicos; for information about this forensic database, there is the following contact: crim-acustica@guardiacivil.es), a Spanish public file belonging to the Spanish Ministry of Interior for scientific purposes. The speaker recognition system used is based on the ATVS GMM-MAP-UBM technology submitted to NIST 2004 SRE (16). Nevertheless, the use of automatic speaker recognition systems is not presently accepted everywhere in the world, particularly in U.S. courts of law, for recordings produced under investigative or forensic conditions. Alexandera et al. (17) have compared human and machine

TABLE 3—*Different scenarios for the speech watermarking system. Only the untouched speech file is able to reproduce the correct stream of time stamps. Any modification, i.e., deletion, insertion, replacement results in a random or mission time stamp, is detected.*

| Scenario	Stream of time stamps			
Original and untampered	2009-06-02 10:23:20	2009-06-02 10:23:21	2009-06-02 10:23:22	2009-06-02 10:23:23
One second in missing (in sync with the watermark)	2009-06-02 10:23:20	2009-06-02 10:23:22	2009-06-02 10:23:23	2009-06-02 10:23:24
One second in missing (at arbitrary position)	2009-06-02 10:23:20	2012-12-05 11:41:52	2009-06-02 10:23:23	2009-06-02 10:23:24
Part of a second is missing	2009-06-02 10:23:20	2029-01-03 03:56:32	2009-06-02 10:23:22	2009-06-02 10:23:23
Part of a second is replaced	2009-06-02 10:23:20	2001-12-31 22:43:11	2009-06-02 10:23:22	2009-06-02 10:23:23
More than a second is inserted	2009-06-02 10:23:20	2001-12-31 22:43:11	2013-03-18 09:45:23	2009-06-02 10:23:22
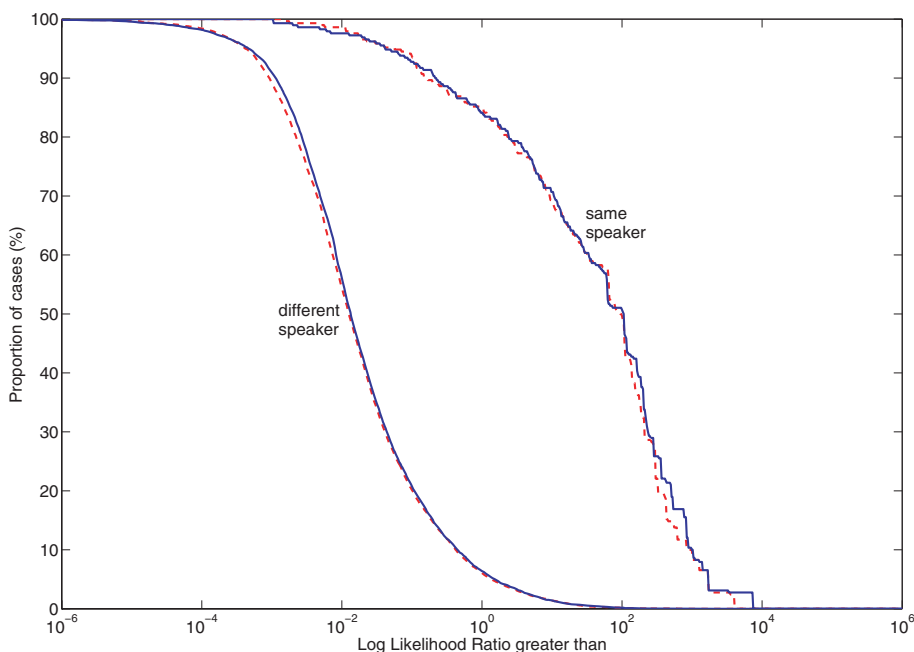
FIG. 7—*Tippet plot for a given Log Likelihood ratio. Clean (solid line) and watermarked (dashed line) signal show no significant difference in performance.*

ability to recognize speakers in forensic applications, and they observed that in matched recording conditions of training and testing, the automatic systems (based on GMM) showed significantly better performance than the aural recognition systems. In mismatched conditions, however, the baseline automatic systems showed comparable or slightly degraded performance compared to the aural recognition systems.

The experiment consists of evaluating the system performance by means of a tool called "EvalIdentiVox2004" (18), a script developed by AGNITIO (19) to calculate DET (detection error trade off) (20), and Tippet plots using test files, speaker models, and reference populations. In this experiment, 300 test files have been used, each one 10 sec long, spontaneous speech, GSM channel, all of them from legal phone tapings carried out from 2000 to 2004 and involving voices recorded on tapes. Ten files per speaker were necessary, with a total of 30 male speakers, all of them speaking Spanish, resulting in the aforementioned 300 files. Each utterance is independent, i.e., it is completely different from the rest.

Two additional minutes of speech from each of the 30 speakers were recorded in conversations independent from the aforementioned ones. These files were used to calculate the unquestioned acoustic model for each speaker.

With regard to the reference population, it comprised 35 male Spanish speaking speakers different from the 30 earlier mentioned. The samples were obtained from legal phone tapings as well, under the same conditions.

Comparing the Tippet (see Fig. 7) plot in the experiment, 300 target scores were calculated along with 8700 nontarget scores. The results from the original files are depicted in blue (solid line) and those obtained using the watermarked files in red (dashed line). The performance differences are irrelevant. The appropriate combination between noise and speech has avoided a substantial difference in performance.

Conclusions

In this article, the authors have discussed the problem of authentication of speech recordings commonly encountered in the forensic audio field. This article has summarized the existing approaches and has proposed a new one based on perceptual speech watermarking, which is independent of the recording device (optical, magnetic, solid state). In addition, a set of experiments on a forensic database has been performed, and it has been found that speech watermarking does not interfere with the posterior biometric speaker identification∕verification.

References

1. Broeders APA. Forensic speech and audio analysis forensic linguistics 1998 to 2001 a review. 13th INTERPOL Forensic Science Symposium; 2001 Oct 16–19. Lyon, France: 2001;pp 53–84, http://www.interpol.int/Public/Forensic/IFSS/meeting13/Reviews/ForensicLinguistics.pdf (accessed on March 8, 2010).
2. Grechishkin RM, Goosev MYu, Ilyashenko SE, Neustroev NS. High-resolution sensitive magneto-optic ferrite-garnet films with planar anisotropy. J Magn Magn Mater 1996;157-158:305–6.
3. Scientific Working Group on Digital Evidence. http://www.swgde.org (accessed on March 8, 2010).
4. Grigoras C. Applications of ENF criterion in forensic audio, video, computer and telecommunication analysis. Forensic Sci Int 2007;167(2):136–45.
5. Cooper AJ. The significance of the Serial Copy Management System (SCMS) in the forensic analysis of digital audio recordings. Int J Speech Lang Law 2005;12(1):49.
6. Alexander A, McElveen JK. An open-standard file format for forensic audio. Proceedings of the 2006 International Association for Forensic Phonetics and Acoustics Meeting; 2006 Jul 23–26; Göteborg, Sweden: Sweden University, 2006.
7. Bender W, Gruhl D, Morimoto N, Lu A. Techniques for data hiding. IBM Systems J 1996;35(3-4):313–36.
8. Steinebach M, Dittmann J. Watermarking-based digital audio data authentication∕EURASIP. J Appl Signal Processing 2003;10:1001–15.
9. Petitcolas FAP, Anderson RJ, Kuhn MG. Information hiding—a survey. Proc IEEE Inst Electr Electron Eng 1999;87(7):1062–78.
10. Hartung F, Kutter M. Multimedia watermarking techniques. Proc IEEE Inst Electr Electron Eng 1999;87(7):1079–107.
11. Faundez-Zanuy M, Hagmüller M, Kubin G. Speaker verification security improvement by means of speech watermarking. Speech Commun 2006;48(12):1608–19.
12. Schilling DL, Pickholtz RL, Milstein LB. Spread spectrum goes commercial. IEEE Spectrum 1990;27(8):40–1, 44–5.
13. Hagmüller M, Kubin G. Speech watermarking for air traffic control. EEC note 2005∕5, Eurocontrol Experimental Centre, Brétigny-sur-Orge, France and Signal processing and speech communication laboratory, Graz University of Technology, Austria, http://www.eurocontrol.int/eec/gallery/content/public/document/eec/report/2005/005_Speech_Watermarking_for_ATC.pdf (accessed on March 8, 2010).
14. Faundez-Zanuy M, Monte-Moreno E. State-of-the-art in speaker recognition. IEEE Aerosp Electron Syst Mag 2005;20(5):7–12.
15. Deller JR, Hansen JHL, Proakis JG. Discrete-time processing of speech signals. Upper Saddle River, NJ: Prentice-Hall, 1993.
16. NIST. 2004 speaker recognition evaluation plan, http://www.itl.nist.gov/iad/mig/tests/sre/2004/ (accessed on March 8, 2010).
17. Alexander A, Botti F, Dessimoz D, Drygajlo A. The effect of mismatched recording conditions on human and automatic speaker recognition in forensic applications. Forensic Sci Int 2004;146(Suppl 1-2):S95–9.
18. Gonzalez-Rodriguez J, Drygajlo A, Ramos-Castro D, Garcia-Gomar M, Ortega-Garcia J. Robust estimation, interpretation and assessment of likelihood ratios in forensic speaker recognition. Computer Speech Lang 2006;20(2-3):331–55.
19. http://www.agnitio.es (accessed on March 8, 2010).
20. Faundez-Zanuy M. Biometric security technology. IEEE Aerospace and Electronic Systems Magazine 2006;21(6):15–26.

Additional information and reprint requests:
Marcos Faundez-Zanuy, Ph.D.
Telecommunication Department
EUP Mataró
Avda. Puig i Cadafalch 101-111
Mataro 08303
Spain
E-mail: faundez@eupmt.es

TECHNICAL NOTE

PATHOLOGY AND BIOLOGY

Nick I. Batalis,[1] M.D.; Bradley J. Marcus,[2] M.D.; Christine N. Papadea,[1] Ph.D.; and Kim A. Collins,[3] M.D.

The Role of Postmortem Cardiac Markers in the Diagnosis of Acute Myocardial Infarction*

ABSTRACT: Sudden cardiac deaths because of acute myocardial infarction (MI) constitute a significant percentage of the caseload for death investigators, coroners, and forensic pathologists. Clinicians use cardiac markers, highly sensitive and specific for myocardial damage, to screen living patients for acute MI; however, to this point, the utility of these markers in the autopsy setting has not been fully established. The current study included 10 decedents, five who died of acute MI, and five subjects who died of noncardiac disease. Samples of pericardial fluid and blood from multiple sites were tested for creatine kinase, creatinine kinase MB, and troponin-I. Three main conclusions were drawn: the levels of cardiac markers from all patients are significantly higher than the reference range for living patients, there are significant differences in cardiac marker levels between samples from different anatomic locations, and only three cardiac marker/anatomic site combinations were significantly different between the control and study groups.

KEYWORDS: forensic sciences, forensic pathology, myocardial infarction, cardiac markers, death, heart, autopsy

Biochemical markers of cardiac damage, especially troponin I and troponin T, are well documented and widely used by clinicians as sensitive and specific indicators of acute cardiac disease (1,2). Patients presenting with symptoms concerning for acute myocardial infarction (MI) are often subjected to a standardized battery of tests including a chest X-ray, an electrocardiogram (EKG), and blood samples submitted for cardiac markers. Often times, if the clinical presentation is suspicious and elevated cardiac markers are noted patients will be taken for immediate cardiac catheterization, even in the absence of EKG changes. Unfortunately, while cardiac deaths comprise a significant volume of a forensic pathologist's caseload, these cardiac markers have not shown the same predictive value in the autopsy literature (3–6). The pathologist, then, is left to diagnose MI at autopsy. While this is possible the vast majority of the time, certainly there are cases of assumed acute myocardial infarction with little or no gross or histologic evidence of myocyte injury. These cases demonstrate the need for a postmortem biochemical, or other, marker of acute myocardial infarction and constitute one reason researchers continue to investigate the role of postmortem cardiac markers (PCM) in death certification (3–19). Studies have also researched markers for other forms of heart disease such as brain naturetic peptide, a protein elevated in many cases of congestive heart failure (20).

Authors have investigated the role of PCM at autopsy and have had varying results (3–13). Some of the studies include a correlation of postmortem and antemortem levels of cardiac markers (3,7), a comparison of postmortem serum and pericardial fluid levels (8–10), and several have attempted to determine if postmortem levels are significantly higher in deaths because of myocardial ischemia than those because of other causes of death (4–6,8–13). Another study suggested using PCM as a way of triaging bodies that present for postmortem examination and using this information, along with other findings, to determine whether to conduct a complete autopsy or just an external examination with ancillary studies (12). In the appropriate setting, this could be a good use for these markers; however, to date there is little information regarding the variability of levels of cardiac markers in serum from different anatomic locations within the same subjects (5,6). As blood can be taken from various sites at external exam, it is important to know if there is analyte variability between sites if these PCM results are to play a role in death certification.

Materials and Methods

Blood samples were obtained from 10 nonconsecutive autopsy patients presenting to the Forensic Pathology Section of the Medical University of South Carolina. In eight of the 10 patients, pericardial fluid was also sampled (insufficient quantity in the other two patients). Patients were chosen according to the circumstances of death, suspected cause of death, availability of blood, and postmortem interval (<24 h). Bodies were refrigerated prior to autopsy in each patient; however, as our facility receives bodies from multiple counties across the state, the intervals between death and refrigeration varied. A complete autopsy, including microscopic examination and toxicology, was performed on each patient. Based on the autopsy findings and final cause of death determination, the patients were assigned either to a MI group (*n* = 5) or a control group (*n* = 5) composed of deaths from a variety of causes, excluding those with significant chest and/or cardiac trauma. All five

[1] Medical University of South Carolina, Suite 309, 171 Ashley Avenue, PO Box 250908, Charleston, SC 29425.
[2] Professional Pathology Services, Columbia, SC.
[3] Fulton County Medical Examiner's Office, Atlanta, GA.
*Presented at the 59th Annual Meeting of the American Academy of Forensic Sciences, February 19–24, 2007, in San Antonio, TX.

patients in the study group had acute, occlusive coronary artery thrombi at autopsy.

Upon internal examination, blood samples were obtained from the femoral veins, subclavian veins, aorta, right cardiac ventricle, and left cardiac ventricle and promptly placed in marble-top serum separator tubes. To ensure peripheral samples were not mixed with more central blood, the subclavian and femoral veins were isolated and cross-clamped, and then blood was drawn distal to the clamp. Samples were then either centrifuged in the autopsy suite and transported to the laboratory after completion of the case, or taken directly to the laboratory. In eight of the 10 patients, pericardial fluid was drawn and placed in a sterile red top tube (no additive) and transported with the blood samples. An insufficient amount of fluid was present in the other two patients.

Testing was performed on the Bayer Advia Centaur® immunology analyzer. This analyzer utilizes direct chemiluminescence assays to quantify the amount of each analyte. All samples were tested for TnI, CK, and CK-MB. A relative index (RI) was then calculated determining the percentage of total CK which CK-MB accounted. Many of the samples yielded concentrations much higher than normally seen in living patients; these samples were serially diluted and retested before a concentration was reported. A statistical analysis was performed using SPSS® version 14.0. The nonparametric Mann–Whitney U-test was used to compare the control and acute MI groups, and the nonparametric Kruskal–Wallis test was used to compare the analyte levels at the different anatomic locations. In both tests, a p-value of 0.05 was considered statistically significant.

Results

Both the control and acute MI groups contained four men and one woman. The median age of the control group was 41.8 (range 34–56) and the median age of the cardiac death group was 49.2 (range 39–61). The cause of deaths of the control groups included two patients with acute cocaine toxicity, two gunshot wounds of the head, and one witnessed terminal seizure. Three of the control patients had no significant cardiovascular pathology, while two exhibited moderate atherosclerosis and interstitial cardiac fibrosis. In addition to the acute thrombi of the study group, all five acute MI deaths demonstrated interstitial fibrosis, mild to severe coronary atherosclerosis, and microscopic contraction band necrosis. Resuscitation was not controlled for as this information was not available in all cases, and previous studies have shown this does not significantly affect PCM levels (4–6,12).

At the time of the study, our laboratory did not grade hemolysis on each specimen, but did note those in which hemolysis was visually present. Varying degrees of hemolysis were identified in nearly all specimens; however, this was not controlled for as to keep the study as practical as possible for the practicing pathologist as the majority of patients presenting for autopsy will have at least low levels of hemolysis.

Data for TnI and CK-MB levels among the control and study groups are presented in Tables 1 and 2. Only three anatomic site/analyte combinations were significantly different between the study and control groups: femoral-TnI, right ventricle-CK-MB, and pericardial fluid-CK-MB. The p-values for these combinations were all extremely low (0.0079–0.0317) and despite markedly elevated PCM concentrations with rather large standard deviations, the standard deviation ranges for femoral-TnI and pericardial fluid-CK-MB do not overlap.

TnI (Table 3) and CK-MB (Table 4) concentrations by anatomic site, including the control and cardiac groups, are presented along

TABLE 1—*Comparison of troponin I (ng/mL) values for all subjects at various anatomic sites.*

Control			Acute Myocardial Infarction			
Site	Mean	SD	Site	Mean	SD	p-Value
LV	18891	31643	LV	62689	44494	0.0952
RV	9051	14153	RV	7057	7924	1.0
SC	2922	5513	SC	3516	2884	0.3095
AO	7581	16104	AO	38203	44012	0.095
FE	7.2	5.9	FE	102	78.4	0.0079
PC	8461	12569	PC	20144	22330	0.3429

SD, standard deviation; LV, left ventricle; RV, right ventricle; SC, subclavian; AO, aorta; FE, femoral vein; PC, pericardial fluid.

Normal range for living patients at our institution: 0–0.2 ng/mL.

TABLE 2—*Comparison of CK-MB (ng/mL) values for all subjects at various anatomic sites.*

Control			Acute Myocardial Infarction			
Site	Mean	SD	Site	Mean	SD	p-Value
LV	864	731	LV	627	117	0.6905
RV	878	482	RV	1493	281	0.0317
SC	457	294	SC	535	313	0.8413
AO	191	141	AO	432	135	0.0556
FE	144	90	FE	153	128	1.0
PC	458	285	PC	2182	1140	0.0286

SD, standard deviation; LV, left ventricle; RV, right ventricle; SC, subclavian; AO, aorta; FE, femoral vein; PC, pericardial fluid.

Normal range for living patients at our institution: 0–4.9 ng/mL.

with the Kruskal–Wallis statistic and p-values. According to this data, only femoral vein blood showed statistically significant differences in its concentrations of TnI and CK-MB when compared to other draw sites. For TnI, when the sites are compared by pairs (Bonferroni contrast), there are significant differences between levels from the femoral veins and all other sites, with exception of the subclavian veins. CK-MB concentrations from the femoral veins were significantly different than CK-MB concentrations from the left ventricle, right ventricle, and pericardial fluid, but not the aorta or subclavian veins. The other draw sites when compared to each other, then, did not show any significant differences in CK-MB or TnI concentrations. It also should be noted that levels of all PCM, in the control and study groups, were all significantly higher than the institutional normal ranges for living patients (TnI <0.2 ng/mL, CK 25–260 IU/L, CK-MB <4.9).

CK concentrations were not statistically different between the study and control groups, likely because of the consistently elevated levels of CK in all patients because of autolysis. Also owing to the markedly elevated CK concentrations, the RI remained minimal and statistically insignificant, even in those patients with acute MI with elevated CK-MB concentrations.

TABLE 3—*Comparison of all troponin I (ng/mL) values (study and control groups) at various anatomic sites.*

Site	TnI Mean	SD	Range	Kruskal–Wallis Statistic
LV	40790	43102	100–139288	22.93
RV	8052	10864	164–33845	p-Value
SC	3219	4160	4–12676	0.0003
AO	22392	35166	6–91793	
FE	54	73	0.2–196	
PC	14228	17956	57–52814	

SD, standard deviation; LV, left ventricle; RV, right ventricle; SC, subclavian; AO, aorta; FE, femoral vein; PC, pericardial fluid.

Normal range for living patients at our institution: 0–0.2 ng/mL.

TABLE 4—*Comparison of all CK-MB (ng/mL) values (control and study groups) at various anatomic sites.*

Site	CK-MB Mean	SD	Range	Kruskal–Wallis Statistic
LV	746	509	254–1850	32.36
RV	1186	494	469–1801	*p*-Value
SC	496	289	96–1070	0.0001
AO	312	182	51–620	
FE	148	104	12–346	
PC	1320	1200	165–3863	

SD, standard deviation; LV, left ventricle; RV, right ventricle; SC, subclavian; AO, aorta; FE, femoral vein; PC, pericardial fluid.

Normal range for living patients at our institution: 0–4.9 ng/mL.

Discussion

The goals of this limited study were to assess the correlation of PCM and deaths because of acute MI and to determine whether the concentrations of these analytes vary according to anatomic site. This study was limited by factors including but not limited to cohort size, hemolysis, variance in postmortem interval, and postmortem refrigeration time; however, data were generated that may help our understanding of PCM.

First, it is apparent that under routine conditions almost all bodies presenting for autopsy will demonstrate some degree of autolysis and hemolysis which will cause an elevation of PCM concentrations above their premortem levels as free hemoglobin interferes with assays for measuring troponin, CK, and CK-MB (3). In the current study, only a single femoral vein blood TnI level from the control group was within our institution's normal range for living patients, while the remainder of samples exhibited levels that would be suspicious, if not diagnostic of, acute myocardial infarction in a living patient. Again, all testing from this study was performed in the same laboratory, on the same instruments, and by the same technicians as clinical samples of living patients received for diagnostic evaluation. While higher PCM levels because of hemolysis seem intuitive and previous studies have also demonstrated this trend (3,4), others have reported levels similar to living persons and have dismissed the role of hemolysis and autolysis (9). Others have theorized that PCM levels may be elevated in many deaths because of nonspecific myocardial damage because of hypoxia during the agonal period (5,6).

Second, it appears there may be significant differences in analyte concentrations based on the site from which the blood/fluid is drawn. In this study, levels of TnI and CK-MB were consistently lower in femoral vein blood than all other draw sites, with the exception of the subclavian veins. In general, PCM concentrations were lowest in peripheral blood (femoral and subclavian) and highest in central blood and pericardial fluid. These data seem to correlate with two other studies that measured PCM levels from different anatomic locations on the same subjects (5,6). Pericardial fluid represents a different matrix than blood and is a filtrate produced via the serous layer of the pericardium. Being so, chemicals and other markers should become elevated in pericardial fluid after becoming elevated in peripheral blood. However, as the fluid also bathes the myocardium, enzymes and proteins may also be released directly into the pericardial fluid from the cardiac muscle and some studies have shown elevated concentrations of PCM in pericardial fluid when compared to peripheral blood (5,6).

As previously mentioned, although this was a small study, three anatomic site/analyte combinations were statistically significant including femoral-TnI, right ventricle-CK-MB, and pericardial fluid-CK-MB. Of note, the femoral-TnI and pericardial fluid CK-MB relationships are particularly strong.

At this point, there is not enough definitive data to determine a cause of death as acute myocardial infarction based on PCM alone. Some authors have noted that PCM may be a useful way of triaging bodies which come for evaluation (12), that is, potentially just performing an external examination on someone with a medical history suspicious acute MI and positive PCM (12). This seems to be a more reasonable use of these markers, if used strictly as an ancillary study when a complete autopsy will not be performed. However, one must understand that there will be moderate numbers of false positive and false negative results depending on the threshold one sets for an abnormal postmortem value, as almost all values in this study were above the normal ranges for living patients. Additionally, if one performs PCM, then one needs to be conscious of the laboratory's normal values, the increased concentrations seen in postmortem control subjects, and the anatomic site from which the sample is drawn. To date, it seems that, although still significantly higher, femoral vein blood best approximates antemortem physiologic levels, while blood from central locations including the aorta and cardiac ventricles yields much higher PCM levels. While our data also suggest that right ventricular and pericardial CK-MB may be elevated in patients with MI, these specimens would not be easily obtainable without performing a complete examination. For these reasons, if one was to use PCM to triage cases, these data suggest that femoral TnI would be the best indicator of acute MI; however, it is important to corroborate the laboratory results with scene, medical history, and toxicology findings. It should also be noted that larger, more comprehensive studies are needed before definitive conclusions can be drawn from these assays.

One problem with the studies to date (including the current study) is that these studies have tested PCM in obvious patients with acute MI and compared them to patients that were not cardiac deaths (4–6,8–13). Not surprisingly, all of these studies concluded the PCM correlated with the cause of death and supported the clear gross and histologic findings. While these studies are necessary to validate the accuracy of PCM in confirmed acute MI, it may prove more difficult to apply this data to the more subtle patients lacking definitive gross or microscopic cardiac pathology. Before one can make this determination, large-scale studies including pronounced and more subtle patients with acute MI (e.g., a heart with 50% stenosis of a coronary artery but without a thrombus at autopsy) are needed to determine the predictive value of PCM. It is hoped that this study can serve as a building block for such larger studies.

References

1. Jaffe AS, Babuin L, Apple FS. Biomarkers in acute cardiac disease: the present and the future. J Am Coll Cardiol 2006;48(1):1–11.
2. Kavsak PA, MacRae AR, Lustig V, Bhargava R, Vandersluis R, Palomaki GE, et al. The impact of the ESC/ACC redefinition of myocardial infarction and new sensitive troponin assays on the frequency of acute myocardial infarction. Am Heart J 2006;152(1):118–25.
3. Davies SJ, Gaze DC, Collinson PO. Investigation of cardiac troponins in postmortem subjects: comparing antemortem and postmortem levels. Am J Forensic Med Pathol 2005;26(3):213–5.
4. Cina SJ, Thompson WC, Fischer JR Jr, Brown DK, Titus JM, Smialek JE. A study of various morphologic variables and troponin I in pericardial fluid as possible discriminators of sudden cardiac death. Am J Forensic Med Pathol 1999;20(4):333–7.
5. Zhu BL, Ishikawa T, Michiue T, Li DR, Zhao D, Oritani S, et al. Postmortem cardiac troponin T levels in the blood and pericardial fluid. Part 1. Analysis with special regard to traumatic causes of death. Leg Med (Tokyo) 2006;8(2):86–93.
6. Zhu BL, Ishikawa T, Michiue T, Li DR, Zhao D, Kamikodai Y, et al. Postmortem cardiac troponin T levels in the blood and pericardial fluid. Part 2: analysis for application in the diagnosis of sudden cardiac death with regard to pathology. Leg Med (Tokyo) 2006;8(2):94–101.

7. Ooi DS, Isotalo PA, Veinot JP. Correlation of antemortem serum creatine kinase, creatine kinase-MB, troponin I, and troponin T with cardiac pathology. Clin Chem 2000;46(3):338–44.

8. Perez-Carceles MD, Noguera J, Jimenez JL, Martinez P, Luna A, Osuna E. Diagnostic efficacy of biochemical markers in diagnosis post-mortem of ischaemic heart disease. Forensic Sci Int 2004;142(1):1–7.

9. Osuna E, Perez-Carceles MD, Alvarez MV, Noguera J, Luna A. Cardiac troponin I (cTn I) and the postmortem diagnosis of myocardial infarction. Int J Legal Med 1998;111(4):173–6.

10. Osuna E, Perez-Carceles MD, Vieira DN, Luna A. Distribution of biochemical markers in biologic fluids: application to the postmortem diagnosis of myocardial infarction. Am J Forensic Med Pathol 1998;19(2):123–8.

11. Khalifa AB, Najjar M, Addad F, Turki E, Mghirbi T. Cardiac troponin T (cTn T) and the postmortem diagnosis of sudden death. Am J Forensic Med Pathol 2006;27(2):175–7.

12. Cina SJ, Brown DK, Smialek JE, Collins KA. A rapid postmortem cardiac troponin T assay: laboratory evidence of sudden cardiac death. Am J Forensic Med Pathol 2001;22(2):173–6.

13. Ellingsen CL, Hetland O. Serum concentrations of cardiac troponin T in sudden death. Am J Forensic Med Pathol 2004;25(3):213–5.

14. Adegboyega PA, Adesokan A, Haque AK, Boor PJ. Sensitivity and specificity of triphenyl tetrazolium chloride in the gross diagnosis of acute myocardial infarcts. Arch Pathol Lab Med 1997;121(10):1063–8.

15. Agdal N, Andersen MT. The value of succinate dehydrogenase stain in the post-mortem diagnosis of early acute myocardial infarction. A forensic study. Forensic Sci 1978;11(3):223–30.

16. Edston E, Grontoft L, Johnsson J. TUNEL: a useful screening method in sudden cardiac death. Int J Legal Med 2002;116(1):22–6.

17. Blanco Pampin J, Garcia Rivero SA, Otero Cepeda XL, Vazquez Boquete A, Forteza Vila J, Hinojal Fonseca R. Immunohistochemical expression of HIF-1alpha in response to early myocardial ischemia. J Forensic Sci 2006;51(1):120–4.

18. Fishbein MC, Wang T, Matijasevic M, Hong L, Apple FS. Myocardial tissue troponins T and I. An immunohistochemical study in experimental models of myocardial ischemia. Cardiovasc Pathol 2003;12(2):65–71.

19. Campobasso CP, Dell'Erba AS, Addante A, Zotti F, Marzullo A, Colonna MF. Sudden cardiac death and myocardial ischemia indicators: a comparative study of four immunohistochemical markers. Am J Forensic Med Pathol 2008;29(2):154–61.

20. Zhu BL, Ishikawa T, Michiue T, Li DR, Zhao D, Tanaka S, et al. Postmortem pericardial natriuretic peptides as markers of cardiac function in medico-legal autopsies. Int J Legal Med 2007;121(1):28–35.

Additional information and reprint requests:
Nicholas I. Batalis, M.D.
Medical University of South Carolina
Suite 309, 171 Ashley Avenue
PO Box 250908
Charleston, SC 29425
E-mail: batalini@musc.edu

TECHNICAL NOTE

PATHOLOGY AND BIOLOGY

Daniele Gibelli,[1] *M.D.; Alberto Brandone,*[2] *B.Sc.; Salvatore Andreola,*[3] *M.D.; Davide Porta,*[1] *B.Sc.;*
Elena Giudici,[1] *M.D.; Marco Aurelio Grandi,*[1] *M.D.; and Cristina Cattaneo,*[1] *Ph.D., M.D.*

Macroscopic, Microscopic, and Chemical Assessment of Gunshot Lesions on Decomposed Pig Skin

ABSTRACT: Very little literature exists on gunshot wounds on decomposed material. In this study, seven pig heads underwent a shooting test. Entrance wounds from the first head underwent neutron activation analysis (NAA) and histological testing immediately after the firing test; the other six heads were exposed to two different environments (open air and soil) and analyzed by radiochemical and histological tests every 15 days. Gunshot wounds in air maintained their morphological characteristics, and those in soil showed severe alteration after 5 weeks. Microscopic testing verified positive results for lead in all gunshot wounds in open air, whereas in most of those in soil lead could not be detected. Radiochemical analysis performed by NAA yielded for all gunshot wounds but one antimony quantities in the range of 0.07–13.89 μg. In conclusion, it may be possible to detect residues of antimony even in degraded tissues.

KEYWORDS: forensic science, forensic pathology, gunshot wounds, neutron activation analysis, sodium rhodizonate, decomposition, inhumation

Gunshot lesions, especially entrance wounds, and gunshot residues are a frequent object of routine macroscopic, microscopic, and chemical analysis, generally performed to determine useful information such as firing distance and manner of death (1–11). However, the large majority of research refers to bodies in a good state of preservation—putrefied or otherwise badly preserved samples are usually not considered for such analyses, mainly because it is assumed that histological and chemical information has been lost. In fact, very little research has been carried out on badly preserved samples (12–14).

Nonetheless, in cases of putrefied remains, an *a priori* evaluation of whether a lesion is actually a gunshot injury may be necessary. For example, in the presence of a severely decayed soft tissue lesion, it may at times be difficult to macroscopically diagnose it as a gunshot wound (as opposed to one produced by other weapons or feeding larvae). On the other hand, actual gunshot lesions may be altered by decomposition and/or entomological activity. The putrefactive alteration of skin often hides fouling, and decomposition alters the characteristics of seared blackened zones and stippling; insects often destroy bullet entrance wounds or create lesions similar to entrance wounds. Thus, it may be useful to perform histological and chemical testing to verify whether a lesion is in fact a gunshot wound.

Although some gunshot residue (GSR) detection techniques are comparatively old, such as neutron activation analysis (NAA) (15–17), neither these nor other more modern techniques have ever been tested on cases of badly preserved corpses, apart from one study on burnt animal bodies (18), where the application of NAA on charred gunshot wounds revealed a detectable antimony concentration; nor have microscopic methods been assessed.

The few articles that have considered the behavior of gunshot wounds on decomposed samples have shown a reasonably good conservation of morphological characteristics when decomposition processes are limited, whereas when they take place quickly, the macroscopic characteristics of the lesion are compromised (19,20). Such studies however did not include radiochemical or histological analyses and were limited to the macroscopic appearance of the wounds. Thus, verifying whether "decomposed" gunshot wounds still yield microscopic and chemical information would be extremely useful.

This study aims at carrying out an analysis of gunshot entrance wounds on pig heads that have undergone decomposition in two different environments (open air and soil) and comparing the alterations observed in the two different conditions. Macroscopic, radiochemical, and sodium rhodizonate tests were performed; NAA, instead of other methods such as FAAS, was included in the experimental design because of the ready availability of a nuclear reactor at one of the two universities involved in the study.

As concerns microscopic testing, a modification of the sodium rhodizonate technique (21) based on the addition of a 5% HCl solution was also assessed (22) to verify whether it could enhance the sodium rhodizonate reaction.

[1]LABANOF, Laboratorio di Antropologia e Odontologia Forense, Istituto di Medicina Legale e delle Assicurazioni, V. Mangiagalli 37, Università degli Studi di Milano, Italy.
[2]Dipartimento di Chimica, Viale Taramelli 12, Università di Pavia, Italy.
[3]Dipartimento di Anatomia Patologica, V. Venezian 1, Istituto Nazionale dei Tumori di Milano, Italy.

Materials and Methods

The heads of seven adult pigs (that died from causes independent from the experiment) were shot at a firing ground; the weapon chosen was a Franchi revolver .38 special, with a cartridge full metal jacket .38. Three entrance wounds, A (high on the forehead), B (between the eyes), and C (low on the nasal region), were produced on each head, for a total of 21 entrance wounds, all at a 5-cm distance and with the same weapon.

Immediately after the firing test, one head (sample 1) was selected as the positive control sample, and all three entrance wounds of the head were subjected to sampling in the following manner: (i) The skin around the entrance wound was divided into four equal parts, two cranial (left and right) and two caudal (left and right); (ii) the two cranial parts and one caudal (right) were swabbed for chemical analysis; (iii) the caudal left zone was not swabbed but was removed with a scalpel for histological analysis, and its perimeter was marked with china ink. Then it was soaked in 10% formalin. One negative skin control sample was taken from the neck of the first head before the firing test.

Three of the six heads were set in open air (n° 2, 4, and 6), and three buried in soil at a depth of 5 cm (n° 3, 5, and 7). Then, one entrance wound per environment (air and soil) was sampled every 15 days, as described earlier. In all, 18 wounds were analyzed in 16 weeks (Table 1). When decomposition was advanced, the swab was taken as usual, but sampling for histological analysis was limited to any soft tissue still preserved around the entrance wound. In the final decomposition phases, the swab was taken from soft tissue residues close to the bone surface.

Each entrance wound was analyzed with three methods: (i) macroscopic observation; (ii) radiochemical testing (NAA on swabs); and (iii) histological testing by sodium rhodizonate; in addition, an improvement of the sodium rhodizonate testing by adding a 5% HCl solution was attempted, as indicated in literature (22).

Macroscopic Analysis

The initial macroscopic analysis was performed to detect specific characteristics of gunshot entrance wounds, especially shape, edge symmetry, fouling, and powder stippling. Fouling refers specifically to the deposition of black powder around the entrance wound, which can be easily wiped away, whereas stippling consists of numerous reddish-brown to orange-red punctuate lesions surrounding the wound entrance, caused by unburnt black powder residues within the skin, or abrasions on the skin surface that occur when the skin is struck by gunpowder particles (8). The characteristics of each entrance wound were recorded, and a comparison was made between the samples in open air and the buried ones. The variations of the different characteristics with time and stage of decomposition were recorded as well.

TABLE 1—*Sampling periods and entrance wounds analyzed during the experiment.*

Week/Environment	Open Air	Soil
1 week	2B	3A
3 weeks	2C	3C
5 weeks	2A	3B
7 weeks	4C	5B
9 weeks	4B	5A
11 weeks	4A	5C
13 weeks	6B	7A
15 weeks	6C	7B
16 weeks	6A	7C

Radiochemical Analysis

Radiochemical analysis was carried out by NAA. The samples were introduced with the reference standards in a Triga Mark II (250 Kw) nuclear reactor of L.E.N.A. (University of Pavia); the neutronic flux was 1×10^{12} neutron/cm^2/sec. After irradiation, the samples decayed for 24 h to eliminate the short-life radioisotope interference, and then they were subjected to γ spectrometry by semiconductor crystal Ge/Li, linked to an analyzer and an electronic processor to measure 564 KeV radiation, typical of ^{122}Sb (counting time 1000 sec.).

Histological Analysis

Sampled entrance wounds were fixed in 10% formalin and paraffined; 5-μ sections were cut and then de-paraffined, hydrated by distilled water, and colored by hematoxylin-eosin. Then other histological sections of the same sample were treated with a 2.8-pH tartaric acid buffer solution and 0.2% sodium rhodizonate; test sensitivity threshold for lead is 0.1 μg. In this test, lead passes from Pb(0) to Pb(II) with formation of tetrahydroxy chinone via a redox reaction and turns scarlet red. Finally, another group of histological sections was further treated with a 5% HCl solution, which should enhance lead detectability by turning blue.

Temperature, humidity, and rainfall were recorded during January, February, March, April, and May 2005 by the meteorological observatory in Milan.

Results

Macroscopic Analysis

In the positive control (sample n° 1), gunshot entrance wounds showed a regular circular shape, with an easily identifiable surrounding sooty area; powder stippling was visible as gray dotting around the gunshot entrance.

The samples in open air during the 16-week experimental period appeared well preserved; the skin turned brown and became tougher, as if mummified (Fig. 1). Samples in soil on the other hand preserved their macroscopic characteristics only for the first weeks; from the 5th week, the putrefaction processes led to degradation of the samples in soil, which made the lesions almost impossible to identify as gunshot wounds (Fig. 2).

On the samples in open air, fouling showed a quick decrease with time but could always be identified. Symmetry of the entrance wound and powder stippling also were easily detectable during the 16 weeks and did not change considerably (Fig. 1).

On the entrance wounds of samples in soil, a more advanced decomposition process took place; fouling and powder stippling quickly became invisible after 3 weeks. Entrance wounds could clearly be identified for the first 8 weeks; then the lesions could no longer be identified, and swabs were therefore made from tissue residues in the area where the gunshot wound could be seen on the underlying bone (Fig. 2).

Radiochemical Analysis

The sample analyzed immediately after the firing test presented antimony concentrations higher than for all the other entrance wounds, as expected, while the negative control sample on skin not affected by a gunshot lesion showed no antimony. The most interesting information obtained was the persistence in all air and most soil samples of antimony (concentrations were higher in air than in
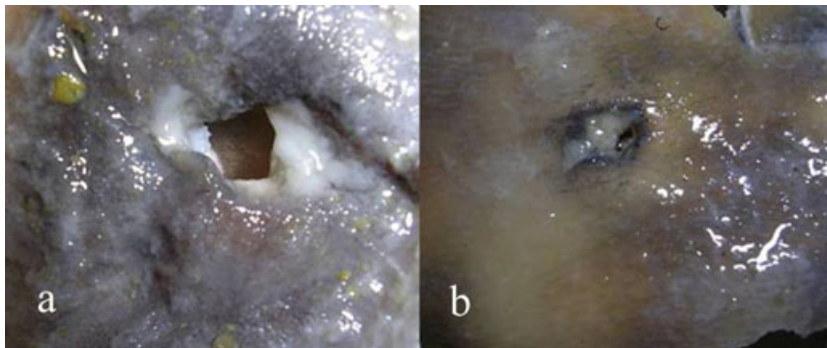
FIG. 1—*Entrance wounds from samples in open air: (a) entrance wound 2C analyzed at the 3rd week; (b) entrance wound 6B analyzed at the 13th week.*

FIG. 2—*Entrance wounds from buried samples: (a) entrance wound 3C analyzed at the 3rd week; (b) entrance wound 3B analyzed at the 5th week.*
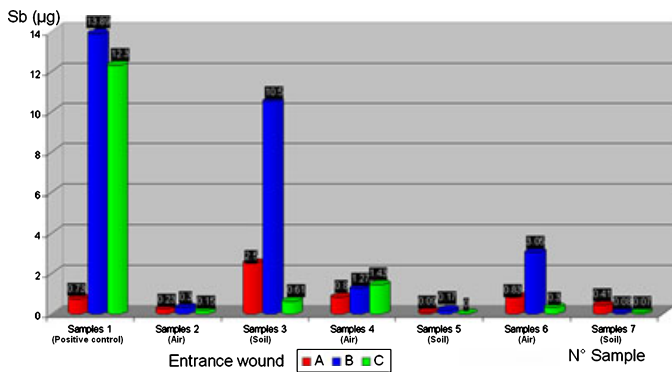

FIG. 3—*Antimony concentrations recorded in entrance wounds analyzed by neutron activation analysis (NAA).*

TABLE 2—*Degree of tattooing found after HE coloration and lead positivity after sodium rhodizonate and treatment by 5% HCl solution in different samples; NV indicates a nonlegible result.*

	HE	SR	SR + 5% HCl solution
1A	+	NV	NV
1B	+++	++	+
1C	+++	+/−	−
2A	+	+	+
2B	++	+/−	−
2C	++	+++	+
3A	+	++	+
3B	NV	+	+
3C	NV	++	+
4A	+++	++	+
4B	++	+	+
4C	+++	+++	+
5A	NV	NV	NV
5B	+/−	NV	NV
5C	NV	NV	NV
6A	++	++	+
6B	++	++	+
6C	+++	+++	+
7A	NV	NV	NV
7B	NV	NV	NV
7C	NV	NV	NV

soil). Samples on all but one of the entrance wounds (entrance wound C on pig head n° 5) showed antimony quantities in the range of 0.07–13.89 μg (Fig. 3).

Higher levels were seen in the open air samples, but residues could also be detected in the buried samples.

Histological Analysis

Use of the hematoxylin-eosin stain on histological sections allowed us to microscopically observe the entrance wounds, as shown in Table 2; stippling was visible in all the three entrance wounds of head n° 1 (the positive control test), fairly numerous in head n° 2 (especially in entrance B and C), 4, and 6, and poor in heads n° 3, 5; at times, the presence of amorphous environmental material (especially fungi and bacteria) prevented the authors from observing useful details. Head n° 7, in soil, was in such an advanced state of decomposition that no soft tissue traits could be identified histologically.

Lesions appeared as clusters of amorphous black material adhering to the entrance wound and were less visible in tissues crossed by the bullet. Burning signs on skin were also visible as protein agglutination on the skin surface, caused by heat emitted by the gunshot, darker than the surrounding areas (Fig. 4). Sodium rhodizonate testing and the analysis of thin sections on the same samples provided more information: highly positive results for lead occurred on head n° 1 (the positive control sample) and on all the heads in open air (Fig. 5): in all samples under soil, except for head n° 3, where all three entrance wounds were positive, no sign of lead was visible, probably because of contamination and advanced decomposition (Fig. 6). GSR appeared as red scarlet clusters assembled around the entrance wound. Results are showed in Table 2.

With subsequent treatment with HCl, the lead clusters acquired the typical blue coloration and appeared more limited in size with respect to the pattern shown by sodium rhodizonate; however, the position of the lead clusters was the same in both groups, and this seems to confirm the greater specificity of the sodium rhodizonate test when completed by adding a 5% HCl solution, although further testing is needed to exclude technical artifacts.

Discussion

In cases of badly preserved corpses, the actual identification of a gunshot wound may be difficult to perform macroscopically. Chemical and microscopic tests could reveal useful gunshot residue markers, but this issue has not been explored up to now.
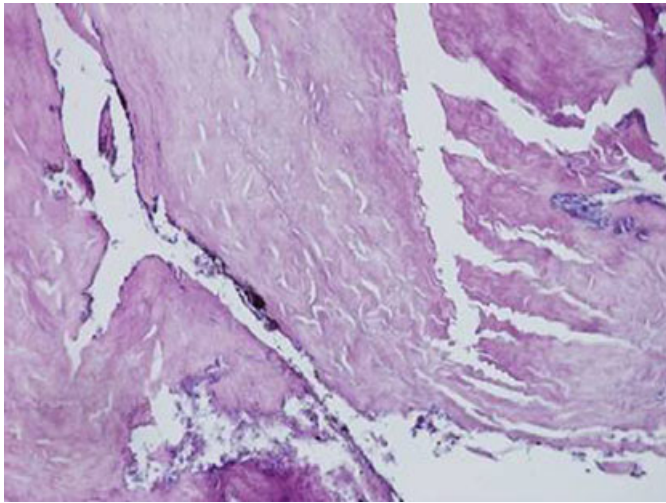

FIG. 4—*Histologic analysis of entrance wound 2C (in open air) analyzed at the 3rd week, HE stain (100×).*

Previous studies (19,20) have pointed out that contact gunshot wounds and those at intermediate range do not radically change in their morphological appearance with the early stages of decomposition but do lose their characteristics when tissue becomes degraded in the late stages of decomposition. In this study, similar results were obtained from a macroscopic point of view.

All samples in open air remained fairly well preserved during the entire period of the experiment: this may be because the skin underwent mummification, which seemed to lead to a good preservation of lesion morphology. Stippling persisted for a longer time than fouling (this was expected because fouling is a powder deposit on skin, whereas powder stippling is produced by the impact of metallic particles on tissues).

Morphological characteristics of the entrance wounds did not change considerably, and the shape and symmetry of entrance wounds could be identified during the entire experimental period. Fouling and powder stippling showed a slight decrease in time, but both could be identified up to the 16th week.

The buried samples underwent a different destiny: up to the 3rd week, the entrance wound shape and symmetry could still be recognized, as for the samples in open air, although fouling and powder stippling were no longer visible. At week 5, tissue consistency decreased, although the entrance wound could still be recognized; at week 7, the gunshot wounds changed morphology drastically, and from week 9 onwards, they could no longer be identified. The skin then underwent severe degradation. Sampling was limited to the remaining tissue fragments, and during the last weeks of the experiment (the 15th and 16th), the heads were almost entirely skeletonized.

Thus, advanced decomposition, such as that of the samples in soil, can modify gunshot wound morphology. This is to be expected and confirms that macroscopic observation can "safely" identify gunshot wounds only in cases of limited decomposition, which is why the study aimed at attempting to verify whether in degraded samples chemical and microscopic information can still be detected.

More surprisingly and importantly, the radiochemical testing pointed out that antimony can still be detected both in samples in open air and in buried ones, as well as in degraded tissues. Even in the last buried samples, which had almost reached skeletonization, NAA tests resulted positive for antimony.

The significance of this finding however needs to be discussed particularly with reference to the environmental threshold level. Some authors have set their threshold levels according to individual experimental studies. Ruch et al. (23) for example applied NAA to 130 control skin samples that had not been exposed to gunshot residues: the environmental antimony concentration threshold found was between 0.01 and 0.03 μg (23). On the other hand, Reed et al.
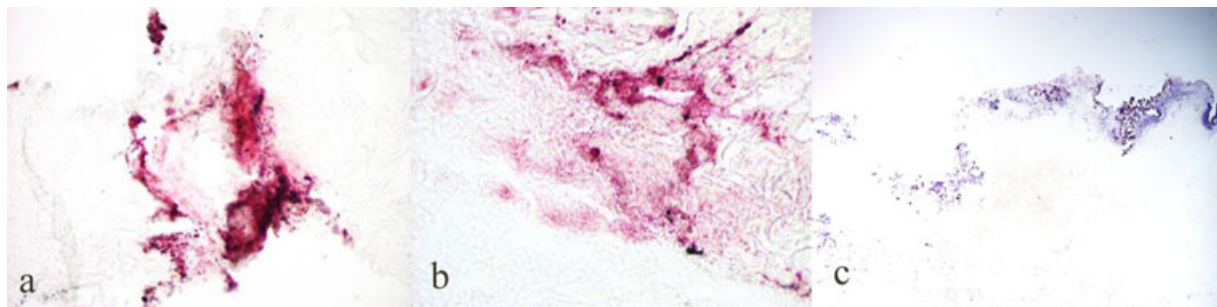

FIG. 5—*Section treated with sodium rhodizonate from entrance wounds in open air: (a) entrance wound 2C analyzed at the 3rd week (100×); (b) entrance wound 6C analyzed at the 15th week (200×); (c) section from entrance wound 2C analyzed at the 3rd week after adding 5% HCl solution (100×).*
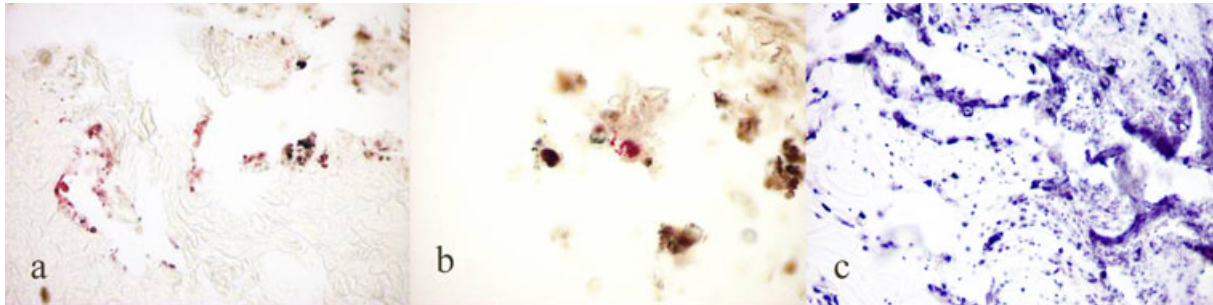
FIG. 6—*Skin section treated with sodium rhodizonate from entrance wounds in buried samples: (a) section from entrance wound 3B at the 5th week (400×); (b) section from entrance wound 3C analyzed at the 3rd week (600×); one can observe contamination by bacteria and fungi; (c) section from entrance wound 3C analyzed at the 3rd week after treatment by adding 5% HCl solution (100×).*

(24) mention higher threshold levels. In fact, they refer to the USA-CIL (US Army Criminal Investigation Laboratory) indications that identify 0.2 μg of antimony as the threshold value for a positive identification of gunshot residue. Observing the results from these two perspectives, if we refer to Ruch et al., all samples (but one that was negative) could be considered positive. If we consider USACIL standards (as reported by literature [24]), all but six samples (Fig. 3) were above this threshold. Regardless of the threshold chosen, most samples (even the ones buried in soil) gave results above 0.2 μg., and the negative control showed no antimony.

Thus, the detection of antimony on skeletonized samples after several weeks suggests that gunshot residues may persist for a long period in degraded soft tissues. It has been reported that gunshot residues, in particular the antimony content, when measured by NAA, radically decrease after 9–12 h after the gunshot explosion (24). Nonetheless, the present results indicate that even decomposed samples may yield chemical markers, which may help orientate toward a diagnosis of a gunshot wound. Given the variability of preservation because of the decomposition processes, a quantitative analysis for an estimation of firing distance should not be considered reliable.

On average, antimony concentrations were higher in samples in air, although with some random behavior. It is difficult to find a single explanation for such a phenomenon. The decomposition processes, faster in samples under soil than in open air, may have caused a decrease in antimony content through the disruption of soft tissues.

In both sets (air and soil), antimony in general decreased progressively with time; this trend was more evident in buried samples, while it was less evident in samples in open air, probably because of the numerous variables in the air affecting the sample. One could hypothesize that soil may have provided "protection" or less dispersion in time of gunshot residues, whereas open weather conditions recorded during the experiment (rainfall, winds, etc.) may have affected antimony concentration of gunshot wounds on samples in open air more severely. However, the limited nature of the experiment prevents us from identifying the actual variables affecting such behavior.

Concerning differences in antimony contents between the three different wounds (A, B, and C), no simple explanation can be given. Wound B (between the eyes) frequently showed higher antimony contents with respect to wounds A and C, whereas wound C antimony content was generally lower, both in open air and in soil. Wound B—situated between A and C—may have been affected by peripheral contamination from both gunshot wounds A and C, with an overall increase in antimony content. On the other hand, gunshot wound C in the nasal area often caused a comminuted fracture of the thinner nasal bones: the larger damage of soft and hard tissues

may have caused a spreading of residues, with consequent decrease in antimony concentration.

As concerns the microscopic analyses, the sodium rhodizonate test was performed to verify its applicability to decomposed material. Results pointed out that, whereas samples in open air showed a high concentration of GSR, only head n° 3 among those buried in soil (the first to be analyzed during the experimental period) was positive; in all other entrance wounds, gunshot residues were no longer visible probably because of the high degree of decomposition, particularly after the 5th–7th week. Although the addition of 5% HCl solution may enhance the specificity of sodium rhodizonate, the sodium rhodizonate test did not prove to be a reliable test for identification of a gunshot wound in decomposed remains, as it seems largely influenced by decomposition and environmental variables.

In our study, the sodium rhodizonate test was performed only to verify the presence of lead around the entrance wound as a simple "marker," which could confirm that the lesion was indeed a gunshot wound. No implications concerning gunshot range were considered, given that conventional thought is that such patterns cannot aid in range determinations (8).

This study has tried to shed some light on the issue concerning the identification of gunshot wounds and detection of gunshot residues on decomposed skin. Three methods of analysis were taken into consideration: classical macroscopic observation, radiochemical analysis, and sodium rhodizonate microscopic testing. Of all these, NAA proved to provide the most useful information concerning the diagnosis of gunshot wounds from suspect lesions on decomposed skin.

Regardless of these encouraging results, more research needs to be performed because, to our knowledge, this is the first systematic chemical and microscopic study on decomposed gunshot wounds.

References

1. Krishnan SS. Determination of gunshot firing distances and identification of bullet holes by neutron activation analysis. J Forensic Sci 1967;12(1):112–22.
2. Krishnan SS, Gillespie KA, Anderson EJ. Rapid detection of firearm discharge residues by atomic absorption and neutron activation analysis. J Forensic Sci 1971;16(2):144–51.
3. White RS, Owens AD. Automation of gunshot residue detection and analysis by scanning electron microscopy/energy dispersive X-ray analysis (SEM/EDX). J Forensic Sci 1983;32(6):1595–603.

4. Brazeau J, Wong RK. Analysis of gunshot residue on human tissues and clothing by X-ray microfluorescence. J Forensic Sci 1997;42(3):424–8.
5. Lebiedzik J, Johnson DL. Rapid search and quantitative analysis of gunshot residue particles in the SEM. J Forensic Sci 2000;45(1):83–92.
6. Stein KM, Bahner ML, Merkel J, Ain S, Mattern R. Detection of gunshot residues in routine CTs. Int J Legal Med 2000;114(1-2):15–8.
7. Adelson L. A microscopic study of dermal gunshot wounds. Am J Clin Pathol 1961;35:393–402.
8. Di Maio VJM. Gunshot wounds. Boca Raton, FL: CRC Press, 1999.
9. Schwoeble AJ, Exline DL. Current methods in forensic gunshot residue analysis, 1st edn. Boca Raton, FL: CRC Press, 2000.
10. Wolten GM, Nesbitt RS. On the mechanism of gunshot residue particle formation. J Forensic Sci 1980;25(3):533–45.
11. Basu S. Formation of gunshot residue. J Forensic Sci 1982;27(1):72–91.
12. Glattstein B, Zeichner A, Vinokurov A, Levin N, Kugel C, Hiss J. Improved method for shooting distance estimation. Part 3. Bullet holes in cadavers. J Forensic Sci 2000;45(6):1243–9.
13. Gibelli D. Alterazioni morfologiche e chimiche di lesioni da CAF su materiale decomposto: studio pilota su modello animale [graduate thesis]. Milan, Italy: Faculty of Medicine, Università degli Studi di Milano, 2004-05.
14. Gibelli D, Brandone A, Andreola S, Poppa P, Porta D, Grandi MA, et al. Alterazioni morfologiche e chimiche di lesioni da colpo d'arma da fuoco su materiale decomposto: studio pilota su modello animale. Riv It Med Leg 2008;1:215–25.
15. Krishnan SS, Nichol RC. Identification of bullet holes by neutron activation analysis and autoradiography. J Forensic Sci 1968;13(4):518–27.
16. McFarland RC, McLain ME. Rapid neutron activation analysis for gunshot residue determination. J Forensic Sci 1973;18(3):226–31.
17. Rudzitis E, Kopina M, Wahlgren M. Optimization of firearm residue detection by neutron activation analysis. J Forensic Sci 1973;18(2):93–100.
18. Cattaneo C, Davide D, Mazzucchi A, Gibelli D, Pasquale P, Brandone A, et al. Morphological alterations of blunt, sharp force and gunshot lesions in bone during the charring process. Am J Forensic Med Pathol 2010. In press.
19. MacAulay L, Barr DG, Strongman DB. Effects of decomposition on gunshot wound characteristics: under cold temperatures with no insect activity. J Forensic Sci 2009;54(2):448–51.
20. MacAulay L, Barr DG, Strongman DB. Effects of decomposition on gunshot wound characteristics: under moderate temperatures with insect activity. J Forensic Sci 2009;54(2):443–7.
21. Marty W, Sigrist T, Wyler D. Determination of firing distance using the rhodizonate staining techniques. Int J Legal Med 2002;116(1):1–4.
22. Bartsch MR, Hilton JK, Wainwright KP. An update on the use of the sodium rhodizonate test for the detection of lead originating from the firearm discharges. J Forensic Sci 1996;41(6):1046–51.
23. Ruch RR, Buchanan JD, Guinn VP, Bellanca SC, Pinker RH. Neutron activation analysis in scientific crime detection. J Forensic Sci 1964;9:119–32.
24. Reed GE, McGuire PJ, Bohem A. Analysis of gunshot residue test results in 112 suicides. J Forensic Sci 1990;35:62–8.

Additional information and reprint requests:
Cristina Cattaneo, Ph.D., M.D.
Professor
LABANOF, Laboratorio di Antropologia ed Odontologia Forense
Istituto di Medicina Legale e delle Assicurazioni
Università degli Studi di Milano
V. Mangiagalli 37
Milan
Italy
E-mail: cristina.cattaneo@unimi.it

TECHNICAL NOTE

PATHOLOGY AND BIOLOGY

Ma. Teresa Valdes-Perezgasga,[1] *Ph.D.; Francisco Javier Sanchez-Ramos,*[1] *Ph.D.;*
Oswaldo Garcia-Martinez,,[2] *Ph.D.; and Gail S. Anderson,*[3] *Ph.D.*

Arthropods of Forensic Importance on Pig Carrion in the Coahuilan Semidesert, Mexico

ABSTRACT: This is the first report of an ongoing research to establish a sarcosaprophagous arthropod database in the Coahuilan semidesert. Seven pigs (*Sus scrofa* L.) were used as human models to determine succession in an open urban area during the 2007 winter–spring period. Arthropods were collected manually and from pitfall traps. Carcass biomass loss, as well as arthropod colonization, was recorded during 71 days postmortem. Five decomposition stages were identified during which most abundant orders were found to be Diptera, Coleoptera, and Hymenoptera. *Lucilia sericata* (Meigen), *Chrysomya rufifacies* (Macquart), *Necrobia rufipes* (DeGeer), *Dermestes maculatus* (DeGeer), *Pheidole hyatti* Emery, and *Pogonomyrmex rugosus* Emery stood out as dominant species.

KEYWORDS: forensic science, forensic entomology, arthropods, succession, semidesert, pigs, Mexico

Arthropods are among the first and more important invertebrates to arrive and colonize a cadaver whether animal or human (1–3). Carrion is a food resource for a diversity of species and maintains a large and changing fauna as it decomposes. Each stage of decomposition attracts a different group of sarcosaprophagous arthropods, especially insects. Some are attracted to the remains, which are used as a medium for oviposition or feeding, while others are attracted by the aggregation of other arthropods that are used as a food source (4).

Sarcosaprophagous arthropods follow a predictable succession sequence in animal remains (5–7), and when this sequence is known for a given area and circumstances, an analysis of the arthropod fauna on the remains can be used to estimate the time since death (8,9). The relevance of a database on arthropod succession on carrion using results obtained from decomposition studies depends on the kind of experimental model used. Pigs (*Sus scrofa* L.) seem to be the best experimental models for calculating the postmortem interval (PMI) (10–12).

The main objective of this research was to determine the arthropod fauna of forensic importance that colonize carrion in a semidesert urban zone of Coahuila, Mexico.

Materials and Methods

This report constitutes a part of ongoing research carried out during the winter and spring of 2007 in the research station (urban area) of the Universidad Autonoma Agraria Antonio Narro, in Torreon Coahuila, Mexico. The site (GPS 25° 33′ 25″ N, 103° 21′ 57″

W) had been recently plowed, so it was devoid of vegetation. Characteristic climate is dry (R.H. ≤30%) with a mean temperature of 25°C and 220 mm of precipitation per annum.

Seven 22-kg pigs were killed *in situ* with a stab to the heart to be used as experimental models to simulate human cadaver decomposition on February 19, 2007, with this date as day zero. Each pig was placed in a 3/8″ steel rod cage 1.2 × 0.8 × 0.5 m covered with iron mesh. In the bottom of each cage, a mesh stretcher was placed to facilitate handling the pigs. Each cage was staked to the ground by a 0.6-m-long steel rod and was surrounded by a wooden fence (2.5 × 2.5 m) to prevent the entry of birds and small scavengers. Pigs were placed 50 m apart so that olfactory orientation by arthropods was independent (13).

Experimental models were divided into three groups. Group 1 included four pig carcasses, which were used for arthropod collection on, in, and under them. Four pitfall traps were placed around each carcass where arthropods that were attracted were collected. Group 2 consisted of two carcasses, which were used to record weight loss over time using an electronic scale (Revuelta H5-30K). The soil beneath each carcass in this group was sampled (250 cc) from a soil profile of 0–0.2 m. These soil samples were processed in separate Berlese funnels for 24 h to determine the soil fauna over time. Group 3 consisted of one carcass. This was labeled as control and was not disturbed at all, serving to determine whether the disturbance of sampling affected the process of decomposition or arthropod colonization.

During the first 3 weeks after death, the carcasses were visited on a daily basis. During each visit, arthropod specimens were collected from group 1 and placed in 70% ethanol. Half the larvae collected were preserved in Khale's solution, and the other half were reared to the adult stage using beef liver in the laboratory at room temperature for identification. Maximum and minimum temperatures and precipitation were registered daily during 70 days postmortem. After the third week, carcasses were visited every other day.

[1]Posgrado en Ciencias Agrarias UAAAN-UL, Periferico Raul Lopez Sanchez km 2, Torreon, Coahuila, CP 27059, Mexcio.
[2]Calzada Antonio Narro 1923, Col. Buenavista, Saltillo, Coahuila, CP 25315, Mexcio.
[3]School of Criminology, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada.

Soil samples from under carcasses in group 2 were collected in 70% ethanol, observed under stereomicroscope, counted, and identified.

Changes during the decomposition process as well as arthropod abundance and activity on carcasses were recorded over time. A photographic record and a detailed written ledger were maintained. All specimens collected were grouped by order and family. Specimens were identified at the laboratory, and some specimens were sent to specialists in certain taxa in Mexico.

Results and Discussion

Decomposition is a continuous process; however, to facilitate comprehension of results, this can be separated into stages. Five stages of decomposition were recorded, in accordance with what several other researchers have stated with slight modifications (4,13–15).

Fresh Stage (0–1 Days After Death)

Since the moment of death, adult flies belonging to the families Calliphoridae, Sarcophagidae, and Muscidae were noted overflying and on the carcasses. In this stage, no putrid odors were perceived, although the day was warm.

Bloated Stage (2–4 Days)

The carcasses were completely bloated and presented extruded anuses because of the gases produced by decomposition. During this stage, discoloration of the skin was noticeable, as well as a strong putrid odor. Abundant calliphorid flies arrived, and small maggots were present in natural orifices. Adult piophilid flies arrived at the carcasses, as well as clerid beetles (*Necrobia rufipes* Macquart), which fed on maggots (4 days after death) inside the ears of the carcasses. Under the carcasses, there were numerous adult cockroaches (*Blattella* sp.), isopods, spiders, and dermestids.

Active Decay Stage (5–13 Days)

Odor was less intense although putrid, with significant liquid under the carcasses where numerous maggots, isopods, cockroaches, earwigs, dermestid, and clerid beetles were found. Inside ocular cavities, nose, and abdomen of the carcasses, large masses of maggots were evident. Maggot migration from the carcasses was recorded at 8 days postmortem. No cockroaches were collected under the carcasses at the end of this stage.

Advanced Decay Stage (14–29 Days)

Fetid odors were less intense, although some fetid liquid was still present under the carcasses. Dry skin was starting to detach from the bodies. Carcasses were completely dehydrated although an oily material covered them. Weight loss at this stage was more gradual than in the previous stages. Emergence of Calliphoridae, Sarcophagidae, and Muscidae started at 14 days after death, attracting black birds (*Quiscalus mexicanus* Gmelin) that consumed mostly tenerals. Piophilidae larvae were collected at this stage. Increasing numbers of clerids, dermestids, and ants were evident.

Dry Remains (30–70 Days)

This stage was characterized by very dry skin with bone disarticulation. Odor emanating from carcasses was like rancid fat, and almost no moisture underneath. The oily residue under the carcasses was solid at the carcass–soil interphase, where abundant pupal cases parasitized by Pteromalidae wasps were collected. Sarcophagidae and Piophilidae adult emergence continued throughout this stage. As this stage progressed, fewer specimens of Staphylinidae, Dermestidae, Cleridae, Anthocoridae, Formicidae, Soliphugae, Isopoda, and Acari were collected. Opportunistic individuals were recovered from the pitfall traps such as lizards (*Podarcis* sp.). The end of this stage was marked by reduced insect activity and no increased biomass loss.

The immediate arrival of blow flies to carrion agrees with findings of other studies (4,10,13,16).

The most abundant forensically important insect species are presented in Table 1. Our findings agree with several studies in other regions where calliphorids (16–18) and sarcophagids (16,19–21) have been mentioned as primary colonizers of carrion.

Lucilia sericata has been mentioned as a dominant species on carrion in various regions of the world (17,18,22–28), although other imported Calliphoridae genera, such as *Chrysomya*, have been

TABLE 1—*Most abundant insects of forensic importance collected from pig carcasses in an open urban area of the Coahuilan semidesert.*

Order	Family	Genus and Species	Insect Stage	Stage of Decomposition
Diptera	Sarcophagide	*Sarcodexia* sp.	L, A	2, 3, 4, 5
		Tytanogrypa sp.	L, A	2, 3, 4, 5
		Neobelleria sp.	L, A	2, 3, 4, 5
		Liopygia sp.	L, A	2, 3, 4, 5
		Bercaea sp.	L, A	2, 3, 4, 5
		Bellieria sp.	L, A	2, 3, 4, 5
		Paraphrissopoda sp.	L	3
		Anicia sp.	L	3
	Calliphoridae	*Lucilia sericata* (Meigen)	L, A	2, 3, 4, 5
		Chrysomya rufifacies (Macquart)	L, A	2, 3, 4, 5
		Lucilia silvarum (Meigen)	L, A	3, 5
		L. cuprina (Wiedemann)	L, A	3, 4, 5
		C. megacephala (Fabricius)	L, A	3, 4, 5
Coleoptera	Cleridae	*Necrobia rufipes* (De Geer)	A	2, 3, 4, 5
	Dermestidae	*Dermestes maculatus* (De Geer)	L, A	2, 3, 4, 5
		Dermestes sp.	L, A	2, 3, 4, 5
Hymenoptera	Formicidae	*Pogonomyrmex rugosus* Emery	A	4, 5
		Pheidole hyatti Emery	A	3, 4, 5
		Camponotus festinatus (Buckley)	A	3, 4, 5

Note: A = adults, L = larvae, 2 = Bloat, 3 = Active decay, 4 = Advanced decay, 5 = Dry remains.

reported as a very important and competitive component of sarcos-aprophagous community in several countries of the American continent since its introduction (29–32).

Piophilid flies were the most abundant dipteran individuals collected from early stages of decomposition; although larvae were not collected until 18 days postmortem, results agree with findings in other regions where they have been categorized as early colonizers of carrion (13,16,33,34).

Necrobia rufipes, Dermestes maculatus, and *Dermestes* sp. were the dominant species in Coleoptera, being collected since the bloat stage, as has been reported in previous studies (13,25,35,36). In our study, few adult dermestid beetles were collected as early as 3 days postmortem, similar to what was reported in Hawai'i (37,38), although larvae were not collected until later (38). Dermestids and clerids were an important component of entomological inhabitants on carrion as in other parts of the world (39).

From the active decay stage until dry remains, ants were collected from the remains, feeding on the carcasses and on other arthropods (mainly maggots), in accordance with the findings of Castillo (25), Martinez-Sanchez et al. (23), and Stoker et al. (40) who categorize them as an important component of the sarcosaprophagous community.

As with previous successional studies, the presence and abundance of mites were restricted to the dry remains stage (41–43).

Biomass Loss

Biomass loss was dramatic during the first 15 days after death, with carcasses losing 20% of their weight 7 days after death (beginning of active decay stage), recording 50% loss of body weight at the end of this stage. Tissue decomposition and water loss during the first three decay stages and arthropod colonization by sarcosaprophagous insects (mainly Calliphoridae and Sarcophagidae) occurred during the first three stages where body weight was greatest.

All carcasses decomposed at the same rate and were colonized by the same insect species, confirming that sampling and carcass manipulation for weight loss recording did not influence insect colonization. These findings agree with the results of studies where disturbed and undisturbed rat (44) and rabbit carcasses (45) presented no differences on overall decomposition stage.

In conclusion, insects colonized the remains in a predictable sequence, as has been reported many times in the literature; however, species and colonization times were specific for this type of habitat and region.

References

1. Nuorteva S. Sarcosaprophagous insects as forensic indicators. In: Tedeschi CG, Eckert WG, Tedeschi LG, editors. Forensic medicine: a study in trauma and environmental hazards. Vol II. Philadelphia, PA: Saunders, 1977;1072–95.
2. Smith KGV. The faunal succession of insects and other invertebrates on a dead fox. Entomol Gaz 1975;26:277.
3. Erzinçlioglu YZ. The application of entomology to forensic medicine. Med Sci Law 1983;25:228–30.
4. Payne JA. A summer carrion study of the baby pig *Sus scrofa* Linnaeus. Ecology 1965;46:592–602.
5. Bornemissza GF. An analysis of arthropod succession in carrion and the effect of its decomposition on the soil fauna. Aust J Zool 1957;5:1–12.
6. Payne JA, King EW, Beinhart G. Arthropod succession and decomposition of buried pigs. Nature 1968;219:1180–1.
7. Jirón LF, Cartín VM. Insect succession in the decomposition of a mammal in Costa Rica. J New York Entomol Soc 1981;89:158–65.
8. Kashyap VK, Pillay VV. Efficacy of entomological methods in estimation of postmortem interval: a comparative analysis. Forensic Sci Int 1989;40:245–50.
9. Benecke M. Six forensic entomology cases: description and commentary. J Forensic Sci 1998;43:797–805.
10. Rodriguez WC. Decomposition of buried and submerged bodies. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton: CRC, 2002;459–68.
11. Schoenly KG, Haskell NH, Hall R. Testing the reliability of an animal model for use in research and training programs in forensic entomology. Washington, DC: US National Institute of Justice, Final Report, Grant No. 94-IJ-CX-0039, 1996.
12. Haskell NH, Schoenly KG, Hall RD. Testing reliability of animal models in research and training programs in forensic entomology, part II. Washington, DC: US National Institute of Justice, Final Report, Grant No. 97-IJ-CX-0046, 2001.
13. Anderson GS, VanLaerhoven SL. Initial studies on insect succession on carrion in southwestern British Columbia. J Forensic Sci 1996;41:617–25.
14. Early M, Goff ML. Arthropod succession patterns in exposed carrion on the island of Oahu, Hawaii. J Med Entomol 1986;23:520–31.
15. Tullis K, Goff ML. Arthopod succession in exposed carrion in a tropical rainforest on Oahu, Hawaii. J Med Entomol 1987;24:332–9.
16. Sharanowski BJ, Walker EG, Anderson GS. Insect succession and decomposition patterns on shaded and sunlit carrion in Saskatchewan in three different seasons. Forensic Sci Int 2008;179:219–40.
17. Campobasso CP, Di Vella G, Introna F. Factors affecting decomposition and Diptera colonization. Forensic Sci Int 2001;120:18–27.
18. Watson EJ, Carlton CE. Spring succession of necrophilous insects on wildlife carcasses in Louisiana. J Med Entomol 2003;40:338–47.
19. Salviano RJB. Sucessão de Diptera Caliptrata em carcaça de Sus scrofa L. [dissertation]. Rio de Janeiro (Brasil): Univ. Fed. Rural do Rio de Janeiro, 1996.
20. Tantawi T, El-Kady EM, Greenberg B, El-Ghaffar HA. Arthropod succession on exposed rabbit carrion in Alexandria, Egypt. J Med Entomol 1996;33:566–80.
21. de Barros RM, Mello-Patiu CA, Pujol-Luz JR. Sarcophagidae (Insecta, Diptera) associados á decomposição de carcaças de *Sus scrofa* Linnaeus (Suidae) em área de Cerrado do Distrito Federal, Brasil. Revista Brasileira de Entomologia 2008;52:606–9.
22. Tessmer JW, Meek CL, Wright VL. Circadian patterns of oviposition by necrophilous flies (Diptera:Calliphoridae) in Southern Louisiana. Southwest Entomol 1995;20:439–45.
23. Martinez-Sanchez A, Rojo S, Marcos-Garcia MA. Annual and spatial activity of dung flies and carrion in a Mediterranean holm-oak pasture ecosystem. Med Vet Entomol 2002;14:56–63.
24. Oliva A. Insects of forensic significance in Argentina. Forensic Sci Int 2001;120:145–54.
25. Castillo MM. Estudio de la entomofauna asociada a cadaveres en el alto Aragon (España). Monografias de la Soc Entomol Aragonesa 2002;6:9–94.
26. Grassberg M, Frank C. Initial study of arthropod succession on pig carrion in a Central European urban habitat. J Med Entomol 2004;41:511–23.
27. Hwang C, Turner BD. Spatial and temporal variability of necrophagous Diptera from urban to rural areas. Med Vet Entomol 2005;19:379–91.
28. Vanin S, Tasinato P, Ducolin G, Terranova C, Zancaner S, Montisci M, et al. Use of *Lucilia* species for forensic investigations in Southern Europe. Forensic Sci Int 2007;177:37–41.
29. de Carvalho CJB, Thyssen PS, Linhares AX, Palhares FAB. A checklist of arthropods associated with pig carrion and human corpses in Southeastern Brazil. Mem do Inst Oswaldo Cruz 2000;95:135–8.
30. Mavarez-Cardozo MG, Espina de Fereira AI, Barrios-Ferrer FA, Ferreira-Paz JL. La entomología forense y el Neotrópico. Cuad Med Forense 2005;11:23–33.
31. Aguiar-Coelho VM, Milward-de-Azevedo EMV. Combined rearing of *Cochliomyia macellaria* (Fabr.), *Chrysomya megacephala* (Fabr.) and *Chrysomya albiceps* (Wied.) (Dipt., Calliphoridae) under laboratory conditions. J Appl Ent 1998;122:551–4.
32. Gruner SV, Slone DH, Capinera JL. Forensically important Calliphoridae (Diptera) associated with pig carrion in rural North-Central Florida. J Med Entomol 2007;44:509–15.
33. DeSouza AM, Linhares AX. Diptera and Coleoptera of potential forensic importance in southeastern Brazil: relative abundance and seasonality. Med Vet Entomol 1997;11:8–12.
34. Gill GJ. Decomposition and arthropod succession on above ground pig carrion in rural Manitoba. Ottawa (Ontario): Canadian Police Research Centre, Technical Report TR-06-2005, 2005.

35. Iannacone J. Artropofauna de importancia forense en un cadaver de cerdo en el Callao, Peru. Rev Bras de Zool 2003;20:85–90.

36. Arnaldos MI, García MD, Romera E, Presa JJ, Luna A. Estimation of postmortem interval in real cases on experimentally obtained entomological evidence. Forensic Sci Int 2005;149:57–65.

37. Early M, Goff ML. Arthropod succession patterns in exposed carrion on the island of O'ahu, Hawai'i. J Med Entomol 1987;23:520–31.

38. Hewadikaram KA, Goff ML. Effect of carcass size on rate of decomposition and arthropod succession patterns. Am J Foren Med Pathol 1991;12:235–40.

39. Kulsreshta P, Satpathy DK. Use of beetles in forensic entomology. Forensic Sci Int 2001;120:15–7.

40. Stoker RL, Grant WE, Vinson SB. *Solenopsis invicta* (Hymenoptera: Formicidae) effect on invertebrate decomposers of carrion in central Texas. Environ Entomol 1995;24:817–22.

41. Battan HM, Arnaldos MI, Rosso B, Garcia MD. Estudio preliminar de la comunidad sarcosaprofaga en Cordoba (Argentina): aplicacion de la entomologia forense. Anales de Biologia 2005;27:191–201.

42. Camacho GPC. Sucesion de la entomofauna cadaverica y ciclo vital de *Calliphora vicina* (Diptera: Callphoridae) como primera especie colonizadora, utilizando cerdo blanco (*Sus scrofa*) en Bogota. Rev Colombiana de Entomol 2005;31:189–97.

43. Magaña C. La entomologia forense y su aplicacion a la medicina legal. Data de la muerte. Bol SEA 2001;28:49–57.

44. De Jong GD, Hoback WW. Effect of investigator disturbance in experimental forensic entomology: succession and community composition. Med Vet Entomol 2006;20:248–58.

45. Adlam RE, Simmons T. The effect of repeated physical disturbance on soft tissue decomposition—are taphonomic studies an accurate reflection of decomposition? J Forensic Sci 2007;52:1007–14.

Additional information and reprint requests:
Ma. Teresa Valdes-Perezgasga, Ph.D.
Departamento de Parasitologia
Periferico Raul Lopez Sanchez km 2
Universidad Autonoma Agraria Antonio Narro-Unidad Laguna
Torreon, Coahuila
México CP. 27000
E-mail: cebolla_55@hotmail.com

CASE REPORT

PHYSICAL ANTHROPOLOGY

Ann W. Bunch,[1] *Ph.D.*

Indoor Wet Screening of Exhumed Skeletal Remains: A Suggested Procedure for the Preparation of Fragile Evidence for Anthropological Analysis

ABSTRACT: The 2007 exhumation of three children's graves, located in rural upstate New York and dating to 1979 and 1980, was warranted as their mother had come under suspicion for the death of a child she had been babysitting in late 2006. The local March weather conditions had been wet, and heavy rains fell during the 2-day process of casket removal. The extremely wet soil and the poor preservation of two wooden caskets increased the likelihood of damage to evidence. After remains' transport to the forensic center, an indoor wet-screening station was established so that skeletal elements could be (i) separated from soil matrix and (ii) preserved carefully for analysis. Not only were the remains relatively small and fragile in comparison with those of an adult, but two of the three remains were known to be fire damaged, thus the use of special laboratory preparation techniques was crucial.

KEYWORDS: forensic science, forensic anthropology, wet screening, forensic recovery, burned skeletal remains, perimortem fractures

In March 2007, an exhumation of three children's graves, dating to 1979 (two individuals) and 1980 (one individual), was ordered and executed in a small, rural cemetery in central New York State. The mother of the three children had come under suspicion because of a recent questioned death of a child she had been babysitting in late 2006. Melting snow and rain had previously created extremely moist soil in the area, and heavy rains fell during the excavation process causing operations to be terminated in the early afternoon on the first day. The surrounding soil matrix was comprised of a dense clay/clay-like loam, excavated first by a backhoe and later by hand once the top of the most recent (1980) casket was located.

Shovels and trowels were employed by personnel to excavate around the casket so that the base could be visualized. Once the base was completely exposed, photos were taken and the container was removed. The same method was used to remove the 1979 wooden caskets that were encountered underlying the more recent plastic one; however, in these cases, an inch-thick plywood board was inserted beneath each container prior to lifting them to ground level and to livery vehicles.

As soil became increasingly moist with falling rain and mist, the matrix took on a more clay-filled consistency that made excavation with hand tools much more challenging. Nonetheless, hand excavation was required because of the fragile nature of the wooden containers and the remains contained within. The 1980 casket appeared to be intact and sealed upon recovery, while the 1979 caskets were wooden and in very poor condition given the water damage and exposure to soil for the past 28 years. Removal of

these caskets was accomplished by the end of the second day at the cemetery site.

Once in the forensic laboratory setting, caskets were opened to assess preservation and processing/analytical requirements. The remains of the 5-month-old child contained in the 1980 casket were well preserved, much soft tissue was present and remains appeared relatively fresh. The remains of the two children (a 3-year-old and an 18-month-old) who had died in association with a house fire in 1979 were completely skeletonized, the wooden caskets surrounding them filled with mud and water at the time of recovery (Figs. 1 and 2). The medical examiner was able to assess the 1980 child's remains without assistance from the anthropologist; however,
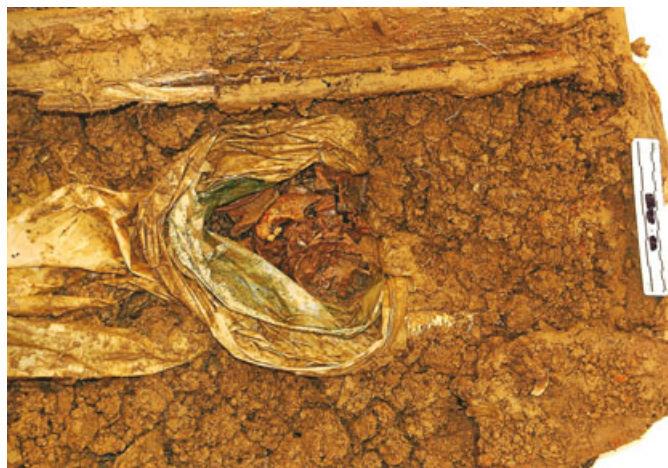

FIG. 1—*Individual 1 in coffin and plastic bags.*

[1]Department of Criminal Justice, College at Brockport, State University of New York, 164 Albert W. Brown Building, Brockport, NY 14420.

FIG. 2—*Individual 2 in coffin and plastic bags.*

anthropological analysis was required for the 1979 remains. The following discussion involves the 1979 remains that were in a much more advanced state of decomposition.

Methods

In the laboratory, the wooden caskets arrived on plywood boards and were transferred to metal gurneys for assessment. Mud and wet soil filled the interior of the coffins and had been left in this state (Figs. 1 and 2) to preserve evidence in the nonlaboratory conditions. However, in the laboratory, the remains had to be cleaned and dried for further anthropological analysis to be conducted. Numerous authors have addressed the proper processing of skeletal remains, especially for subsequent DNA-testing purposes (1–3), however special cases such as the present situation are not specifically dealt with in the literature.

The remains were known to be fire damaged at or during the time of death; however, blunt force trauma had not been considered at autopsy in 1979, despite noted cranial fractures. Upon present inspection, it was confirmed that all osseous elements were in an extremely fragile and fragmentary state and would necessitate exceedingly gentle treatment during the cleaning process so that no crucial information would be lost. The goal of the preanalysis cleaning was to preserve fracture patterns and coloration to the greatest degree possible, to reconstruct the death scenarios of both children. It should be noted that analysis of fire-damaged skeletal remains is challenging in ideal (dry) conditions, as coloration patterns must be analyzed together with fracture patterns, to determine timing and cause of fractures (4–7).

As cleaning procedures for sturdier skeletal remains (8) were deemed unacceptable in this case, an indoor wet-screening station was established to address this need (Fig. 3). A hose was set up near to a wooden excavation screen lined with quarter-inch wire mesh. The screen was set on two sawhorses and positioned near to the floor drain. The use of the floor drain with a wider diameter than a typical sink drain is preferable in cases such as this that involve much soil, vegetation, and stone matter.

Water from the hose drained directly into the screen on a very low-pressure setting. All material from the coffin(s) were removed by hand and placed in the screen, inundated with water, so that the soil matrix dissolved and harder evidence remained in the screen. No hose nozzle attachment was employed, rather the low water flow was utilized along with gloved hands to gently separate

FIG. 3—*Indoor wet-screening station.*

mud/soil from remains. In some cases, osseous element fragments were extremely small.

It should be noted that the screen is kept stable during the cleaning process unlike the "dry screen" process where screens are shaken for the soil to sift through the base mesh. The mud and soil will typically dissolve in the wet-screen process as cleaning proceeds, and thus will go through the mesh without shaking. If nonevidentiary items (e.g., gravel, wood, plant roots, etc.) accumulate in the screen, the screen can be inverted so that build-up does not hamper the sorting and cleaning procedure.

The location of evidence as it was found within the casket was documented and sketched in a notebook and immediately "reconstructed" on an adjacent gurney lined with dry brown paper so that the provenance of the remains was not lost during the cleaning process (Figs. 4 and 5). This allowed for later inventory of osseous elements present as well as body position at the time of interment.

Such wet-screen stations have been used routinely by the author for outdoor scenes such as aircraft crash sites where large amounts
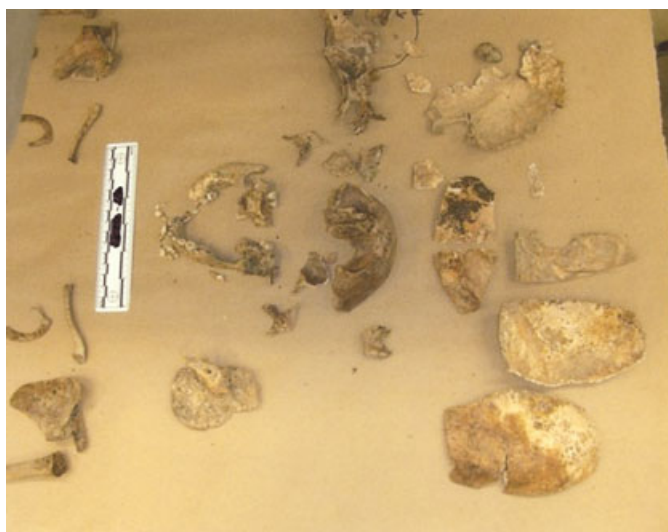

FIG. 4—*Individual 1: cranial, upper limb, shoulder girdle elements. Note various coloration of heat-affected portions.*

FIG. 5—*Individual 2: cranial, upper limb, and shoulder girdle elements. Varied colors of bone surface demonstrate differing levels of heat exposure.*

of evidence are required to be recovered from wet and clayey soil matrices (9).

Single stations for cases similar to the present one could conceivably be set up outdoors for increased drainage purposes. However, indoor wet screening was preferable in this situation for the following reasons: to avoid cold outdoor air temperatures, to avoid wind/precipitation that could hamper complete evidence recovery, and to maintain privacy in handling potential evidence of a homicide.

The cleaning process required a full work day (approximately 8 h) of two individuals (anthropologist and death investigator). The presence of the anthropologist for this screening was requisite as bone and nonbone (wood, stone) specimens needed to be separated/assessed throughout the process. After cleaning was completed, remains, which had been separated and placed/reconstructed on two paper-lined gurneys representing two individuals, were allowed to air dry at room temperature for 48 h (Figs. 4 and 5).

The black, gray coloration signifying exposure to fire/heat could be observed clearly once remains were completely dry (Figs. 4 and 5). In addition, numerous fractures, some heat related and some related to traumatic injury, were able to be assessed.

Results and Discussion

The manners of deaths of these children had been ruled accidental (caused by the house fire) at the time of autopsy despite surprisingly low carbon monoxide blood levels. Thus, the present ability to differentiate heat- and trauma-caused cranial fractures was critical. Such fractures were indeed discovered upon close examination of the cranial remains; both children shared similar perimortem fracture patterns that did not mimic heat-induced damage.

The indoor wet-screening method proved especially important in carefully maintaining provenance of the two individuals as well as preserving perimortem fractures without confounding these with postmortem damage. In addition, the preservation of coloration of

in situ remains was successfully carried out with this method, as only gloved hands and low-pressure water flow were used to clean skeletal surfaces, no brushes or other typically used tools were employed. Therefore, fire damage and burn patterns could be assessed in concert with fractures to sort out heat- versus trauma-caused injury (4–6).

Faced with the well-preserved and documented evidence in this case, the murder suspect pled guilty to the murder of her 5-month-old son (1980) and was given a 25-year sentence in state prison, as well as a concurrent 20-year sentence (in another county) for the death of the 13-month-old she was babysitting in 2006. Unfortunately, according to the ground rules established for the plea bargain, she was not required to address the 1979 deaths but knew about the findings in the case when discussion of suspicious patterns in the deaths of her children was evinced during grand jury hearings. It is likely that the knowledge of evidence that could be brought against her had an effect on her decision to plea bargain rather than face a jury of her peers.

References

1. Steadman D, DiAntonio L, Wilson J, Sheridan K, Tammariello S. The effects of chemical and heat maceration techniques on the recovery of nuclear and mitochondrial DNA from bone. J Forensic Sci 2006;51(1):11–7.
2. Fenton T, Birkby W, Cornelison J. A fast and safe non-bleaching method for forensic skeletal preparation. J Forensic Sci 2003;48:274–6.
3. Rennick S, Fenton T, Foran D. The effects of skeletal preparation techniques on DNA from human and nonhuman bone. Proceedings of the 57th Annual Meeting of the American Academy of Forensic Sciences; 2005 Feb 21–26; New Orleans, LA. Colorado Springs, CO: American Academy of Forensic Sciences, 2005;306.
4. Pope EJ, Smith OC. Identification of traumatic injury in burned cranial bones: an experimental approach. J Forensic Sci 2004;49(3):431–40.
5. Ubelaker D. The forensic evaluation of burned skeletal remains: a synthesis. Forensic Sci Int 2009;183(1):1–5.
6. Schmidt C, Symes S, editors. The analysis of burned human remains. London: Academic Press, 2008.
7. Dupras T, Schultz J, Wheeler S, Williams L. Forensic recovery of human remains. Boca Raton, FL: Taylor and Francis, 2006.
8. Cox M, Flavel A, Hanson I, Laver J, Wessling R. The scientific investigation of mass graves: toward protocols and standard operating procedures. Cambridge: Cambridge University Press, 2007.
9. Webster A. Recovery of a Vietnam-era aircraft crash: use of cross-cultural understanding and dual forensic recovery methods. J Forensic Sci 1998;43(2):277–83.

Additional information and reprint requests:
Ann Webster Bunch, Ph.D.
Department of Criminal Justice
College at Brockport, SUNY
164 Albert W. Brown Building
350 New Campus Drive
Brockport, NY 14420
E-mail: abunch@brockport.edu

CASE REPORT

PHYSICAL ANTHROPOLOGY

Tania Delabarde,[1] *Ph.D. and Bertrand Ludes,*[2] *M.D., Ph.D.*

Missing in Amazonian Jungle: A Case Report of Skeletal Evidence for Dismemberment*

ABSTRACT: This case study presents the results of the recovery and analysis of three sets of disarticulated and incomplete human remains found in Ecuador, within the Amazonian jungle. Recovered body parts sustained extensive sharp force trauma situated on different aspect of the skeleton. The anthropological examination (bone reassembly, biological profile) was followed by a detailed analysis of cut marks, including a basic experimental study on pig bones to demonstrate that dismemberment may have occurred within a certain amount of time after death. Despite the location (deep into the Amazonian jungle) and the perpetrator's actions (dismemberment and dispersion of body parts in a river), forensic work both on the field and in laboratory allowed identification of the victims and the reconstruction of the sequence of events.

KEYWORDS: forensic science, dismemberment, cut marks, chain saw, machete, postmortem interval

Dismemberment, the intentional separation of body segments, is known to be the major postmortem activity performed by humans on the remains of others (1). The main objective is to prevent identification of the victim by destroying body integrity as well as criminal evidence. Consequently, dismemberment is generally associated with homicide. Nevertheless, some authors presented cases where dismemberment injuries were linked to the cause of death (e.g., dismemberment as means of torture) (2,3). Consequently, it is essential if possible to determine the nature and sequence of injuries (perimortem/postmortem). Dismemberment patterns are of particular importance as the actions and instruments used are a clear expression of the perpetrator's intentions (4). Localized dismemberments involve the separation of some parts of the body and are commonly associated with hindered identifications, easing transport, or body disposal, while generalized dismemberments involve cuts to all aspects of body and are more often correlated with specific intentions (i.e., disregard of the victim or hiding evidence in sexual homicides). Specific capacities of the perpetrator (knowledge of anatomy, use of specific instrument) could be also determined based on the analysis of cut marks.

Sharp force trauma refers to injury produced by instrument that may incise, cut, chop, dent, or crush the bone tissue. Generally, injuries related to dismemberment are caused by cutting (knives), chopping (axes), and chiseling (saw). When a body is well preserved, analysis of soft tissue injuries represents the first step. Various aspects of the lesion usually present sharp force characteristics (e.g., incised wound with an associated groove or cut in the

underlying bone) with severed skin and adjacent tissues (5). Nevertheless, soft tissue injuries linked to dismemberment do not provide as much relevant data as bone tissue. The response to sharp force injury will greatly differ according to the biomechanical properties of the involved tissue. Because of their elastic limit and tolerance, soft tissue will be more damaged and stretched, which is not indicative of the causative instrument. Many publications demonstrate the potential of cut marks in bone to identify the causative instrument and lead to the perpetrator (4,6–13). Like other mechanisms of injury, analysis of sharp force trauma in bone is based on fracture patterns and defect morphology. Variability of wounds depends on the physical properties of causative instrument and the biomechanical properties of the affected bone (7). Because of the cost and gain of time, toothed instruments, and more specifically saws, are preferably used in dismemberment cases. During cutting, striae are formed by the uneven action of teeth on the wall of the cut. Power saws with small teeth leave the finest striations, while hand saws with fewer teeth cause most prominent striae. The number of striae correlates to the number of teeth. Therefore, marks will be used to determine general class characteristics (size, shape, width, and set of teeth) as well as the sawing action of the user (12).

Case History and Methods

Between December 2006 and March 2007, three sets of disarticulated human remains exhibiting extensive sharp force trauma were found in different locations within a small area of the Amazonian Jungle in the Ecuadorian territory. In August 2006, two European tourists were reported missing in this area. On December 19, 2006, the lower half of a body designated as set #1 (partially skeletonized pelvic girdle, both legs and foot) was found by local fishermen, lying on the border at the confluence between Pastaza and Puyo Rivers. The body part was still wearing underwear and pants. On February 22, 2007, the upper half of a body designated as set #2 was found in a plastic bag lying on the shore of the Paztaza River. The remains (head, trunk, and arms) were partially disarticulated and skeletonized and associated with a sports jacket and tee-shirt.

[1]French Institute of Andean Studies, Whymper 442 y Coruña, Quito, Ecuador.

[2]Institute of Legal Medicine, 11 rue de Humann, Université de Strasbourg, 4 rue Kirschleger, 67000 Strasbourg, F-67085 Strasbourg Cedex, France.

Evidence of carnivore scavenging was found on distal extremity of right humerus and proximal extremity of right radius and ulna. On the March 11, 2007, a third set of human remains designated as set #3 (pelvic girdle articulated with both femurs) was exhumed from a grave dug by the Natives. According to their testimony, they found the body parts lying on the shore of the Paztaza River in January 2007 and decided to bury them and then called police investigators. The delay between the finding and exhumation is correlated with the location of the scene deep inside the Amazonian forest. The remains were consistent with a saponified pelvic girdle articulated to both femurs. This body part was still wearing male underwear. The three sets of remains were sent to the Quito mortuary, and our participation was requested at this stage.

The forensic examination was performed by two pathologists and one anthropologist (TD). Examination of the soft tissues and organs was very limited because of the state of decomposition. The possibility to assess time since death was restricted, considering that body parts usually decompose at a slower rate than the whole corpse and also taking into account the delay between discovery and forensic examination (no refrigeration was available in this region, and no specific transport either). The first issue of the anthropological analysis was to determine the exact number of

TABLE 1—*Description and characterization of sharp force injuries sustained by Victim 1 and Victim 2.*
(First number in SFT# refers to set of remains: 1.3 designated set #1 injury 3).

SFT#	Type	Location and Measurements	Characteristics	Estimated Causative Tool
1.1	False start kerf	At the junction between left iliopubic ramus and acetabulum (18 × 6 × 5 mm)	Concave defect associated with a fracture running to the medial aspect of left iliopubic ramus and bone spalling	Power saw (chain saw?)
1.2	Complete transverse-sectioned bone cut	Circular defect located on the antero-superior aspect of the left femoral head (19 mm diam approx)	Ablation of a circular portion of left femoral head, clear section bone cut (polish appearance), no associated fracture. Anatomically linked to the false start described above (#1.1)	Power saw (chain saw?)
1.3	Cut mark (slipped mark)	At the level of the superior anterior margin of acetabulum (12 × 3 × 1 mm)	Superficial defect with polish appearance; this could be the result of the tool blade slipping during cutting intent previously described (#1.1)	Power saw (chain saw?)
1.4	False start kerf	At the level of the anterior inferior iliac spine (19 × 7 × 15 mm)	Kerf wall exhibits a clean cut section followed by irregular kerf floor with bone spalling and fractures; this may be linked to vibration and/or resistance that led to abandoning this location	Power saw (chain saw?)
1.5	False start kerf	At the level of the inferior gluteal line (34 × 10 × 13 mm)	Very irregular kerf wall exhibiting tooth bite that reflects a problematic blade progress (vibrations/resistance?) that led to abandoning this location	Power saw (chain saw?)
1.6	Complete transverse-sectioned bone cut	Superior aspect (1/4) of both ilium bones is missing, posterior aspect of sacrum is missing at the level of S1-S2, left ala is missing	Entrance just above the previously cut section described (#1.5) the blade progressed through left ilium to sacrum and right ilium (notch on right sacro-ilac joint and notch on right anterior superior iliac spine); clear bone section on right os coxa; the anterior aspect of pelvic girdle exhibits bone spalling associated with running fractures, exit chipping on the posterior aspect of both os coxae, exit at the level of the anterior superior iliac spine of right ox coxa (left to right/below to above)	Power saw (chain saw?)
2.1	Complete transverse-sectioned bone cut	At the level of the inferior aspect of body and spinous process of lumbar 4 (one-third of the body and inferior articular facets are missing)	Entrance on the anterior aspect of vertebral body in the midline, cut section exhibited weavy-edged walls followed by a bone island located on the posterior portion of vertebral body, suggesting a slight change of plane stroke but without interruption of the cut section with an exit in spinous process (front to back/above to below)	Power saw (chain saw?)
2.2	Irregular incomplete defect	At the level of the medial border of left scapula, there is a linear incomplete defect (22 × 5 mm) at 21 mm from the junction with the scapular line	Anteriorly, the defect is more regular, and the inferior end showed a quadrangular shape associated with a slight displacement with plastic deformation of the adjacent bone. The medial wall of the defect is almost linear with some irregularities. Posteriorly, the medial wall showed jagged edges and detached fragments of bone associated in the inferior end of the injury with a radiating fracture line running to the scapular spine. The defect is consistent with penetrating injury caused by a bladed instrument. The indentation of the medial wall and detached bone fragments are correlated with the rotation and withdraw of the blade.	Bladed instrument (knife?)
3.1	Chop cut mark	(12 × 2 mm) is located on the medial aspect of patellar surface (left femur)	Small concave defect exhibits bone flaking associated with fracture (left patella was missing; tibia and fibula are absent)	Heavy blade (machete?)
3.2	Chop cut mark	(10 × 3 mm) is located on the lateral aspect of patellar surface (left femur)	The defect exhibits multiple facets and bone flaking associated with fracture. In both cases, the impact was oblique to the femur, and injury was inflicted from front to back	Heavy blade (machete?)
3.4	Complete transverse-sectioned bone cut	A quadrangular cut section (53 × 30 × 1.5 mm) is located on the medial condyle (left femur)	Entrance on the posterior aspect of medial condyle, cut section exhibits a bone island suggesting a slight change of plane stroke but without interruption of the cut section with an exit in the anterior aspect of medial condyle (back to front)	Power saw (chain saw?)
3.5	Ovalar cut section	(25 × 15 mm) is located on the lateral condyle (left femur)	The defect exhibits a clear cut surface with polish appearance and no fracture, probably related to previous cut section (#3.4) as both condyles of left femur sustained same cutting plane	Power saw (chain saw?)
3.6	Ovalar cut section	(20 × 15 mm) is located on the inferior margin of lateral condyle (right femur)	The defect exhibits a clear cut surface with polish appearance and no fracture (right patella is present and exhibits no evidence of injury; tibia and fibula are absent)	Power saw (chain saw?)

individuals through bone reassembly and exclusion. The three sets of remains showed very different long bone proportions. Following Konisgberg's recommendations (14), stature was estimated with two variables (humerus and femur length). These measurements allow associating set #2 and set #3 and excluding set #1. The determination of sex was based upon the morphology of both pelvic girdles (set #1 and set #3) and the skull (set #2) combined with measurements of long bones from all sets. Age estimation was performed with the combination of traditional methods for adults based on 4th rib and pubic symphisis (1). The medial extremities of both clavicles from set #2 were partially fused, and fusion lines on femoral head were still visible on set #1. The three sets of remains were determined to belong to two young Caucasian adult men. Individual characteristics were also relevant on set #2 (head, trunk, and arms) with the presence of a healed fracture on right clavicle midshaft and conclusive dental data. Results of the forensic examination were consistent with the antemortem data of the two missing persons. Clothing associated with the remains was also recognized by the families. Considering that osteological data from set #3 were not relevant to missing person reports, DNA samples were sent from each of the three sets of remains. Reassembly of the victims' remains and identification were later confirmed by DNA testing.

The body parts sustained a number of cut marks situated on different aspects of the skeleton; some of them were not visible when soft tissue was still attached, showing the importance of bone cleaning. Each cut mark is listed in Table 1 with the location, number, type (e.g., false start), and measurements (e.g., width, length, and depth) following standard anthropological protocols (4,6). According to Symes recommendation, false starts were distinguished from slipped marks of the blade (12) (Table 1).

Results

Victim 1 (set #3)

Victim 1 is only represented by lower half of body and sustained a complete separation at the level of the pelvic girdle (Fig. 1). False starts at the level of the left acetabulum were followed by a successful separation at the level of upper portion of pelvic. Direction of blade progress is traced from hesitation marks to breakaway spur: exit chipping is because of bone spalling as teeth exit the bone and indicate direction of cutting stroke (6). False starts on left os coxa (Fig. 2), notch on right sacro-iliac joint and right anterior superior iliac spine, and exit chipping on the posterior aspect of

both os coxae (Fig. 3) suggested that the victim was lying on the lateral right side, the saw blade entered on left os coxa and progressed through the sacrum to the right os coxa. Therefore, the perpetrator probably stood at the back of the victim.

Victim 2 (set #1 and set #2)

Victim 2 was almost complete with the exception of lower legs and feet. The left scapula exhibited a defect located on the medial border that is consistent with a penetrating perimortem injury (plastic deformation, jagged edges associated with fractures) caused by a bladed instrument (Fig. 4). Shape and size of the defect suggest that the blade could have penetrated enough to perforate the left lung. It is important however to stress that no evidence of sharp force trauma was observed on adjacent ribs. This could be correlated with the position of the scapula (e.g., the medial border could have been almost parallel to the ribs to allow the penetration without damage). The location of this injury is considered as lethal if untreated. Victim 2 also presented the successful separation of trunk at the level of lumbar 4th. Orientation of striae and bone islands suggests that the saw penetrated on the anterior right aspect
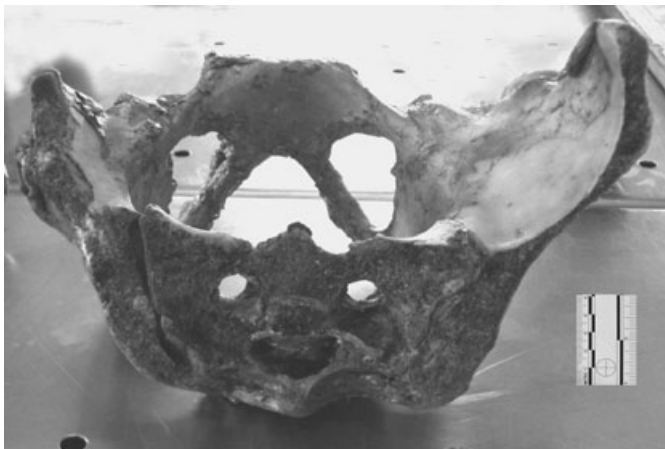

FIG. 2—*View of left coxo-femoral articulation from Victim 1 with false starts.*
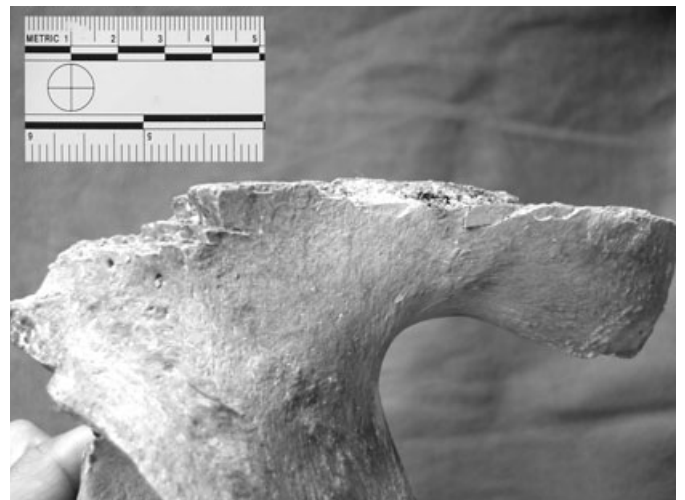
FIG. 3—*Posterior view of left os coxa from Victim 1 with false start and tooth impression.*
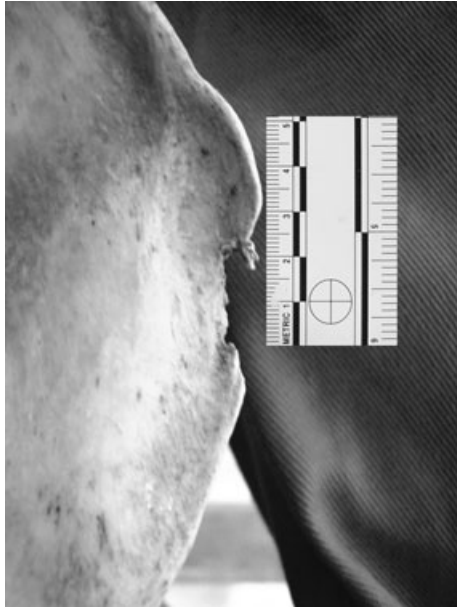

FIG. 1—*General view of sectioned pelvic girdle from Victim 1.*

FIG. 4—*Anterior view of left scapula from Victim 2 and detail of the defect located on medial border.*

FIG. 6—*Lateral left view of 4th lumbar from Victim 2 with complete bone section (blade entered on the anterior aspect of vertebral body and exited throught spinous process).*

of lumbar body at the level of midline and exited through the spinous process (Figs. 5 and 6). The victim was probably lying on the back, and the perpetrator stood at the right side of the victim according to blade progress. The distal epiphysis of left femur sustained two injuries inflicted by a heavy blade and two others related with the use of a power saw (Fig. 7). There was almost certainly a first attempt to separate legs with a bladed instrument (anterior aspect of condyles), and then a power saw was successfully used (entrance on the posterior aspect of condyle and exit anteriorly). Chop marks were relevant enough to suspect the use of a machete. Victim 2 also exhibited some evidence of scavenger activity with distal end of right humerus and proximal end of right radius and ulna showing tooth marks. Tooth impressions were consistent with small carnivores. It was then established that the remains of this individual were accessible for a certain time to animals.

The total number of sharp force injuries identified on the remains of the two victims is $n = 13$. Injuries can be divided according to their morphological characteristics and linked with estimated causative instruments:

- $n = 10$ injuries sustained by both victims showed recognizable features related with the action of a power saw, probably a chain saw (e.g., weavy-edges walls). However, cut marks were described individually and a few of them could be related to the same cutting stroke (e.g., injuries on lateral and medial condyle of left femur, victim 2).
- $n = 2$ injuries are consistent with chop wounds (concave defect, bone flaking) inflicted with a heavy blade, probably a machete (victim 2).
- $n = 1$ injury on left scapula is consistent with the penetration and withdrawal of a blade; size and shape of the defect suggest a knife (victim 2).

FIG. 5—*Caudal view of 4th lumbar from Victim 2 sustaining weavy-edges walls on complete cut section.*

FIG. 7—*Left femur of Victim 2 showing quadrangular cut surface resulting from a change of direction in blade progress on medial condyle and ovular cut surface on lateral condyle.*

Victim 1 (lower half of body) sustained one type of sharp force injury (cut marks produced by power saw), while victim 2 (almost complete body) sustained at least three different types of sharp force trauma (cut marks produced by power saw, cut marks produced by heavy blade, and one penetrating injury produced by a blade).

The timing of injuries is one important consideration in the evaluation of any skeletal trauma with respect to death (15). In cases of dismemberment, the main limitation in determining cause of death is the absence of some organs and parts of the body. The cause of death was ascertained for victim 2 (perimortem injury on left scapula) but remains unascertained for victim 1. The location of the other sharp force trauma (trunk and knees) is characteristic of a dismemberment activity with the successful separation of body segments. It is probably not a coincidence if the victim who sustained cut marks both at the level of trunk and at knee joint was the taller of the two missing persons (ease transport?). The position of the bodies and successful separation of body segments (trunk and legs) suggest postmortem dismemberment. No evidence of defense wound was found on both victims' remains. Saw cut characteristics indicate a mechanically powered saw (high transfer energy suggests polish and wide kerf widths, and deep false starts) and more specifically a chain saw (weavy-edges walls, irregular concave bending striae, bone islands). A review of forensic literature helps to support the hypothesis that cut marks were not presenting characteristics expected, while fresh bone is exposed to a major trauma inflicted with a chain saw (4,6–10,12). A basic experimental study on pig bones was performed to evaluate the postmortem interval between death and dismemberment as well as confirm causative agent related to cut sections. Two adult pigs recently killed were separated in two groups. Group A represents fresh body parts, while pig parts from group B were buried wrapped in clothing for 5 months in a place with similar outside environmental conditions (temperature/precipitation) as the jungle. A chain saw commonly used in the Amazonian jungle was employed to dismember fresh pig parts and then buried samples from group B. Because of biomechanical properties (e.g., elasticity), fresh pigs bone exhibited less damage, the clearest bone sections (Fig. 8) and noticeable characteristics (e.g., exit chipping) than any body parts buried for 5 months. Comparisons between bone cutting sections from pig bones and the actual forensic case were similar enough to determine the chain saw as the causative instrument used for the successful separation of body segments from the two victims. Despite the lack of laboratory facilities, this experimental study helps us to confirm the causative instrument and establishes the possibility of an interval between death and dismemberment of the two victims as cut marks exhibit characteristics of bone response in drying process.

Discussion

After receiving our forensic report, the police informed us that a chain saw had been observed but not collected during a search in a property of colonists close to locations where the three sets of victims' remains were discovered. We then participated in a new search of this property as we suspected that victim bodies could have remained buried during a certain time before being dismembered and thrown into the river. The chain saw had disappeared in the meantime, and despite intensive searches, including a canine team, no more human remains were found. Nevertheless, a grave was finally located in the proximity of the property previously mentioned, on the other side of the river. The size and shape of the grave recovered two complete extended bodies, no clothing was found but a strong smell and remains of decomposition as well as the rest of tied ropes were located in the bottom of the grave. During the search in the property, few tools including a machete were collected as well as a wood support exhibiting numerous cut marks characteristic of a chain saw. A fireplace was excavated, and remains of burned synthetic material were identified. Fibers later analyzed were consistent with synthetic material known to be used for sleeping bag production. Facing the results of both recovery scene and the forensic autopsy, the alleged perpetrators confirmed the burying and dismemberment of the two victims but denied committing the homicide. In the beginning of December 2006, when police investigators were searching around, owners of the property were afraid of the possibility to be link to the homicide and decided to exhume, dismember, and throw the dead bodies into the river. According to their testimony, three native men came on August 6, 2006 transporting the dead bodies of two tourists in their sleeping bags. They explained that death occurred accidentally during a jungle trip, after the drinking of a powerful hallucinogenic beverage. The first set of remains was found on December 19, 2006, and the two victims went missing on August 5, 2006. Consequently, the interval between death and dismemberment was confirmed to be 4 months.

Three forensic challenges were faced by teams working on this case: the condition of the remains (dismembered bodies of two tourists), the location (Amazonian jungle), and the lack of laboratory facilities. In forensic literature, dismembered cases generally focused on soft tissue injury patterns or bone analysis through cut marks to identify the causative instrument, but very few linked the analysis of injuries with data from the recovery scene. This case study illustrates the value of the timing of injuries linked to dismemberment. The interval between death and dismemberment was determined by the results of forensic examination and basic experimental study and confirmed by evidences collected on the scene (e.g., grave). The analysis and interpretation of sharp force injury and involved instruments were also essential to identify alleged perpetrators in a specific region with little human settlement. Native people (*Shuar* Indians) living in this area belong to Jibaro family and have achieved their notoriety through their customary practice of head-shrinking (*tsansa*) and war hunting. It was then important to exclude any ritual practices or murders already known in this region because of territory rivalry between native and colony communities. Despite the complexity of the case and the location, a
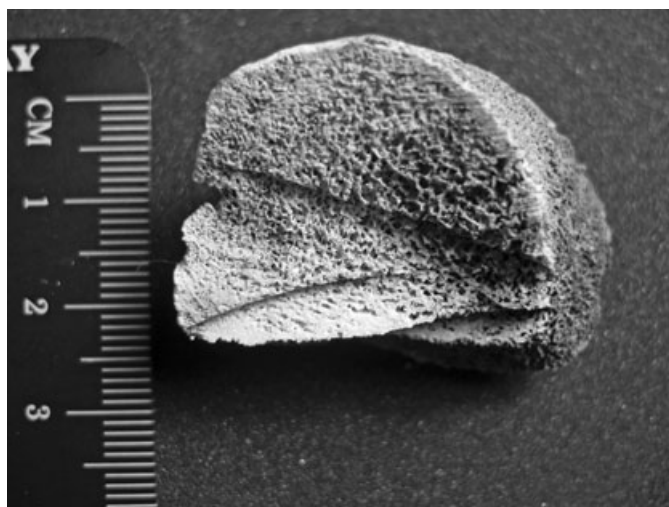

FIG. 8—*Cut section produced by chain saw on pig femur from group A.*

strong collaboration on the field and in the mortuary from all the parts involved allowed identification of the victims and to some extent the reconstruction of the sequence of events. Unfortunately, to date, circumstances surrounding the death of the two victims remain unknown, and trial has been delayed many times.

References

1. Byers SN. Introduction to forensic anthropology, a textbook, 2nd edn. Boston, MA: Pearson Education, 2005.
2. Kimmerle EH, Baraybar JP. Skeletal trauma, identification of injuries resulting from human rights abuse and armed conflict. New York, NY: CRS Press, 2008.
3. Konopka T, Strona M, Bolechala F, Kunz J. Corpse dismemberment in the material collected by the Department of Forensic Medicine, Cracow, Poland. Leg Med Tokyo 2007;9(1):1–13.
4. Reichs KJ. Postmortem dismemberment: recovery, analysis and interpretation. In: Reichs KJ, editor. Forensic osteology, advances in the identification of human remains. Springfield, IL: Charles C. Thomas, 1998;353–98.
5. Dimaio VJM, Dana SE. Handbook of forensic pathology, 2nd rev. edn. New York, NY: CRC Press, 2007.
6. Symes SA, Berryman A, Hugh E, Smith OC, Black C. Saw dismemberment of human bone: characteristics indicative of saw class and type. Proceedings of the 44th Annual Meeting of the American Academy of Forensic Sciences; 1992 Feb 17–22; New Orleans, LA. Colorado Springs, CO: American Academy of Forensic Sciences, 1992.
7. Smith OC, Pope EJ, Symes SA. Look until you see: identification of trauma in skeletal material. In: Steadman DW, editor. Hard evidence: case studies in forensic anthropology. New Jersey: Prentice Hall, 2003;138–54.
8. Madea B, Driever F. Cadaver dismemberment by chain saw. Arch Kriminol 2000;205(3-4):75–81.
9. Reuhl J, Bratzke H. Death caused by chain saw: homicide, suicide or accident? A case report with a literature review. Forensic Sci Int 1999;105:45–59.
10. Schultz Y, Mossakowski H, Albrecht K, Bretmeier D. Postmortem dismemberment/mutilation; medico legal and criminalistic evaluation of the autopsies performed by the Institute of Legal Medicine at the Hanover Medical School. Arch Krimnol 2008;221(1-2):1–16.
11. Quatrehomme G. A strange case of dismemberment. In: Brikley MB, Ferlini R, editors. Forensic anthropology: case studies in Europe. New Jersey: Humana Press, 2007;99–119.
12. Symes SA. Morphology of saw marks in human bone: introduction and examination of residual kerf contour. In: Reichs KJ, editor. Forensic osteology, advances in the identification of human remains. Springfield, IL: Charles C. Thomas, 1998;389–409.
13. Saville PA, Hainsworth SV, Rutty GN. Cutting crime: the analysis of the "uniqueness" of saw marks on bone. Int J Legal Med 2007;121:349–57.
14. Koningsberg LW, Ross AH, Jungers WL. Estimation and evidence in forensic anthropology: determining stature. In: Schimtt A, Cunha E, Pinheiro J, editors. Forensic anthropology and medicine, complementary sciences from recovery to cause of death. New Jersey: Humana Press, 2006;317–31.
15. Sauer NJ, Lovis WA, Blumer ME, Fillion J. The contribution of archaeology and physical anthropology to the John McRae case. In: Steadman DW, editor. Hard evidence: case studies in forensic anthropology. New Jersey: Prentice Hall, 2003;117–26.

Additional information and reprint requests:
Tania Delabarde, Ph.D.
French Institute of Andean Studies
Whymper 442 y Coruña
Quito
Ecuador
E-mail: tania.delabarde@gmail.com

CASE REPORT

GENERAL: PSYCHIATRY & BEHAVIORAL SCIENCES

Robert J. Morton,[1] *M.S.; Carlo P. Campobasso,*[2] *M.D., Ph.D.; James J. McNamara,*[1] *M.S.;*
Massimo Colonna,[3] *M.D.; Felice Carabellese,*[4] *M.D.; Ignazio Grattagliano,*[4] *Psy.D.;*
Roberto Catanesi,[4] *M.D.; and Jennifer M. Lawrence,*[1] *M.A.*

Cross-Cultural Comparison of Two Serial Sexual Murder Series in Italy and the United States

ABSTRACT: There have been few documented comparisons of serial murder cases committed in the United States with cases occurring internationally. The authors contrasted two unique serial murder series: one in Italy and one in the United States by examining the details of both series, including the M.O., motivation, crime scene interactions, sexual acts performed, and the general backgrounds of both offenders. The comparison revealed a number of similarities. Both offenders specifically targeted elderly women, who were attacked in their residences. The two series involved sexually motivated crimes, although the sexual interactions were different. Both offenders stole objects from their victims after the murders and each kept newspaper accounts of their crimes. In addition, both offenders claimed to have abusive upbringings, including sexual abuse.

KEYWORDS: forensic science, serial murder, elderly victims, sexually motivated murder, cross-cultural, sexual homicide

Serial Murder

There has always been great interest in serial murder from law enforcement, the medicolegal community, academicians, and the general public. This interest has generated numerous academic studies as well as a plethora of books, movies, and articles (1). Despite the intense interest, serial murder is a relatively uncommon occurrence, comprising <1% of all murders committed in the United States (1–3).

Although serial murder historically has occurred all over the world, the majority of studies concerning serial murder were based on cases that have occurred in the United States (2). There have been, however, a number of other studies on serial murder in countries outside of the United States, including England, Germany, Canada, South Africa, and The Netherlands (4–8). Throughout all of these studies, serial murder has been defined in a variety of ways, including the number of murders required, temporal aspects, and motivations of the offenders. The shared components of all these definitions include two or more victims, killed in separate events, and committed for a number of different motivations (1,4–8).

Sexually Motivated Serial Murder

There are a variety of different classification systems used by law enforcement, academicians, and members of the mental health community to identify and analyze serial murders and the offenders who commit these crimes (1,2,9,10). These classifications divide serial murder into different categories and most include sexual motivation (1,2,4,10). As a result, sexually motivated murder is considered by many experts to be a distinct subset of serial murder (2,3,10).

Murder of the elderly is a rare event, occurring in <3% of all homicides committed in the United States (11). Among those murders, there is a smaller set of murders that are sexually motivated (11). Serial sexual murder of the elderly is an even smaller subset within the category of serial sexual murder (11).

Ben Mohamed Ezzedine Sebai Case History

Ben Mohamed Ezzedine Sebai was born in 1962, in Tunisia, Africa. In 1990, he immigrated illegally to northern Italy. In 1991, he was charged with attempted murder and rape by the Police headquarters of Bolzano in northern Italy, and they issued an expulsion order. Sebai fled before trial and relocated to the Province of Foggia, located in southern Italy, where he worked occasionally as a farm laborer.

Between 1996 and 1997, there were 12 homicides of elderly women over the age of 70, in the Apulia territories of southern Italy. The chronological distribution of the victims were three in 1996 and nine in 1997 (Fig. 1). All of the victims were found in their own apartments, all of which were located on the ground

[1]National Center for the Analysis of Violent Crime, Federal Bureau of Investigation, FBI Academy, Quantico, VA 22135.

[2]Department of Health Sciences (S.pe.S), University of Molise, via De Sanctis, 86100, Campobasso, Italy.

[3]Section of Legal Medicine (DI.M.I.M.P.), University of Bari, Policlinico, Piazza Giulio Cesare, 70124 Bari, Italy.

[4]Section of Forensic Psychiatry (DI.M.I.M.P.), University of Bari, Policlinico, Piazza Giulio Cesare, 70124 Bari, Italy.

2010 American Academy of Forensic Sciences

Published 2010. This article is a U.S. Government work and is in the public domain in the U.S.A.

1111

level, and there were no signs of forced entry. The 12 victims were killed by stab wounds to the neck. DNA evidence was found in only one of the cases.

The murders were committed in two areas: the Province of Foggia and the Province of Taranto, which are located approximately 300 km south of Foggia (Fig. 2). These territories are under the jurisdiction of three different tribunals: the Tribunal of Foggia and the Tribunal of Lucera located in the north and the Tribunal of Taranto located in the south. The murder investigations were conducted separately by three different Judicial Authorities based upon the location where the particular victim lived.

In the early part of 1997, because of the similarities in the cases, the local Italian press purported a theory that a single, serial killer was at work and had murdered all of the elderly victims. This theory continued to gain momentum throughout 1997 as more elderly victims were discovered murdered under similar circumstances.

Based on increasing public pressure, a multidisciplinary team was assembled to review each of the crimes. The team included a forensic pathologist and a forensic psychiatrist. After their review, the team determined that all of the murders appeared to be sexually motivated and had the following common characteristics: the victims were women over the age of 70, who lived alone in their own ground floor apartments; there were no signs of forced entry; all of the victims' apartments were ransacked; and in some of the cases money and/or jewelry was taken. All of the victims were stabbed in the neck, primarily on the left side, and none of the victims had defensive injuries. The offender apparently brought the knife to the scene and took it with him afterward. In several of the cases, diluted blood evidence indicated that the killer washed his hands in the victims' residences.

The team concluded that the method of operation was unusual for the local Italian criminal element, suggesting that the offender could be an immigrant. Furthermore, the team opined that the offender might have a prior arrest for sex-related crimes or attempted murder because the murders appeared to be sexually motivated.

On September 16, 1997, Sebai was arrested after being identified by a neighbor of the last murder victim, who had observed Sebai discarding his blood-stained clothing. In 2000, Sebai was tried for 4 of the 12 murders. During the trial, Sebai continually claimed he was innocent; nevertheless, he was convicted and sentenced to a life term.

Even though all of the murders appeared to have been committed by the same offender and Sebai had been convicted of four of them, the investigation to link Sebai to all 12 homicides stalled (Table 1). This was because a number of other people had previously been tried and convicted of killing four of the victims included in the series of 12 cases.

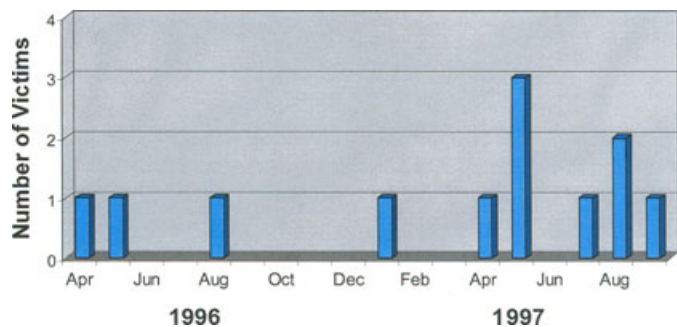After several years of incarceration, Sebai confessed to a total of 15 murders, including three cases not previously linked to the series (Table 2). During his confession, Sebai admitted to masturbating at every homicide, even though his DNA was recovered at only one scene. Ironically, Sebai left very little forensic evidence that could conclusively link him to all 15 murders.

Michael Darnell Harris Case History

Michael Darnell Harris was an African American, born in 1963, and was the oldest of five children. He had two younger sisters and two younger brothers, and he spent his early childhood in Muskegon, Michigan. When Harris was 10 years old, his mother got a job at the U.S. Veteran's Hospital in Ann Arbor, Michigan. Harris' family, including his mother's boyfriend, moved to Ann Arbor. In May of 1980, Harris moved by himself to Lansing, Michigan.

During the period of October 18, 1980 through September 30, 1982, there was a series of murders involving nine elderly women in southern Michigan (Fig. 3). These murders occurred in the cities of Lansing, Portland, Ann Arbor, and Ypsilanti (Fig. 4). The victims ranged in age from 77 to 91 years. All of the victims were elderly white women with the exception of one victim, who was a light complected, African American woman. Every one of the victims lived alone, and all were assaulted and murdered inside their residences. The offender apparently entered the residences of the victims through unlocked doors, most often the rear entry doors.
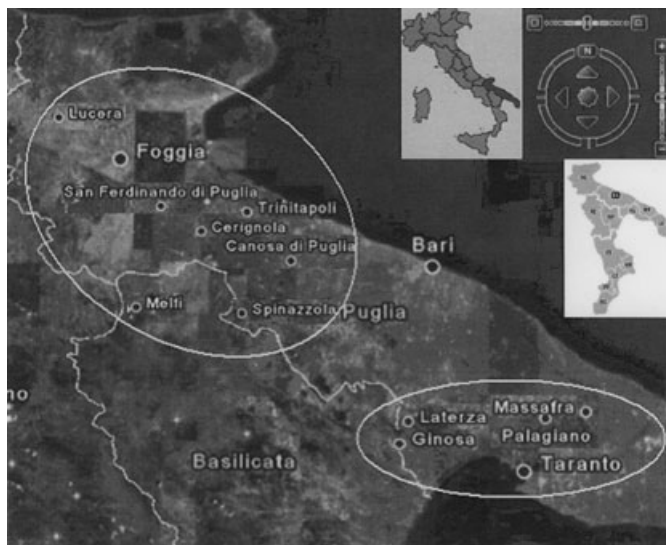

FIG. 2—*Distribution of Sebais' victims along two main areas of the Apulia territories, under the jurisdiction of three different Tribunals.*

TABLE 1—*Homicides of elderly women aged over 70 that Ben Ezzedine Sebai confessed to in 2005 and 2007.*

Case	Date	Name	Age	City	Cause of Death
1	04/24/1996	M C	81	Lucera	1 Stab Wound
2	05/29/1996	G G	72	S. Ferdinando	17 Stab Wounds
3	08/10/1996	S A	85	Ginosa	4 Stab Wounds
4	01/15/1997	T M	75	Cerignola	11 Stab Wounds
5	04/04/1997	M G	75	Massafra	1 Stab Wound
6	05/01/1997	S A	70	Trinitapoli	1 Stab Wound
7	05/09/1997	L S	82	Canosa	8 Stab Wounds
8	05/14/1997	L P R	86	Castellaneta	3 Stab Wounds
9	07/29/1997	V M	83	Palagiano	25 Stab Wounds
10	08/21/1997	L L	90	Laterza	2 Stab Wounds
11	08/27/1997	S A	84	Spinazzola	5 Stab Wounds
12	09/16/1997	N L	75	Palagianello	5 Stab Wounds

FIG. 1—*Distribution of Sebais' victims between 1996 and 1997.*

TABLE 2—*Homicides of elderly women aged over 70 that Sebai confessed to in 2005 and 2007.*

| Case	Date	Age	City	Cause of Death	Disposition
1	06/1995	–	Foggia	Unknown	Confessed in 2007
2	07/08/1995	83	Melfi	Unknown	Confessed in 2007
3	08/13/1995	76	Palagiano	Unknown	Confessed in 2005
4	04/24/1996	81	Lucera	1 Stab Wound	Confessed in 2007
5	05/29/1996	72	S. Ferdinando	17 Stab Wounds	Confessed in 2007
6	08/10/1996	85	Ginosa	4 Stab Wounds	Confessed in 2007
7	01/15/1997	75	Cerignola	11 Stab Wounds	Life Sentence in 2000 Confessed in 2007
8	04/04/1997	75	Massafra	1 Stab Wound	Confessed in 2005
9	05/01/1997	70	Trinitapoli	1 Stab Wound	Confessed in 2007
10	05/09/1997	82	Canosa	8 Stab Wounds	Confessed in 2007
11	05/14/1997	86	Castellaneta	3 Stab Wounds	Confessed in 2005
12	07/29/1997	83	Palagiano	25 Stab Wounds	Life Sentence in 2000 Confessed in 2005
13	08/21/1997	90	Laterza	2 Stab Wounds	Confessed in 2007
14	08/27/1997	84	Spinazzola	5 Stab Wounds	Life Sentence in 2000 Confessed in 2007
15	09/16/1997	75	Palagianello	5 Stab Wounds	Life Sentence in 2000 Confessed in 2007
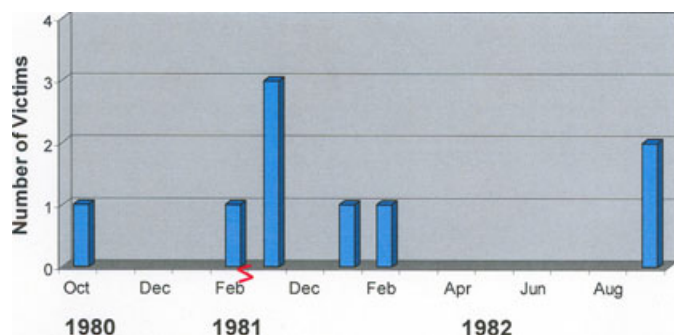
FIG. 3—*Distribution of Harris' victims between 1980 and 1982.*

The first five victims were killed by asphyxia, either manual or ligature strangulation, or both. The sixth victim had her throat slashed. The seventh victim was stabbed in the neck. The last two victims were strangled with nylon stockings and pantyhose used as ligatures. None of the victims suffered any defensive wounds.

All the victims also suffered various forms of blunt force trauma, to include bruising, broken ribs, and blows to the face and head. Several had vaginal lacerations, and two had postmortem incised wounds to their necks. All the victims, with the exception of the African American woman, had been sexually assaulted vaginally. In addition, two of the victims had been vaginally assaulted by foreign object insertion. In one case an umbrella was used, and in the other case, an unknown object was used and not recovered. In at least one case, the forensic pathologist opined that the victim was sexually assaulted postmortem.

In 1983, Harris was arrested and convicted of the rape of a 68-year-old, elderly woman, who was attacked inside her residence. The victim had observed a young man walking up and down the street in front of her home. A short time later, she observed the same young man in her residence, where he confronted her and told her he was going to rape her. He strangled her until she passed out, and when she awoke, he was gone. In addition, she suffered several knife wounds to the throat. There were also physical injuries indicating that she had been sexually assaulted. A total of three hundred dollars was taken from her purse, which had been located in her bedroom. Harris was subsequently identified by a number of witnesses in the neighborhood, including the victim. He was eventually convicted of assault, breaking and entering, and criminal sexual conduct, and he received sentences totaling up to 90 years in prison.

The similarities of the rape to the unsolved murders led the investigators to focus on Harris as a possible suspect (Table 3). In 1983, Harris was tried separately for two of the murders, convicted of both, and received a sentence of life imprisonment. Harris denied his guilt and filed numerous appeals in an attempt to overturn his convictions.

In 2000, the prosecutor's office decided to re-investigate Harris for additional murders of elderly women that he was suspected of committing in southern Michigan. At this time, Harris was petitioning the Michigan State Parole Board for an early release, claiming that he no longer thought of elderly women in a sexual way.

Based upon a recommendation from the Federal Bureau of Investigation's (FBI) National Center for the Analysis of Violent Crime (NCAVC), Behavioral Analysis Unit (BAU), a routine search of Harris' prison cell yielded press clippings relating to the homicides Harris was suspected of, as well as fantasy writings of Harris', including an autobiography and poetry. Investigators also recovered a pornographic magazine titled *Women Over Fifty* from under Harris' pillow, which featured elderly women posed naked, in sexually provocative ways. In 2003, Harris was convicted of two additional murders of elderly women and received three additional life sentences (Table 4).

Discussion

Although serial murders are relatively infrequent events, the comparison of these two cases involving Sebai and Harris clearly shows similarities in the way both offenders committed their crimes and selected their victims. Both chose elderly victims purposely and attacked elderly women exclusively.

Sexually motivated offenders choose victims from a number of parameters best described as vulnerability, availability, and desirability (12,13). Vulnerability is defined as the factors that allow easy predation of a victim by an offender (13). These factors include the potential victims' age, sex, health, occupation, or temporal circumstances (13). Availability reflects the ease of access to a victim by an offender (13). Desirability refers to the fantasy element of the crime, where the offender chooses a victim because the victim fills some element of his sexual fantasy (4,11,13,14).

Sebai chose his victims randomly, through surveillance, but had no prior personal connection with them. Once he identified a potential victim, he studied the elderly woman's habits and chose a time when the victim was alone. He would approach the victim at her residence, using the pretext of selling holy pictures. Then he would blitz attack the victim.

Harris identified potential victims while he walked through their neighborhoods. He did not know any of the victims personally and did not even live in the same neighborhoods as his victims. He
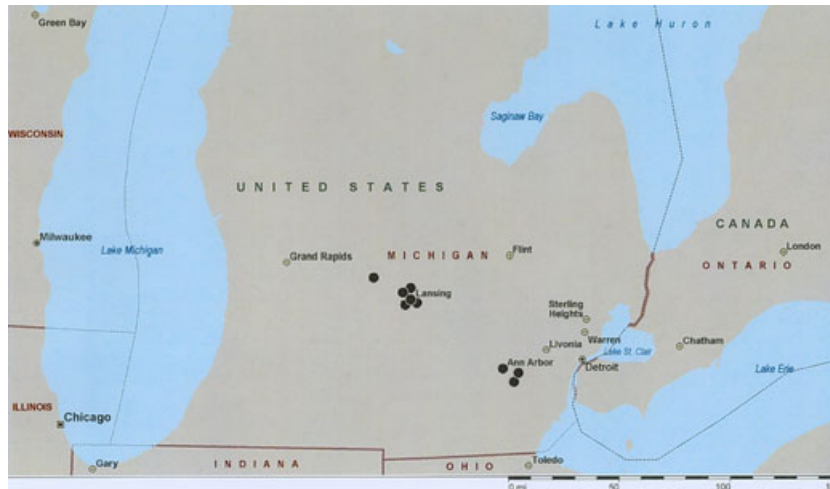
FIG. 4—*Distribution of Harris' victims in southern Michigan.*

would enter the apartments of the victims through unlocked doors and blitz attack them to gain control. He would then sexually assault and murder them.

Throughout other studies of sexually motivated murder, the most common causes of death are strangulation, stabbing or cutting wounds, and blunt force trauma (4,10,11,15). Sebai related that during his attacks, the neck was his favorite target because, according to him, stab wounds there produced rapid death. Harris utilized different means of killing: primarily ligature and manual strangulation; however, he also stabbed several other victims with a knife. He also battered his victims with his fists, inflicting serious blunt force injuries.

Historically, in cases of sexually motivated murder, the offender is acting out a sexual fantasy (2,10,13–15). The fulfillment of the fantasy is the driving force for his crime (2,10,13–15). The choice of a specific victim or type of victim, including age, race, sex, and physical appearance, can provide glimpses into the sexual preferences of the offender (14,15). The elderly women in both of these murder/rape series were specifically selected as victims because of the sexual desires of both offenders.

Sebai confessed to experiencing sexual gratification at every homicide. He admitted that during the course of each murder he would ejaculate, sometimes multiple times. Although Harris was uncooperative and made no admissions, there was physical

evidence that his victims were sexually assaulted. Two of his victims were also penetrated vaginally by a foreign object, which is a theme seen in other studies of sexually motivated murder (14,15). Additionally, 17 years after Harris had been incarcerated, he was found to possess explicit material that portrayed elderly women in a sexually provocative way.

In both cases, the primary motivation for the attacks was sexual in nature; the theft of objects from the crime scene was secondary. After each homicide, Sebai ransacked the apartment and took small items. Harris stole money and other objects from the victims after he assaulted and murdered them. In addition, Sebai and Harris collected copies of newspaper articles that dealt with their respective murders. Both of these activities have also been well documented in other studies of sexually motivated murderers (11,15).

In other studies of serial murderers, some offenders have had dysfunctional and abusive childhoods (1,2,4,9,15). According to Sebai, his father was physically and sexually abusive. He would strike Sebai on his feet after hanging him upside down, and he would also insert objects into his anus. Sebai related during his confession that his grandmother played a central role in his upbringing. He advised that she was overbearing and abusive, allowing him to be physically abused by his uncles. Sebai further confessed to having been sexually abused by his first Imam at his Muslim school and also in Turkish baths by "*black dressed*," old women, who were friends of his mother and grandmother. Although Sebai gave extensive details regarding his physical and sexual abuse, there is no

TABLE 3—*Harris' victims murdered and/or raped between 1980 and 1982 in southern Michigan.*

Case	Date	Name	Age	City	Cause of Death
1	10/18/1980	E R	77	Lansing	Asphyxia (ligature strangulation)
2	02/14/1981	E C	81	Lansing	Strangulation; combination of manual and ligature
3	11/02/1981	U C	77	Lansing	Asphyxia (characteristic of strangulation)
4	11/17/1981	B L	78	Portland	Smothering
5	11/30/1981	D S	78	Lansing	Asphyxia (manual strangulation)
6	01/08/1982	F B	91	Ypsilanti	Slash to throat
7	02/27/1982	D S	74	Lansing	Stab wound to neck
8	09/28/1982	M U	84	Ypsilanti	Ligature strangulation w/pantyhose
9	09/30/1982	L K	84	Ann Arbor	Ligature strangulation w/nylon stocking
10	12/04/1982	L N	68	Jackson	N/A (rape)

TABLE 4—*Elderly women that Harris was suspected of and/or convicted of their deaths/rape.*

Case	Date	Name	Age	City	Disposition
1	10/18/1980	E R	77	Lansing	
2	02/14/1981	E C	81	Lansing	
3	11/02/1981	U C	77	Lansing	
4	11/17/1981	B L	78	Portland	
5	11/30/1981	D S	78	Lansing	Two Life Sentences in 1983
6	01/08/1982	F B	91	Ypsilanti	
7	02/27/1982	D S	74	Lansing	
8	09/28/1982	M U	84	Ypsilanti	Two Life Sentences in 2003
9	09/30/1982	L K	84	Ann Arbor	Life Sentence in 2003
Rape	12/04/1982	L N	68	Jackson	60 years in 1983

independent confirmation or any official accounts or records of any such abuse.

Harris claimed he had a wonderful childhood up to the time he was 10 years of age when his family moved from Muskegon to Ann Arbor. He admitted that this was when his behavioral problems began. He further stated that his mother's boyfriend was physically abusive towards him, causing Harris to run away several times. There was, however, no independent confirmation of any abuse suffered by Harris.

The comparison of these two serial murder series demonstrates similar behavior from two culturally different offenders. The similarity in behavior is interesting in light of the significant differences in crime rates between the United States and Europe.

It is well documented that the United States has a much higher rate of murder than most countries in Europe (16,17). According to the United States Department of Justice, Uniformed Crime Reports (UCR) for 2004, the murder rate in the United States was 5.5 per 100,000 inhabitants, for a total of 16,137 murders (16). By contrast, murder rates in Europe are strikingly lower (17). According to United Nations Survey of Crime Trends and Operations of Criminal Justice Systems, murder rates in Europe in 2004 varied from .78 (Norway) to 2.94 (Scotland) per 100,000 inhabitants (17). Italy's murder rate in 2004 was 1.23 per 100,000 inhabitants (17). UCR data also reflect when a murder is committed in conjunction with another crime and under what general circumstances the murder occurred. However, neither the UCR nor the United Nations data comprehensively delineate the number of murders that were sexually motivated (11,16,17).

The empirical and predictive value of a single comparison, based upon only two serial murder series, is limited. However, there were a number of unique features shared by both of these crime series that reinforce previous studies (2,4,11,14,15). Both murder series were sexually motivated, and the offenders purposefully chose elderly women as their victims. Both offenders utilized similar methods to identify their victims, and both blitz attacked the victims in the victims' residences. There were differences in the type of sex acts engaged in by the offenders, but this is reflective of the diversity in sexual proclivities observed among other sexually violent offenders (2,4,14,15). Both offenders also stole objects from the victims' residences after the murders. In light of these apparent parallels, further cooperation and collaboration is needed to explore and compare similar cross-cultural issues in other violent, sexually motivated crimes.

References

1. Morton RJ, editor. Serial murder: multi-disciplinary perspectives for investigators. Washington, DC: United States Department of Justice, Federal Bureau of Investigation, 2007.
2. Schlesinger LB, editor. Serial offenders, current thought, recent findings. Boca Raton, FL: CRC Press, 2000.
3. McNamara JJ, Morton RJ. Frequency of serial sexual homicide victimization in Virginia for a ten-year period. J Forensic Sci 2004;49(3):529–33.
4. Hickey EW. Serial murderers and their victims, 3rd edn. Belmont, CA: Wadsworth Group, 2002.
5. Labuschange G. An empirical evaluation of interactional analysis of serial murder. Acta Criminologica 2000;13(2):22–30.
6. Brantley AC, Kosky RH. Serial murder in the Netherlands. FBI Law Enforcement Bulletin 2005;74(1):26–32.
7. O'Reilly-Fleming T. Serial and mass murder: theory, research and policy. Toronto, Canada: Canadian Scholar Press, 1996.
8. Roberts JV, Grossman MG. Sexual homicide in Canada: an integrative review. Aggress Violent Behav 2000;5(1):6.
9. Kraemer GW, Lord WD, Helibrun K. Comparing single and serial homicide offenses. Behav Sci Law 2004;22:325–43.
10. Malmquest CP. Homicide: a psychiatric perspective. Washington, DC: American Psychiatric Press, 1996.
11. Safarik ME, Jarvis J, Nussbaum K. Elderly female serial sexual homicide. Homicide Studies 2000;4(3):294–307.
12. Boudreaus MC, Lord WD, Jarvis JP. Behavioral perspectives on child homicide, the role of access, vulnerability, and routine theory. Trauma Violence Abuse 2001;2(1):56–76.
13. Morton RJ, McNamara JJ. Serial murder. In: Byard R, Corey T, Henderson C, Payne-James J, editors. Encyclopedia of forensic and legal medicine. Oxford, UK: Elsevier, 2005;47–53.
14. Meloy JR. The nature and dynamics of sexual homicide: an integrative review. Aggress Violent Behav 2000;5(1):1–32.
15. Ressler RK, Burgess AW, Douglas JE. Sexual homicide: patterns and motives. New York, NY: Lexington Books, 1992.
16. Uniform Crime Reports, 2004. Washington, DC: United States Dept. of Justice, Federal Bureau of Investigation, http://www.fbi.gov/ucr/ucr.htm#cius.
17. Ninth United Nations Survey of Crime Trends and Operations of Criminal Justice Systems. United Nations Office on Drugs and Crime, 2006; 1–9.http://www.unodc.org/pdf/research/9th_survey/CTS9ByIndicator Extract.pdf

Additional information and reprint requests:
Robert J. Morton, M.S.
NCAVC
FBI Academy
Quantico, VA 22135
E-mail: robert.morton@ic.fbi.gov

CASE REPORT
GENERAL; TOXICOLOGY

Kathleen F. Gensheimer,[1],[†] *M.D., M.P.H.; Vicki Rea,*[1] *R.N., M.P.H.; Dora Anne Mills,*[1] *M.D., M.P.H.;*
Christopher P. Montagna,[1] *M.S., M.P.A.; and Karen Simone,*[2] *Pharm.D.*

Arsenic Poisoning Caused by Intentional Contamination of Coffee at a Church Gathering—An Epidemiological Approach to a Forensic Investigation

ABSTRACT: An outbreak of apparent food-borne illness following a church gathering was promptly reported to the Maine Bureau of Health. Gastrointestinal symptoms among church attendees were initially attributed to consumption of leftover sandwiches that had been served the previous day. However, a rapid epidemiological and laboratory assessment revealed the etiology of illness, including the death of an elderly gentleman, was not food-borne in origin. A criminal investigation determined that deliberate arsenic contamination of the brewed coffee by one of the church members was the source of the outbreak. Public health officials and criminal investigators must be aware that intentional biologic aggression can initially present as typical unintentional disease outbreaks. Practitioners must also consider the need to properly maintain and preserve potential forensic evidence. This case demonstrates the key role public health practitioners may play in criminal investigations.

KEYWORDS: forensic science, arsenic, terrorism, epidemiology, poisoning, emergency response

With the result of the infusion of federal funding to enhance public health infrastructure following the events of September 11, 2001, the Maine Center for Disease Control (MeCDC), formerly the Maine Bureau of Health (BOH), instituted a formal 24-h toll-free after hours on call emergency and consultation system (AHOC) to respond to reports of infectious disease and other public health emergencies. The system was critical in the response to a call received on Sunday evening, April 27, 2003, when the Infection Control Practitioner of a small community hospital notified the BOH that six people had presented to the hospital emergency room throughout the afternoon with vomiting and diarrhea. All six patients had attended a social gathering held at their church that morning. Leftover food from the previous day's church bake sale, including turkey, ham, and egg sandwiches and assorted cookies had been served. Later that evening, an additional six people with similar gastrointestinal symptoms, all of whom attended the Sunday morning gathering, sought medical attention at the emergency room. Four of the twelve patients were admitted to the hospital; three of them were transferred to a regional medical center for intensive care. Within 16 h of symptom onset, two patients developed multisystem organ failure and one patient died. An epidemiological investigation was initiated at the time of receipt of the disease report. Because of the apparent short incubation period with

the possibility of a chemical toxin being the causative agent and initial findings that the coffee may have been the vehicle of illness, the cases were also reported to the Northern New England Poison Control Center (NNEPC), which offered a differential diagnosis and recommended directed laboratory testing.

Methods

The church, where the social gathering took place, is located in a rural community (population 621) and had no designated minister or leader. Church members were asked to provide the names of people who had attended the Saturday church bake sale, the Sunday morning gathering, a list of foods served, and a general description of the Sunday gathering. All attendees and family members were interviewed using a standard questionnaire regarding food and beverages consumed during either the Saturday church bake sale or the Sunday morning gathering, signs and symptoms of illness, and onset and duration of illness.

Food and drink preparation areas and equipment were inspected by a health sanitarian. The preliminary epidemiologic investigation found illness to be associated with attendance at the Sunday morning gathering. A case was defined as a person who attended the morning gathering on April 27 and developed vomiting or diarrhea (defined as three or more stools in a 24-h period) within 12 h of the gathering. Stool and emesis specimens were obtained from ill attendees and tested for *staphylococcus, salmonella, shigella,* and *campylobacter.* Emesis specimens were tested for staphylococcal toxin and arsenic. Coffee, blood, and urine specimens were tested for arsenic.

Data on clinical characteristics of cases who sought medical attention were obtained from medical records. Samples of all remaining food and drinks served, including discarded food remains

[1]Maine Center for Disease Control and Prevention, Maine Department of Health & Human Services, Station #11 Augusta, ME 04333.
[2]Northern New England Poison Center, 22 Bramhall Street, Portland, Maine 04102.
[†]Present address: sanofi pasteur, 38 Sidney Street, Cambridge, MA 02139.

and drinking water, were obtained for laboratory analysis. Food and drink preparation areas and equipment were inspected. Air quality inside the church and the premises was tested for evidence of carbon monoxide, low oxygen, sulfur dioxide, and the lower explosive limit for substances including propane. Data from the food intake and symptom questionnaire were analyzed using Epi-Info 2002 (1). Relative risks for illness and corresponding 95% confidence limits were calculated for food and drink served at the gathering.

Despite the incident's initial classification as a routine food-borne investigation, because of the clinical suspicion that the patients' symptoms might be caused by a chemical toxin and early findings that the coffee may have been the vehicle for illness, the Criminal Investigation Division of the Maine State Police was notified on April 28. A chain of custody was established employing BOH protocols over clinical and environmental specimens submitted to the Maine Health and Environmental Testing Laboratory (HETL). The federal Centers for Disease Control and Prevention (CDC) and the Department of Homeland Security were notified of the outbreak on April 28, and surveillance for similar symptoms throughout the state and nation was initiated to evaluate whether the outbreak could have been part of a more widespread biologic attack. Ongoing surveillance for additional patients with acute gastrointestinal disease or toxic syndromes was established at the hospital.

Results

Approximately 80 people attended the Saturday bake sale, and 27 persons attended the Sunday gathering. Eight people attended both events. Remaining food from the bake sale, including cookies, egg, ham, and tuna salad sandwiches, was refrigerated in the church and served with coffee at the Sunday morning gathering. Coffee was freshly brewed in a 30-cup canister coffee maker for both events.

The median age of the Sunday gathering attendees was 58 (range 1–79); 56% were men and 46% women. Thirteen attendees (48%) became ill with vomiting and diarrhea within 1 h following the gathering (Fig. 1). Consumption of coffee was the only factor significantly associated with illness (RR 8.9, 95% CI 1.4–58.8) (Table 1). The median incubation period from 10:30 AM, when coffee consumption began to onset of illness, was 45 min.

The coffee was prepared by church members prior to the church service on the morning of April 27. The container of the coffee urn was filled with water from the church kitchen tap. Coffee grounds were taken from an opened 3 lb can of commercial coffee stored in the church's kitchen refrigerator and placed in the filter in the top of the coffee maker. The coffee brewed during the church service was ready for consumption at 10:30 AM when the Sunday morning social gathering began. Coffee and leftover bake sale goods were offered to the attendees. A total of 16 people consumed coffee. The amount of coffee consumed ranged from one sip to 30 sips (median 3 sips). Several of the coffee drinkers mentioned that
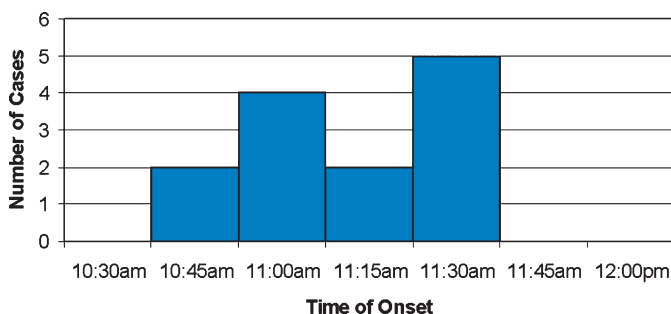

FIG. 1—*Time of symptom onset.*

the coffee tasted so bitter that they refrained from drinking more coffee. Thirteen (81%) coffee drinkers sought medical attention on April 27. Eight of these were hospitalized, and five were evaluated and discharged home. One person died 16 h after drinking coffee. Three persons who consumed coffee did not report any illness.

Stool and emesis samples from 12 cases were cultured for bacterial pathogens by the Maine HETL. Cultures of emesis yielded a pure growth of *Staphylococcus aureus* from one case and scarce *S. aureus* from another case; all other cultures yielded either no growth or diverse bacterial flora. Blood was cultured from eleven cases; one grew diphtheroids, and all others showed no bacterial growth. Samples of discarded sandwiches tested by the federal Food and Drug Administration showed no evidence of staphylococcal enterotoxin.

Elevated blood arsenic concentrations were noted in eight cases on clinical specimens collected on April 27 with a range of 68–192 µg/L (reference range 60 or less µg/L). Three people, who drank coffee and were not ill, had a range of blood arsenic concentration of 0–14 µg/L. Emesis samples from two cases had arsenic concentration of 190 and 240 µg/L.

Analysis conducted by the HETL on April 27 and 28 found arsenic (6300 mg/L) in a leftover cup of coffee found in the church and in residual coffee from the brewing pot (3500 mg/L). Unbrewed coffee from the leftover can of commercial coffee found at the church had no detectable arsenic. The water source, a dug well on the church property, was reported to have arsenic detected below the safe drinking water level of <0.015 mg/L. Water taken from the sink tap at the church had no detectable arsenic. Inspection of the church found no source of arsenic on site. The church is located in a rural potato-farming area where arsenic was extensively used in the past as a pesticide and herbicide up until the late 1960s. Although no longer used, containers of arsenic compounds may still be found in local barns and other storage facilities. Attempts at speciation of the form of arsenic were inconclusive.

Based upon the findings of the HETL, the focus of the event switched from a public health investigation to a criminal investigation. Despite the incident's initial classification as an epidemiological event, proper chain of custody was established employing BOH protocols. As a result, no problems were encountered in retroactively establishing the chain of custody for the clinical samples as related to criminal justice purposes.

Discussion

The report of this outbreak initially appeared similar to one of innumerable reports of illness following the classic church supper scenario that has been used to teach basic epidemiologic methods: a rural church gathering, where food that had been prepared a day earlier was served at room temperature, providing a likely vehicle for bacterial overgrowth and production of bacterial toxins. The epidemic curve, Fig. 1, demonstrates that the majority of symptoms began while the contaminated coffee was still being served. The shape of the curve indicates a common source outbreak because of a noncontagious agent. Staphylococcal food poisoning was initially considered the etiological agent when one culture of emesis yielded a pure growth. However, the clinical presentation of severity of illness and rapid onset of symptoms associated with the epidemiologically incriminated coffee was more consistent with chemical intoxication rather than the toxic or infectious gastroenteritis that typically results from bacterial or viral contamination of food.

Consultation with both the Maine CDC and the NNEPC was initiated within several hours of patient presentation to the local

TABLE 1—*Food-specific attack rates for persons who attended church gathering: Maine, 2003.*

Food Item	# Cases Who Consumed	Total # Who Consumed	Attack Rate of Those Who Consumed (%)	# Cases Who Did Not Consume	Total # Who Did Not Consume	Attack Rate of Those Who Did Not Consume (%)	Risk Ratio	95% Confidence Interval
Coffee	13	16	81.3	1	11	9.1	8.9	1.4–58.8
Punch	2	6	33.3	12	21	57.1	0.6	0.2–1.9
Cookies	2	7	28.6	12	20	60.0	0.5	0.1–1.6
Egg salad sandwich	1	2	50.0	13	25	52.0	1.0	0.2–4.0
Ham salad sandwich	2	4	50.0	12	23	52.0	1.0	0.3–2.8
Tuna salad sandwich	3	7	42.9	11	20	55.0	0.8	0.3–2.0

hospital. The AHOC system resulted in prompt receipt of the disease report by the state public health agency. The clinical features in acute arsenic poisoning include the nonspecific syndrome of nausea, vomiting, abdominal pain, and diarrhea (2,3), Despite the initial assumption that this was a typical food-borne outbreak, a rapid epidemiologic and laboratory response confirmed, <24 h within receipt of the outbreak report, that arsenic-contaminated coffee was the source of illness. Prompt reporting of the outbreak with rapid identification of the etiology of illness resulted in immediate mobilization and transport of 100 vials of British Anti-Lewisite (BAL) from the state stockpile to the local hospitals where symptomatic patients were being treated. Such action undoubtedly proved to be lifesaving and prevented more serious sequelae. All 15 surviving individuals who consumed the coffee were treated with chelation therapy agents including dimercaptopropanol or BAL, dimercaptosuccinic acid, and dimercaptopropane sulfonic acid protocols in consultation with the NNEPC. The NNEPC played a key role in diagnosing, locating, and facilitating the transport of antidotes and ongoing management of the patients involved.

Plausible risk factors for illness were rapidly identified employing standard epidemiologic techniques including: establishment of a case definition, case finding, identification of a study population, generation of hypotheses, collection of clinical and food samples, and implementation of an analytic study. These steps helped rule out infectious causes and yielded an elevated relative risk of illness associated with drinking coffee at the Sunday morning social. The concern expressed by patients regarding the unusual taste of the coffee also helped focus investigative efforts on coffee as a possible source of illness. The church is located in a rural potato-farming area where arsenic was extensively used in the past as a pesticide and herbicide. Elevated arsenic concentrations in patients' emesis, blood, and in brewed coffee initially raised concern that the outbreak might have resulted from accidental poisoning. As arsenic was an active ingredient in chemical fertilizer used in this farming community in the 1960s, it was critical to make a rapid determination as to criminal versus unintentional poisoning.

The finding of arsenic in brewed coffee but not in unbrewed coffee grounds indicated the coffee was contaminated with arsenic during the brewing process. There was no indication that accidental contamination of the coffee had occurred. A criminal investigation into the outbreak determined the arsenic contamination was the result of intentional poisoning. The person responsible for the poisoning was determined to be a church member who committed suicide 5 days after the incident. The church member left behind a suicide note implicating his involvement with the poisoning. The criminal investigation was closed in 2006.

Conclusion

Central to the dogma of the practicing epidemiologist is to evaluate all facts and information with diagnostic support provided by an experienced laboratory. The public health response to this outbreak serves as a model for responding to the intentional release of chemical or biologic agents as acts of terrorism. As in this situation, public health agencies will often be the first responders to reports of unusual clusters or outbreaks of illness. The implementation of standard epidemiologic techniques provides useful information on the scope, source, and etiology of an outbreak and helps focus further investigation into the most likely cause. A coordinated response with public safety and criminal investigative teams to establish chain of custody for testing biological, physical, and environmental samples is critical to ensure the integrity of evidence for both investigation and prosecution of criminal cases.

Such collaboration can provide crucial information on the cause of the outbreak that might not be forthcoming from the public health response alone. As in this case, a rapid coordinated response by public health, medical providers, and public safety agencies provided an effective means to identify the source of an intentionally caused disease outbreak, preventing further illness from occurring and promoting utilization of life-saving agents.

Deliberate contamination of food during its production and preparation has been reported infrequently (4,5). Unintentional contamination of food by chemicals, including reports of contamination by pesticides, occurs sporadically (6,7). Methods of analysis for the intentional cases involved restriction endonuclease digestion and pulse-field gel electrophoresis. Analytical methods used in the unintentional cases were not cited. Vigilance and heightened awareness is required on the part of the health care community and public health agencies that investigate reports of food-borne illnesses to ensure that unusual events are quickly identified, reported, investigated, and properly mitigated.

In this incident, two ongoing investigations with different goals were conducted simultaneously. The BOH worked to determine the source of illness, the extent of contamination, assure treatment of those exposed, and to limit further exposure. The goal of the Maine State Police was to determine whether criminal activity was involved and if so, to identify, apprehend, and prosecute the perpetrator. The modes of operation of public health and criminal investigation agencies are often different and sometimes conflict. For example, the state public health agency felt communication to be critical to inform the public of any possible threat, whereas the investigative team was opposed to any public release of information out of concern for introducing bias or disrupting the criminal investigation underway. In contrast, law enforcement officials released the name of the patient who died, whereas public health practitioners adhere strictly to patient confidentiality and do not identify patients' names during or after the conduct of an investigation. Despite the conflicting roles of the agencies involved, the overall investigation was coordinated, information shared, and conclusions regarding the cause of illness were rapidly identified, allowing life-saving intervention.

Multiple agency coordination and cooperation between public health and law enforcement officials at the local, state, and federal levels are critical for the detection and response to similar events, whether they are intentional or unintentional. This event speaks to the need of a solid public health infrastructure that supports a robust surveillance and laboratory diagnostic system, the value of clinical awareness for prompt reporting to public health agencies, and to the power of exercising complementary skills between invested public health surveillance and response activities. This investigation benefited from such a coordination which contributed to a rapid response to a serious event.

References

1. Dean AG, Arner TG, Sunki GG, Friedman R, Lantinga M, Sangam S, et al. Epi Info TM, a database and statistics program for public health professionals. Atlanta, GA: Centers for Disease Control and Prevention, 2002.
2. Stenehjem AE, Vahter M, Nermell B, Aasen J, Lierhagen S, Morland J, et al. Slow recovery from severe inorganic arsenic poisoning despite treatment with DMSA. Clin Toxicol 2007;45:424–8.
3. Lai MW, Boyer EW, Kleiman ME, Rodig NM, Ewald MB. Acute arsenic poisoning in two siblings. Pediatrics 2005;116:249–57.
4. Török TJ, Tauxe RV, Wise RP, Livengood JR, Sokolow R, Mauvais S, et al. A large community outbreak of salmonellosis caused by intentional contamination of restaurant salad bars. JAMA 1997;278:389–95.
5. Kolavec SA, Kimura A, Simons SL, Slutsker L, Barth S, Haley CE. An outbreak of *Shigella dysenteriae* type 2 among laboratory workers due to intentional food contamination. JAMA 1997;278:396–8.
6. CDC. Endrin poisoning associated with taquito ingestion—California. MMWR 1989;38(19):345–7.
7. CDC. Epidemiologic notes and reports aldicarb food poisoning from contaminated melons—California. MMWR 1986;35(16):254–8.

Additional information and reprint requests:
Kathleen F. Gensheimer, M.D., M.P.H.
Director Scientific and Medical Affairs
sanofi pasteur
38 Sidney Street
Cambridge, MA 02139
E-mail: Kathleen.Gensheimer@sanofipasteur.com

CASE REPORT

PATHOLOGY AND BIOLOGY

Claudia Liuzzi,[1] M.D.; Felice Carabellese,[2] M.D.; and Francesco Vinci,[1] M.D.

Perineal-Vaginal Injuries in Children: Accident or Abuse?

ABSTRACT: A large number of conditions have been mistaken for abuse. Differentiating accidental injuries from inflicted injuries is important in the management of injured children. In this work, the authors describe two cases of accidental perineal-vaginal injury in children. In case 1, a 4-year-old girl suffered a vaginal tear caused by violent stretching during play; in case 2, a 3-year-old girl had minor lacerations of labia minora and majora. The intervention of a multi-specialist team including a forensic pathologist and forensic psychiatrist was the key factor in being able to exclude abuse by third parties in the described cases.. This leads to the appropriate recommendations to be adopted in the forensic medicine setting.

KEYWORDS: forensic science, forensic pathologist, child sex abuse, accidental perineal-vaginal injury

The management of suspected child sex abuse involves various professionals; indeed, the diagnosis itself requires the intervention of a multi-specialist team such as a forensic pathologist, a pediatrician, a gynecologist, a psychiatrist, etc. An etiological diagnosis based only on a physical examination is often difficult because physical signs are not always present and, when they are present, can be ambiguous. Genital and paragenital injury can be accidental and so not necessarily an indication of sexual abuse. Distinguishing accidental from abusive causes of genital trauma is often critical. Here, we report two cases of accidental injury mimicking sexual trauma.

Case 1

After returning home from kindergarten a 4-year-old girl had blood spots on her underwear. As no explanation was forthcoming for the origin of the bloodstains, the mother took her daughter to the local hospital emergency department. The gynecological examination revealed a vaginal-perineum tear with blood infiltration at the 6 o'clock position directly below the hymen insertion. Suspecting abuse by means of penetration or attempted penetration (e.g., by fingers or some other object), the intervention of a forensic pathologist was requested, who, together with a psychologist expert in this field, questioned the mother and child to obtain the historical information. This inevitably led to the kindergarten and the people who had been in direct contact with the child. Questioning of the four kindergarten teachers, all of whom were qualified professionals and respected for their morals, was unsuspicious. All four stated that at all times the child had been playing together with her 25 classmates and under the supervision of the teachers.

The investigations carried out by the medical-legal team and by the authorities were totally consistent, as was the information given by the mother, the child, the teachers, and the child's playmates. Having excluded the possibility of any third party sexual activity and taking into account the stable family history, the law enforcement officers concluded that the injuries were caused accidently during play activity which including doing the "splits." The rapid abduction of the girl's legs had caused an excessive traction of the tissues so exceeding their elasticity and resistance (Fig. 1).

Case 2

A 3-year-old girl with vaginal trauma was brought to the local hospital emergency department by her parents who reported that they had been washing the girl in the sink when it had collapsed and she had hit her vulvar region against the tap. Physical

[1]Section of Legal Medicine, University of Bari, P.zza Giulio Cesare, 11, 70124 Bari, Italy.
[2]Section of Forensic Psychiatry, University of Bari, P.zza Giulio Cesare, 11, 70124 Bari, Italy.

FIG. 1—*Case 1: vaginal-perineum tear at 6 o'clock below the hymen insertion.*

examination revealed a small tear with blood infiltration on the left labia minora at 11 o'clock and a bruised swelling of the opposite labia majora (Fig. 2).

Suspecting sexual abuse by penetration of one or more fingers, the intervention of a forensic pathologist was requested, who together with a psychologist then questioned the parents and the girl to obtain a clear and a consistent history of the event. The subsequent visit to the house by the forensic pathologist and law enforcement officers revealed the remains of the broken sink.

Discussion

The general definition of child abuse is "*all forms of physical and/or emotional ill treatment, sexual abuse, neglect or negligent treatment or commercial or other exploitation, resulting in actual or potential harm to the child's health, survival, development or dignity in the context of a relationship of responsibility, trust or power.*"

Abuse is therefore identifiable as physical, sexual, or psychological trauma arising from the actions of an adult on a child. It has been estimated that each year around 40 million children are subject to this type of violence. The differential diagnosis of findings suggestive of child abuse is very difficult, especially in cases of sexual abuse (1).

Apart from nonabuse cases involving congenital problems with the genital-urinary apparatus, for example erythema of the vestibule, periurethral bands, labial adhesion, urethral dilatation, vaginal ridges (2), in the literature there are reports of accidental mechanisms that could be erroneously interpreted as a result of sexual abuse when in fact they are accidental, as in the two cases here described (3–5).

Various surgical cases have been described in the repair of perineal and rectal impalement injuries even from gunshot (6–8).

Scheidler describes 358 female pediatric patients observed between 1993–1997 with contusion secondary to perineum trauma as a result of motor vehicle collisions, falls, and bicycle-related incidents (prevalent in children under 9 years of age), assaults (prevalent in children between 0–4 years), or use of playground-related equipment (9).

In cases of suspected child sex abuse after examination in the hospital emergency department, the intervention of a multi-specialist team, including a psychologist, is required, which then proceeds to conduct a psychological and physical exam (Table 1) (10).

In the diagnosis of child sex abuse cases, it is also fundamental to inspect the place where the event took place, after having gathered complete information from the child and its parents or carers with the help of a psychologist trained in this field. Particular attention must be given to any inconsistencies between the history of event as told by the parents and the most probable etiology of the trauma. In identifying a probable abuse case, there are various indicators that can be gleaned from an accurate family history and from the child's health history and his or her current status. These are: attempts to minimize the gravity of the injuries; delays in seeking medical assistance; reports of substance abuse by the parents; mental illness of one or both parents; signs of neglect or malnutrition in the child; vaccinations not up-to-date. The child also is an important source of information that is best obtained without the adult caregivers present. This privacy may allow the child to be more candid about the injuries. The child's behavior can show sudden changes in schoolwork or socializing or in eating. Others signs can be extreme withdrawal or aggressiveness or self-destruction, or inappropriate sex play or premature understanding of sex.

The goal is to gather as much information as possible from as many different sources (family history, health and development of the child, behavior indicators, etc.) as well as the history of the event. These steps are necessary to allow the subsequent investigation to be thorough (11–14).

The multidisciplinary approach to information gathering for child sex abuse cases requires specialized knowledge in forensic pathology, forensic psychiatry, and criminology, aligned with the procedural protocols. We must also remember that the entity and type of trauma depends on the sexual practice employed (object use, vaginal or anal coitus), the resistance of the victim (practically none in children and the elderly), violent and/or sadistic behavior by the aggressor, which could or could not be indicative of a psychopathological disorder. The forensic pathology investigation is aimed at achieving a precise reconstruction of the traumatic event based on the environment in which it occurred and the objects there present (11,15).

The age and sex of the victim is important in determining the precise methodology to be adopted in sex abuse cases. For female pediatric patients, sexual abuse is highly probable if the examination of the genital or perigenital regions reveals hyperthemia, ecchimosis, or laceration of the vulvo-vestibular mucous and (in virgins) of the hymenal membrane, of the clitoris, and/or of the vagina.

The presence of such signs varies in function of the victim's age, and Gisbert Calabuig (1991) published a classification for sexual abuse injuries (Table 2) caused by penis penetration in relation to the different developmental stages in girls (16).

FIG. 2—*Case 2: tear with blood infiltration on the left labia minora at 11 o'clock and a bruised swelling of the opposite labia majora.*

TABLE 1—*(adapted from James A Monteleone, 1998): Conditions to note during the physical examination.*

Evidence of other abuse (loop marks and scars)	Semen on clothing
General hygiene	Grasp marks
Scars form other injuries	Blood stains and/or discharge stains on underwear
Nutritional status; including height and weight	Absent or suppressed gag reflex
Anatomic or congenital defects	Extreme compliances, child emotionally distant during examination
Difficulty in walking or sitting	Child seductive

TABLE 2—*(adapted from Gisbert Calabuig, 1991): Age of female pediatric victim, possibility of vaginal penetration and injury.*

Age	Vaginal Penetration	Lesions
<6 years	Impossible	Probable nonspecific paragenital injury
6–11 years	Possible	Hymen laceration
		Minimal genital or paragenital injury
11 years	Possible	Serious vulvo-vaginal, perineal, and rectal laceration
>18 years (virgin)	Present	Possible hymenal laceration

In girls under 6 years of age carnal penetration is impossible for anatomic reasons: the subpubic angle is still very acute and therefore creates a true bone barrier which obstructs penetration by the penis. In girls between the ages of 6 and 11 years, because of the disproportion in size of the genital organs, carnal penetration causes serious tears with serious bleeding to the vulvo-vaginal walls, to the perineum, and to the rectal walls.

From 11 years of age until menarche there is hymenal laceration and minimal trauma to the genital and paragenital areas. In adult virgins, penetration can cause only hymenal alteration dependent on the anatomical variations of the membrane, unless there is particular violence used (16,17).

In the cases here presented, the application of the correct procedures provided clear and detailed histories of the nature of the mechanism of the trauma that caused the injuries. Differential diagnosis in cases of suspected child sex abuse is a demanding task and can be arrived at only through the intervention of a multi-specialist team coordinated following the international guidelines regarding child sex abuse (1,18–20). The gathering of both historic and objective information before the subsequent investigation of where the event happened allows the correct interpretation of the place and the objects found there, giving the possibility to confirm or negate an accidental manner of injury.

Acknowledgment

References

1. World Health Organization. Managing child abuse. A handbook for medical officers. New Delhi: World Health Organization, Regional Office for South-East Asia, 2004.
2. Berkoff MC, Zolotor AJ, Makoroff KL, Thackeray JD, Shapiro RA, Runyan DK. Has this prepubertal girl been sexually abused? JAMA 2008;300(23):2779–92.
3. Boos SC. Experience and reason: accidental hymenal injury mimicking sexual trauma. Pediatrics 1999;103(6):1287–90.
4. Heppenstall-Heger A, McConnell G, Ticson L, Guerra L, Lister J, Zaragoza T. Healing patterns in anogenital injuries: a longitudinal study of injuries associated with sexual abuse, accidental injuries, or genital surgery in the preadolescent child. Pediatrics 2003;112(4):829–37.
5. Boos SC, Rosas AJ, Boyle C, McCann J. Anogenital injuries in child pedestrians run over by low-speed motor vehicles: four cases with findings that mimic child sexual abuse. Pediatrics 2003;112(1 Pt 1):e77–84.
6. Reinberg O, Yazbeck S. Major perineal trauma in children. J Pediatr Surg 1989;24(10):982–4.
7. Joos AK, Herold A, Palma P, Post S. Perianal and rectal impalement injuries. Chirurg 2006;77(9):781–9.
8. Grisoni ER, Hahn E, Marsh E, Volsko T, Dudgeon D. Pediatric perineal impalement injuries. J Pediatr Surg 2000;35(5):702–4.
9. Scheidler MG, Shultz BL, Schall L, Ford HR. Mechanism of blunt perineal injury in female pediatric patients. J Pediatr Surg 2000;35(9):1317–9.
10. Monteleone JA. Quick reference child abuse. St. Louis, MO: G. W. Medical Publishing Inc., 1998.
11. Vimercati F, Vinci F, Dell'Erba A. Approccio metodologico alle indagini medico-legali e criminalistiche in tema di decessi conseguiti ad attività a sfondo sessuale. Jura Medica 1997;2:135–62.
12. Makoroff KL, Brauley JL, Brandner AM, Myers PA, Shapiro RA. Genital examinations for alleged sexual abuse of prepubertal girls: findings by pediatric emergency medicine physicians compared with child abuse trained physicians. Child Abuse Negl 2002;26(12):1235–42.
13. Heger A, Ticson L, Velasquez O, Bernier R. Children referred for possible sexual abuse: medical findings in 2384 children. Child Abuse Negl 2002;26(6–7):645–59.
14. Kelly P, Koh J, Thompson JM. Diagnostic findings in alleged sexual abuse: symptoms have no predictive value. J Paediatr Child Health 2006;42(3):112–7.
15. Brown DA, Lamb ME. Helping abused children talk about their experiences in forensic interviews. Minerva Med 2006;126(2):155–68.
16. Gisbert Calabuig JA. Medicina legal y toxicologia, 5th rev. edn. Barcelona: Masson, 1998.
17. Edgardh K, Von Krogh G, Ormstad K. Adolescent girls investigated for sexual abuse: history, physical findings and legal outcome. Forensic Sci Int 1999;104(1):1–15.
18. American Academy of Pediatrics Committee on Child Abuse and Neglect. Guidelines for the evaluation of sexual abuse of children: subject review. Pediatrics 1999;103(1):186–91.
19. Kellogg N. American Academy of Pediatrics Committee on Child Abuse and Neglect. The evaluation of sexual abuse in children. Pediatrics 2005;116(2):506–12.
20. Kellogg ND. American Academy of Pediatrics Committee on Child Abuse and Neglect. Evaluation of suspected child physical abuse. Pediatrics 2007;119(6):1232–41.

Additional information and reprint requests:
Claudia Liuzzi, M.D.
University of Bari
Section of Legal Medicine
P.zza Giulio Cesare, n. 11
70124 Bari
Italy
E-mail: claudialiuzzi77@hotmail.com

CASE REPORT
PATHOLOGY AND BIOLOGY

Bryan C. Patonay,[1,2] M.D. and William R. Oliver,[2] M.D.

Can Birth Trauma Be Confused for Abuse?

ABSTRACT: An unexpected infant death is usually investigated with a complete autopsy. If evidence of prior trauma is found at autopsy in these cases, suspicion is raised for nonaccidental trauma. In a young infant, the residua of trauma received during birth has the potential to be incorrectly interpreted as nonaccidental trauma. We report a the findings of a 4 1/2-month-old-infant that died unexpectedly with a healing linear skull fracture and a circular lesion over the calvarium found at autopsy. Though this lesion was concerning, the remainder of the autopsy and the histological findings did not support a diagnosis of recent trauma. Review of the literature describing birth injuries made the diagnosis of healing, residual birth trauma more convincing in this case.

KEYWORDS: forensic science, forensic pathology, birth injuries, histology, child abuse, sudden infant death, autopsy

Fetal injuries obtained during birth are not uncommon and range in severity from caput succedaneum to clinically apparent depressed skull fracture (1). Such injuries are observed in spontaneous vaginal deliveries, instrument deliveries (2–5), and even rarely in the setting of cesarean section deliveries utilizing forceps (6). Unfortunately, finding birth trauma at autopsy can complicate the investigation of an unexpected infant death. Because it is important to rule out a nonaccidental cause of death in such circumstances, the presence of residua of birth trauma may sway the pathologist to incorrectly diagnosis nonaccidental trauma (7).

Case Report

The decedent, a 4 1/2-month-old boy, was born at 35 weeks gestational age as part of a twin gestation. Delivery was via cesarean section after prolonged labor. The decedent was delivered before his twin sister. Per the operative note, the procedure was uncomplicated and the infant was delivered without any difficulty or need for instrumentation. After delivery, the neonate was noted to have some pallor and truncal ecchymoses with an Apgar score of 8 at 1 min and 9 at 5 min. Otherwise, the decedent was appropriate for gestational age. The birth records do not mention any postpartum clinical findings that would be consistent with birth trauma such as cephalohematoma or palpable skull fracture. After discharge home, development was unremarkable and appropriate pediatric preventative care was received. Other than an uneventful emergency room visit on the 28th day of life for a choking episode, the infant's medical history was unremarkable until his death.

On the day of his death, the infant was put down for a nap. Sometime, later, when his mother checked on him, he was not breathing. Paramedics were called and resuscitation was attempted.

However, the child was declared dead on arrival to the hospital. Law enforcement officials investigated with interview of the caretakers and by examination of the home where the infant died. These investigations did not uncover evidence of abuse or neglect. The deceased infant's hospital records were reviewed at the time of autopsy; no admissions were ever made for traumatic injury. The decedent's twin sister and 8-year-old sister are healthy. There is no documentation of any genetic diseases in the infant's family.

At autopsy, the only remarkable findings were on the calvarium. One finding was a variegated, circular, purple-red discoloration on the outer table of the left parietal bone that measured approximately 5 cm in maximum diameter (see Fig. 1). This lesion was apparent on the inside of the skull as well (see Fig. 2). The other finding was a linear, well-healed fracture visible on the inner table of the left occipital bone that stretched from anterior to posterior for a total of 5 cm (see Fig. 3). There was no evidence of recent trauma in the scalp. No hemorrhage, hematoma, or brain contusions were seen upon opening the scalp. On histology, the sections of the linear fracture had abundant fibrous tissue with old, extravasated

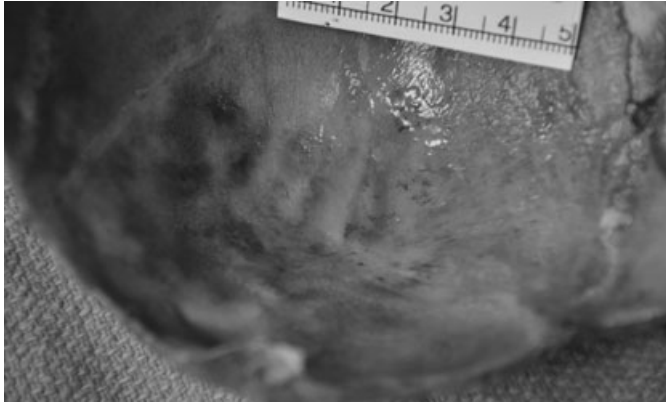FIG. 1—*Circular lesion over the left occiput seen from the external aspect of the calvarium.*

[1]Pitt County Memorial Hospital, Pathology and Laboratory Medicine, 2100 Statonsburg Road, PO Box 6028, Greenville, NC 27835.
[2]Brody School of Medicine, East Carolina University, Pathology and Laboratory Medicine, Brody Medical Science Building, Greenville, NC 27834.

FIG. 2—*Circular lesion seen from the internal aspect of the calvarium.*
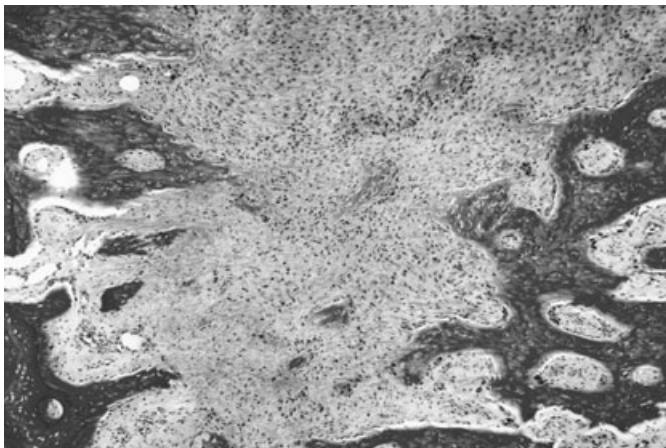

FIG. 3—*Linear skull fracture (left, next to ruler).*

FIG. 4—*Photomicrograph of the linear fracture. (Original magnification 40×; H and E).*

erythrocytes and a paucity of hemosiderin-laden macrophages on iron stains (see Fig. 4). Histologically, the appearance is similar to the mechanical phase of craniofacial bone healing described by Gannon et al. (8). The circular lesion showed an abundance of fibrous tissue with some old red blood cells outside of the bony part of the calvarium, corresponding to an expansion of the pericranium (see Fig. 5). As cephalohematomas are sub-periosteal lesions,
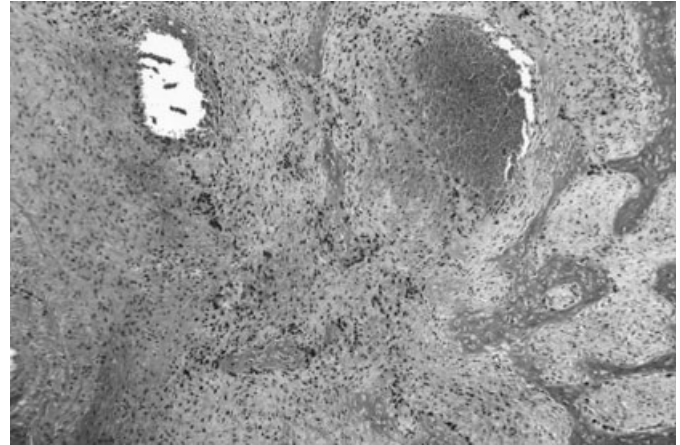
FIG. 5—*Photomicrograph of the circular lesion. (Original magnification: 40×; H and E).*

the location of the abnormal tissue in the circular lesion is consistent with a healing cephalohematoma, which are known to be grossly visible even 3 months after birth (1). There was no fresh hemorrhage as would be expected in death because of traumatic head injury. No evidence of vascular malformation, such as a hemangioma, was seen. In the absence of any other anatomic finding, we ultimately diagnosed the cause of death as sudden infant death syndrome.

Discussion

Certain injuries, when seen in young children, may raise the suspicion for child abuse (9). However, many of these "red flag" injuries can occur as a result of birth trauma. These include posterior rib fractures (10), fracture of the clavicle (11), fractures of the femur and humerus (12,13), scalp hemorrhage, subgaleal hemorrhage, cephalohematoma, skull fracture (14), occipital osteodiastasis, intracranial hemorrhage, subarachnoid hemorrhage, and subdural hemorrhage (1). Cephalohematomas can occur in up to 2% of live births and 9.4% of forceps-assisted deliveries (1). In one series of 3340 live births in Iran, birth trauma was observed in 0.8% of deliveries, and cephalohematomas made up 37.5% of these, though they were only seen in low-birthweight children (15).

In our case, we noted the presence of both a periosteal discoloration and healed skull fracture. Because of the presence of two distinct injuries, we were initially suspicious for a nonaccidental etiology for these findings. But, cephalohematomas are associated with underlying linear skull fractures in 5–25% of cases (1,16). And, as there was no evidence of recent trauma at autopsy, we concluded that the healing injuries described earlier most likely represented a birth-related linear skull fracture with an associated subperiosteal hematoma, that is, a cephalohematoma, even though the birth record does not document clinical findings of birth trauma.

References

1. Reichard R. Birth injury of the cranium and central nervous system. Brain Pathol 2008;18:565–70.
2. Dupuis O, Silveira R, Dupont C, Mottolese C, Kahn P, Dittmar A, et al. Comparison of "instrument-associated" and "spontaneous" obstetric depressed skull fractures in a cohort of 68 neonates. Am J Obstet Gynecol 2005;192:165–70.
3. Hickey K, McKenna P. Skull fracture caused by vacuum extraction. Obstet Gynecol 1996;88:671–3.

4. Gorbani DK. Comminuted skull fracture in a neonate. BMJ 1970;1(5698):756.

5. Garza-Mercado R. Intrauterine depressed skull fractures of the newborn. Neurosurgery 1982;10:694–7.

6. Heise RH, Srivatsa PJ, Karsell PR. Spontaneous intrauterine linear skull fracture: a rare complication of spontaneous vaginal delivery. Obstet Gynecol 1996;87:851–3.

7. Rupp W, Ropohl D, Bohnert M. Zur Defferenzialdiagnose von traumatischen Schadelbefunder bei Saulingen: Residuen nach Vakuumextraktion. Arch Kriminol 2005;215:70–6.

8. Gannon FH, Thompson LDR. Traumatic fracture callus. Ear Nose Throat J 2007;86(4):200.

9. Kemp AM, Dunstan F, Harrison S, Morris S, Mann M, Rolfe K, et al. Patterns of skeletal fractures in child abuse: systematic review. BMJ 2008;337:1518.

10. van Rijn RR, Bilo RAC, Robben SGF. Birth-related mid-posterior rib fractures in neonates: a report of three cases (and a possible fourth case) and a review of the literature. Pediatr Radiol 2009;39(1):30–4.

11. Monjok E. Clavicle fractures during birth. Am Fam Physician 2008;78(6):697.

12. Matsubara S, Izumi A, Nagai T, Kikkawa I, Suzuki M. Femur fracture during abdominal breech delivery. Arch Gynecol Obstet 2008;278(2):195–7.

13. Husain SN, King EC, Young JL, Sarwark JF. Remodeling of birth fractures of the humeral diaphysis. J Pediatr Orthop 2008;28(1):10–3.

14. Stoll BJ, Kliegman RM. Nervous system disorders. In: Behrman RE, Kliegman RM, Jenson HB, editors. Nelson textbook of pediatrics, 17th edn. Philadelphia, PA: W. B. Saunders, 2004;561–9.

15. Mosavat SA, Zamani M. The incidence of birth trauma among live born term neonates at a referral hospital in Rafsanjan, Iran. J Matern Fetal Neonatal Med 2008;21(5):337–9.

16. Nicholson L. Caput succedaneum and cephalohematoma: the Cs that leave bumps on the head. Neonatal Netw 2007;26(5):277–81.

Additional information—reprints not available from author:
William R. Oliver, M.D.
East Carolina University
Department of Pathology and Laboratory Medicine
2100 Statonsburg Road
PO Box 6028
Greenville, NC 27835
E-mail: oliverw@ecu.edu

CASE REPORT

PATHOLOGY AND BIOLOGY

Asser H. Thomsen,[1] *M.D.; Anne Grethe Jurik,*[2] *D.M.Sc.; Lars Uhrenholt,*[1] *Ph.D., D.C.; and Annie Vesterby,*[1] *D.M.Sc.*

Traumatic Death in Ankylosing Spondylitis

ABSTRACT: Ankylosing spondylitis (AS) is a chronic rheumatic disease that causes spinal rigidity with an increased risk of spinal fractures. We present a case report where a middle-aged man, in apparent good health, died following a fall from his bike. Postmortem computed tomography (CT) showed several fractures in the cervical and thoracic spine, with displacement into the spinal canal as well as spinal changes consistent with AS. The cause of death was determined to be upper spinal cord injury caused by cervical spinal fractures that were facilitated by spinal rigidity from AS. Further investigation into the medical records revealed that the decedent had previously been treated for AS. This case report illustrates the importance of obtaining a detailed medical history when investigating deaths, including nonfatal conditions, such as AS. Furthermore, it shows the value of CT in the evaluation of the mechanism and manner of death.

KEYWORDS: forensic science, forensic pathology, ankylosing spondylitis, spinal fractures, spinal cord injuries, radiology

Ankylosing spondylitis (AS) is a chronic rheumatic disease primarily involving the sacroiliac and spinal joints. In advanced disease, the main structural features are fusion of the sacroiliac joints and rigidity of the spine caused by bridging bone between vertebral bodies (syndesmophyte formation) as well as ankylosis of the apophyseal joints. This rigidity contributes to an increased risk of spinal fracture from even low-energy trauma, often in the thoracolumbar and cervical regions (1).

The aim of this case report is to review the possible impact of pre-existing structural skeletal disease on traumatic death and to bring attention to the importance of getting a complete medical history.

Case Story

While alcohol intoxicated a middle-aged man rode his bike home from a bar. A witness saw him swaying, and at low speed, he rode the bike into a curb stone. He went over the handlebars and collided with the pavement face first. His breathing ceased immediately, soon followed by cardiac arrest. Immediate resuscitation was unsuccessful. Because of the rapidity of the cardiac arrest, the emergency physician suggested that death was caused by a cardiovascular event prior to the fall.

According to the antemortem information obtained from the police report and the general practitioner, the deceased was in good health with no prior cardiovascular disease.

Postmortem computed tomography (CT) revealed multiple spinal fractures, including a fracture of the odontoid process of C2 and disco-vertebral separation through the region of C3-C4 and Th10-Th11 (Figs. 1–3). At the C2 fracture displaced fragments in the spinal canal compressed the spinal cord. There was ossification of the anterior longitudinal ligament throughout the spinal column and extensive bridging syndesmophytes, particularly in the lumbar spine, giving the characteristic configuration of a
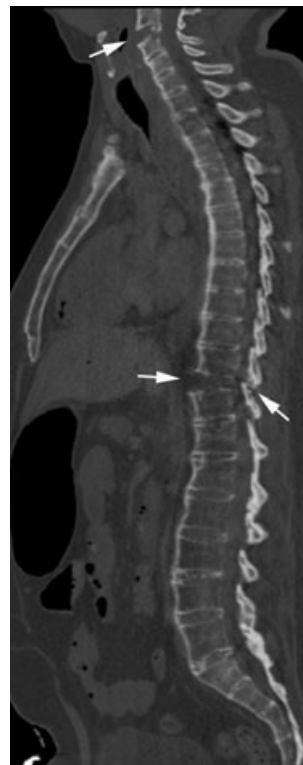
FIG. 1—*Sagittal 2D computed tomography reconstruction of the thoracolumbar and lower cervical spine shows changes typical for advanced ankylosing spondylitis with fusion of vertebral bodies due to bridging syndesmophytes. There is discontinuation at the disk between Th. 10 and Th. 11 with a widening anteriorly and fracture of the posterior column (arrows). A cervical fracture is apparent at the upper part of the image (arrow).*

[1] Department of Forensic Medicine, Aarhus University, Brendstrupgaards Vej 100, Aarhus N DK-8200, Denmark.
[2] Department of Radiology, Aarhus University Hospital, Noerrebrogade 44, 8000 Aarhus C, Denmark.

FIG. 2—*Sagittal 2D computed tomography reconstruction of the cervical spine demonstrating two lesions: (i) fracture at the base of the odontoid process with anterior displacement of the body part (black arrow), and (ii) rupture with widening of the disc space between C3 and C4 (white arrow) accompanied by avulsion of the vertebral plate of C4 posteriorly. The non-traumatized part of the spine shows typical ankylosing spondylitis changes.*

bamboo spine in AS (Fig. 1). In addition, there was osseous fusion of the sacroiliac joints. All CT findings were in agreement with a diagnosis of AS.

The medicolegal autopsy showed abrasions on the face (Fig. 4), on the back, the hands, and the legs. In addition, there was spinal cord injury at the level of C2, hemorrhage in and around the spinal fractures (Fig. 5), rib fractures with sparse bleeding comparable with sequels of resuscitation, osseous bridging of several intervertebral joints, an enlarged heart and insignificant atherosclerosis, a fatty liver, and an enlarged spleen. Blood alcohol concentration was 189 mg/dL. The microscopical examination revealed hypoxic changes in the brain, granuloma formation in the lungs consistent with sarcoidosis, and bone marrow (fatty) emboli in the lung arteries (Fig. 6). The bone marrow emboli were believed to come from the primary spinal fractures or the secondary rib fractures caused by the resuscitation effort.

Additional hospital records were requested and revealed that 9 years previously, the deceased had received treatment in an out-patient rheumatic clinic for AS with rigidity of the thoraco-lumbar spine. Furthermore, he had been under evaluation for pulmonary sarcoidosis.

The manner of death was determined to be accidental, while the cause of death was attributed to cervical spinal cord injury in the region of the fracture. The pre-existing spinal rigidity because of AS predisposed the victim to multiple fractures.

FIG. 3—*Coronal 2D computed tomography reconstruction of the lower thoracic and upper lumbar spine (Th7 – L1 region) showing disruption of otherwise bridging syndesmophytes at the intervertebral space between Th. 10 and Th. 11 laterally (arrows).*
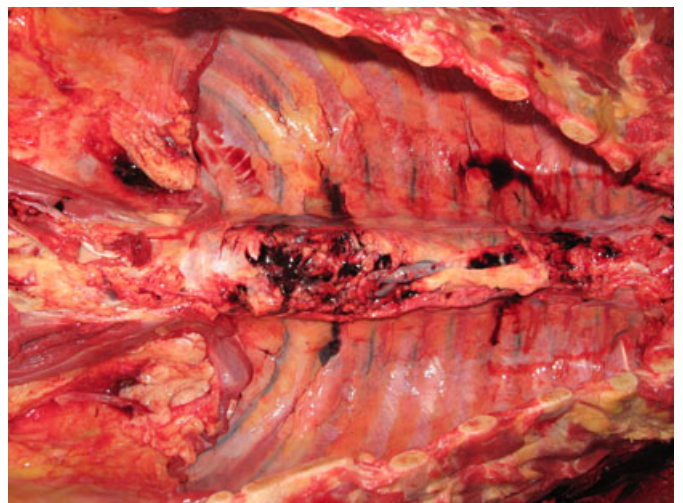

FIG. 4—*Facial abrasions.*

FIG. 5—*Haemorrhage around thoracic spinal fractures (center).*

Discussion

This case report illustrates the fatal outcome of spinal injuries in an individual suffering from spinal stiffness because of AS. The pre-existing structural changes were predisposing factors leading to his death.
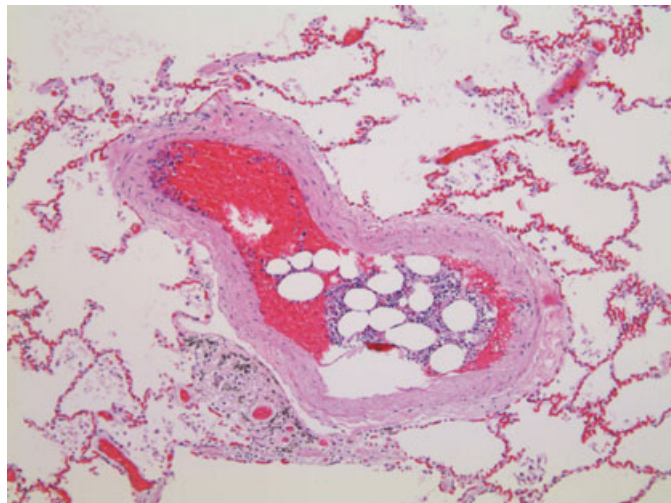

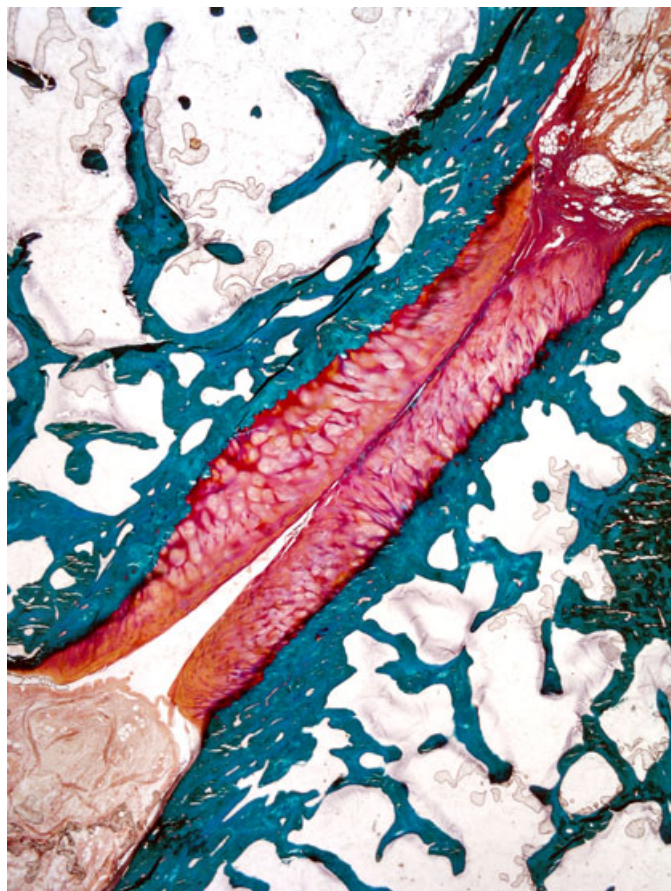FIG. 6—*Bone marrow embolism in small pulmonary artery.*

Ankylosing spondylitis is a chronic systemic inflammatory disease with a variable course. The disease manifests predominantly during the third decade. Sacroiliitis is usually the first clinical sign of AS, and the diagnosis can be made based on this with additional criteria according to the modified New York criteria (2). Although there is known to be a genetic predisposition with a strong association with HLA-B27, the exact cause of the disease remains unclear. An auto-immune response may lead to fibrosis and ossification of spinal joints and ligaments, and in the end-stage to fusion of spinal segments (Figs. 7 and 8).

Patients with AS are at risk for sustaining spinal fractures following trauma. The predisposing factors are stiffness and spinal deformity with thoracic kyphosis, and, frequently osteoporotic bone quality. Spinal nonfatal fractures have been reported to occur in up to 6% of patients with AS and especially in patients with long disease duration (3,4). Furthermore, traumatic fractures have been reported to decrease with age (5). Trivial or even unrecognized trauma may cause fracture (3,6). The frequent comorbid osteoporosis may be a contributing factor (7). Because of the spinal ankylosis, the fractures are often unstable and may give rise to spinal cord lesion (6). Neurological complications, both reversible and irreversible, are frequent, occurring in about half of the patients (4).

Spinal fractures may occur without being diagnosed primarily and can then result in pseudo-arthrosis. Radiologically, this may correspond to the local destructive lesions described by Romanus

FIG. 7—*Normal cervical apophyseal joint with well-defined joint margins covered by hyaline cartilage and with synovial folds at the anterior and posterior margins of the joint space surrounded by fatty tissue. (Stain: Masson Goldner-Trichrome).*
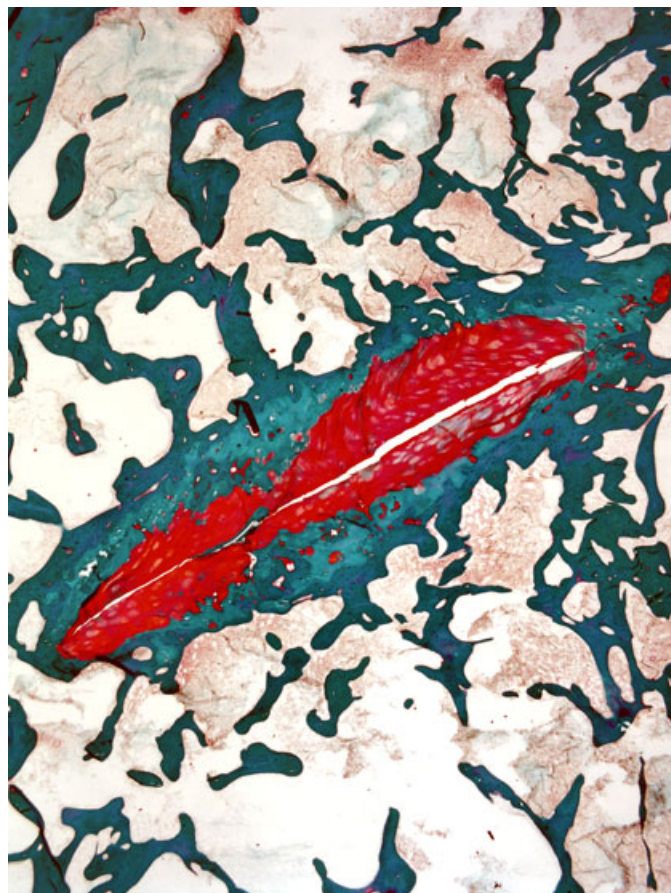

FIG. 8—*A cervical apophyseal joint from the victim that exhibits osseous union of the joint margins corresponding to ossification of the joint capsule and peripheral structures. The hyaline cartilage is thin and partly replaced by osseous tissue, but there is still a visible joint space. There are no detectable synovial folds. (Stain: Masson Goldner-Trichrome).*

and Yden in 1955 (8). Such lesions were not observed in the present patient.

The knowledge of pre-existing skeletal disease is important in the medicolegal evaluation, as diseases otherwise considered nonfatal can contribute to the cause of death. In this case report, conditions leading to death were recognized and supported by the postmortem CT.

This case report illustrates the importance of obtaining a detailed medical history, including nonfatal conditions, when investigating deaths. Furthermore, it shows the value of CT in evaluation of the mechanism and the manner of death.

References

1. Jacobs WB, Fehlings MG. Ankylosing spondylitis and spinal cord injury: origin, incidence, management, and avoidance. Neurosurg Focus 2008;24(1):E12.
2. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27(4):361–8.
3. Feldtkeller E, Vosse D, Geusens P, van der Linden S. Prevalence and annual incidence of vertebral fractures in patients with ankylosing spondylitis. Rheumatol Int 2006;26(3):234–9.
4. Vosse D, Feldtkeller E, Erlendsson J, Geusens P, van der Linden S. Clinical vertebral fractures in patients with ankylosing spondylitis. J Rheumatol 2004;31(10):1981–5.
5. Hitchon PW, From AM, Brenton MD, Glaser JA, Torner JC. Fractures of the thoracolumbar spine complicating ankylosing spondylitis. J Neurosurg 2002;97(2 Suppl):218–22.
6. Erlendsson J, Mortensen NP, de Carvalho A. Fractures of the spine in patients with ankylosing spondylitis. Ugeskr Laeger 1984;146(17): 1292–4.
7. Geusens P, Vosse D, van der Linden S. Osteoporosis and vertebral fractures in ankylosing spondylitis. Curr Opin Rheumatol 2007;19(4): 335–9.
8. Romanus R, Yden S. Pelvo-spondylitis ossificans, 1st edn. Copenhagen: Munksgaard, 1955.

Additional information and reprint requests:
Asser H. Thomsen, M.D.
Department of Forensic Medicine
Aarhus University
Brendstrupgaards Vej 100
Aarhus N DK-8200
Denmark
E-mail: aht@retsmedicin.au.dk

CASE REPORT

PATHOLOGY AND BIOLOGY; CRIMINALISTICS

Sibyl R. Bucheli,[1] Ph.D.; Joan A. Bytheway,[2] Ph.D.; and David A. Gangitano,[2] Ph.D.

Necrophagous Caterpillars Provide Human mtDNA Evidence*

ABSTRACT: Decomposition of large mammalian carcasses is greatly accelerated through the action of insects. Specialized feeders capable of digesting keratin and collagen found in skin, hair, and tendons and ligaments are attracted to corpses in late stages of dry decomposition and include *Tinea pellionella*, the casemaking clothes moth, and *Tineola bisselliella*, the webbing clothes moth (Lepidoptera; Tineidae). Until now, details of the caterpillar behavior as necrophagous insects were vague. Here, we detail the behavior of each species and document the incorporation of human hair into the portable larval shelters constructed by the caterpillars of *T. pellionella*. Hair of the decedent used as building material for caterpillar shelters provided enough starting template to amplify and sequence the HVI and HVII sections of the control region (mtDNA) of the decedent.

KEYWORDS: forensic science, forensic entomology, necrophagous caterpillars, lepidoptera, tineidae, human hair, control region, case report

Decomposition of large, mammalian carcasses is greatly accelerated through the action of insects, particularly those belonging to the order Diptera (flies). As decomposition proceeds, a large portion of the corpse is consumed by these insects resulting in a reduction of biomass and an overall drying out of the corpse (1). Other insects are attracted to corpses in late, dry stages of decomposition. They are specialized feeders capable of digesting keratin and collagen found in skin, tendons and ligaments, and hair and include beetles in the family Dermestidae and moths in the family Tineidae (1,2). The activities of dermestid beetles are well known: larval and adult dermested beetles are urban pests commonly encountered feeding on carpets and drapes. They are also common members of the successional community on mammalian carcasses (1) and are widely used in taxidermy practices to remove dried flesh from bone preparations (3). The activities of tineid moths as urban pests are also well known and similar to that of dermestid beetles; they are frequently encountered feeding on textiles derived from natural fibers (3). Two common species are the casemaking clothes moth (*Tinea pellionella* Linnaeus) and the webbing clothes moth (*Tineola bisselliella* [Hummel]). They are frequent pests of stored textiles, especially those that are naturally derived, such as wool carpets, rugs, upholstery and blankets, piano felts, brush bristles, and pet hair, and both have been documented feeding on feathers and insect exoskeletons, as well (4,5). Their role as necrophagous insects is less well understood, especially in human decomposition.

Case Study

The desiccated remains of an individual were found in an abandoned house in an urban area of Galveston County, Texas, in August

[1]Assistant Professor, Department of Biological Sciences, Sam Houston State University, Huntsville, TX.
[2]Assistant Professor, College of Criminal Justice, Sam Houston State University, Huntsville, TX.

of 2007. The house was in poor condition with many openings allowing easy access for insects. The individual, clothed in winter attire, was mainly skeletonized with small amounts of dried tissue adhered to joint areas and the pelvic girdle. Head hair that had sloughed off during decomposition was found adhered to the clothing. Based on the clothing, condition of the body, entomological evidence, and witness statements, postmortem interval was determined to be approximately nine to 10 months (6). An unusually dense population of actively feeding larvae of the casemaking clothes moth *T. pellionella* (approximately 50 individuals) and living pupae of the webbing clothes moth *T. bisselliella* (two individuals) (Lepidoptera; Tineidae), composing nearly half of the total nonfly diversity, was reported from the hair of the decedent (6). Here, we detail the necrophagous behavior of *T. bisselliella* and *T. pellionella* and document the incorporation of human hair into their larval shelters of *T. pellionella*, which provided enough starting template to amplify and sequence the HVI and HVII sections of the control region (mtDNA) of the decedent.

Methods and Results

Rearing and Identification of Moths

For each species found on the individual, larvae were kept separately in culture feeding on the hair of the decedent and held at room temperature and average humidity to complete their life cycle. This was important to document that features of the shelter architecture are unique and sufficient for identification of each species. As is common in studies of moths, features of the male genitalia were used for positive species identification. Emerged adult males were dissected and prepared according to standard practices (7,8) and viewed under a microscope at 10–100× magnification. Slide-mounted genitalia were compared to published literature for identification purposes (4,9).

One dead adult male moth and two pupae were recovered from the hair of the decedent and were positively identified as *T. bisselliella*, the webbing clothes moth. The webbing clothes

moth typically spins silken tunnels or patches on the surface of fabrics as it feeds. The caterpillar remains inside the tunnel moving in search of food, adding on silk to the tunnel as it goes. When the larva is ready for pupation, it leaves its tunnel and spins a spindle-shaped cocoon, usually incorporating fecal matter and food particles on the outside (5). In this situation, the pupae were found inside their characteristic pupal cocoons, which were attached to clumps of hair but showed no evidence of hair being used as building material.

A dense population of approximately 50 caterpillars, actively feeding, was recovered from the hair of the decedent and was positively identified as *T. pellionella*, the casemaking clothes moth. Several larvae were also recovered from inside shed fly puparia associated with the remains. The casemaking clothes moth normally constructs a portable silken shelter that is carried with the caterpillar as it moves and feeds. The caterpillar typically projects its head from one end of the shelter, while its abdomen remains inside. The shelter is enlarged with silk as the larva grows. Frequently, the casemaking clothes moth will incorporate wool and other fibers from the food source into its larval shelter (4). In this situation, the necrophagous individuals had constructed their portable shelters with silk as normally performed but had also incorporated a large amount of the decedent's hair (Fig. 1A,B). This is the first time this activity has been documented for this species of moth under conditions of necrophagous feeding. In addition to human hair, the silken shelters also contained pieces of shed fly puparia (Fig. 1B). To enlarge the shelter to accommodate their growing bodies, *T. pellionella* caterpillars add on silk to either end of their portable shelter. In this situation, the caterpillar cut the hair and puparia debris into smaller pieces and added it at an angle perpendicular to the length of the shelter, working in concentric rings, embedding the material in silk as it was attached (Fig. 1A). In some instances, the hair was left very long and allowed to project from the shelter (Fig. 1A). One such hair had the decedent's follicle intact.

Extraction and Amplification of mtDNA Control Region

Using the entire larval shelter and the decedent's hair incorporated within as extraction material, we amplified and sequenced the control regions of the mtDNA HVI and HVII sections. Samples were decontaminated and washed following the protocol outlined by Jehaes et al. (10). Extraction of mtDNA, amplification, and sequencing of the HVI and HVII sections were carried out according to Budowle et al. (11). Sequencing reactions were run in a ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, CA) using the Big Dye Terminator Sequencing kit v1.1 (Applied Biosystems). *Tinea pellionella* COI mtDNA was extracted and amplified according to Bucheli and Wenzel (12). Sequences were assembled and proof-read using the Sequencher software (Gene Codes Corp., Ann Arbor, MI). Sequence data of the hair contained within the larval shelter was compared to reference DNA obtained from the nail tissue from the decedent. Sequences were submitted to Genbank (*T. pellionella* accession #FJ424518; decedent accession #FJ561375).

Using a minimum of two independent reactions for each region, the sequence data was analyzed using the revised Cambridge sequence as reference. The following polymorphisms were found in both the hair contained within the case and the nail used as reference tissue: 16174T, 16189C (HVI) and 73G, 146C, 150T, and 263G (HVII). According to these results, the hair contained within the case and the nail used as reference tissue may belong to the same individual. We conclude that the hair incorporated to the larval shelter is suitable for identification purposes.
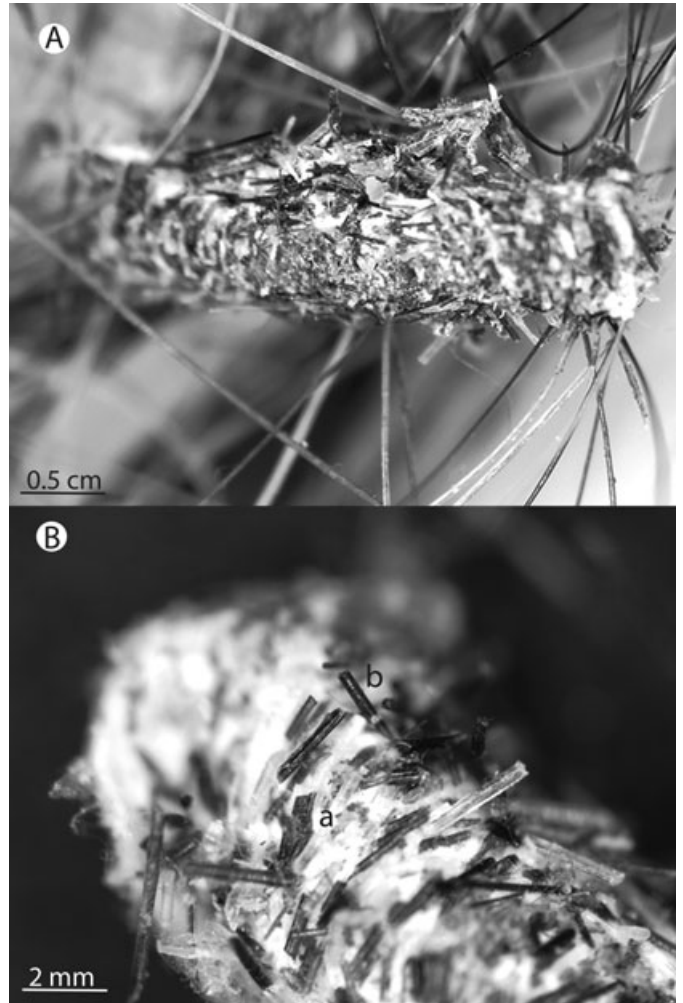

FIG. 1—*A.* Tinea pellionella *caterpillar inside portable shelter feeding on human hair of mummified corpse; B. portable shelter magnified showing shed fly puparia (a) and human hair (b) that have been cut into smaller pieces and embedded in the silk matrix.*

Discussion

Tineola bisselliella and *T. pellionella* are found around the world but are much more common in temperate areas (4,5). Both Robinson (4) and Cox and Pinniger (5) note that both *T. bisselliella* and *T. pellionella* are common in urban areas, but *T. bisselliella* tends to be much more common than *T. pellionella*. Robinson (4) states that house-hold infestations of *T. pellionella* are on the decline, unless houses are damp, poorly heated, and abandoned, while numbers remain common in warehouses and stores. In nonurban situations, they have been found in association with insect remains, bird nests, and owl pellets. While tineid moths have been recorded from mammalian carcasses, details have been deficient. Catts and Haskel (2) document the presence of Tineidae on mammalian carcasses. Both Early and Goff (13) and Tullis and Goff (14) note the presence of tineid caterpillars under human corpses in the Hawaiian Islands. Bourel et al. (15) record *T. bisselliella* from a mummified corpse in France. Oliva (16) documents the presence of larvae of both species on the head hair of corpses found indoors in Argentina. The above-mentioned studies do not document details of the biology of the caterpillars as necrophagous insects nor do they note incorporation of human remains into the larval shelters.

Tineola bisselliella and *T. pellionella* are each small moths with a wingspan of about 1–1.5 cm, possessing narrow front and hind wings. The hind wings are fringed with long hairs. Adults of each are shiny buff to golden in color and possess tufts of scales on their head, which gives the heads the appearance of being "rough." *Tineola bisselliella* is slightly lighter in color and sometimes has head scales that are more red in color than *T. pellionella*. *Tinea pellionella* has a well-defined, slightly darker kidney-shaped spot on the forewing that is lacking in *T. bisselliella* (17). Adults are weak fliers and remain close to the food source, although they do not feed. Both caterpillars and adults of each species are repelled by light and prefer dark, undisturbed places such as basements, closets, and attics (5,18).

Females of both species are reported to lay approximately 50 eggs over a span of 2–3 weeks and then die, while the males continue to mate multiple times. The period to hatching and time of larval development varies based on temperature and humidity. Up to 45 larval developmental stages have been recorded, with a larval period lasting between 35 days and 2.5 years (18). Larvae of both are similar in appearance; each is translucent white with a dark head capsule; however, larvae can be easily recognized by features of their larval shelters (see Methods and Results section previously for those details). It has been noted that each species has a preference for dirty fibers, perhaps suggesting a need for bacterial or fungal associates to aid in keratin digestion (4,5). There are typically two generations per year, although moths may feed and mate year round in areas that are warm and humid (4,5,18).

Pupation lasts about 1 week in summer and up to 4 weeks in winter. Heat and humidity may affect mortality rate and can greatly accelerate completion of larval and pupal stages (18). At the time of pupation, *T. pellionella* leave the food source to find an appropriate location for pupation, sometimes traveling several meters before a suitable area is located. In domestic situations, larvae attach the shelter to a room ceiling for pupation. Pupation occurs inside the portable shelter (4). Larval *T. bisselliella* are for the most part sedentary, completing their lifecycle on the food source where eggs were deposited (5). After emergence, the adult moths of both species mate and females locate an appropriate food source for oviposition (4,5).

This is the first documented instance of *T. pellionella* incorporating human hair into the construction of their portable shelter. It is typical of the tineid larvae to leave the food source at the time of pupation and seek a site removed from the corpse. In an indoor setting, the larvae may choose corners of ceilings to complete the life cycle. While the application may prove limited because of the nature of the circumstances, we feel this wandering behavior, when considered along with their unique biology of incorporating human remains into their shelter, makes them potentially a crucial piece of forensic evidence. The hair incorporated into the shelter may be the only means of potentially obtaining a DNA sample of the individual. Typically, investigators are not trained to collect insects nor to search in areas away from the body for insect evidence. However, in forensic situations where the deceased individual is in later stages of decomposition and the structure in which he/she is found is damp, cold, or abandoned, it may be beneficial to the investigation to examine corners or ceiling areas for possible identification and collection of *T. pellionella* shelters.

References

1. Byrd JH, Castner JL, editors. Forensic entomology: the utility of arthropods in legal investigations. Boca Raton, FL: CRC Press, 2000.
2. Catts EP, Haskell NH. Entomology and death: a procedural guide. South Carolina: Clemson, 1990.
3. Pedigo LP, Rice ME. Entomology and pest management, 5th edn. Upper Saddle River, NJ: Prentice Hall, 2005.
4. Robinson GS. Clothes-moths of the *Tinea pellionella* complex: a revision of the world's species (Lepidoptera: Tineidae). B Brit Mus Entomol Nat Hist 1979;38(3):57–128.
5. Cox PD, Pinniger DB. Biology, behaviour and environmentally sustainable control of *Tineola bisselliella* (Hummel) (Lepidoptera: Tineidae). J Stored Prod Res 2007;43(1):2–32.
6. Bucheli SR, Bytheway JA, Pustilnik SM, Florence J. Insect successional pattern of a corpse in cooler months of subtropical southeastern Texas. J Forensic Sci 2009;54(2):452–5.
7. Clarke JFG. The preparation of slides of the genitalia of Lepidoptera. B Brooklyn Entomol Soc 1941;36:149–61.
8. Robinson GS. The preparation of slides of Lepidoptera genitalia with special reference to the microlepidoptera. Entomol Gazette 1976;27: 127–32.
9. Robinson GS, Nielsen ES. Tineid genera of Australia, monographs on Australian lepidoptera. Vol. 2. Melbourne: CSIRO Publishing, 1993.
10. Jehaes E, Gilissen A, Cassiman J-J, Decorte R. Evaluation of a decontamination protocol for hair shafts before mtDNA sequencing. Forensic Sci Int 1998;94:1–2.
11. Budowle B, editor. DNA typing protocols: molecular biology and forensic analysis. Natick, MA: Eaton Pub, 2000.
12. Bucheli SR, Wenzel JW. Gelechioidea (Insecta: Lepidoptera) systematics: a reexamination using combined morphology and mitochondrial DNA data. Mol Phylogenet Evol 2005;35:380–94.
13. Early M, Goff ML. Arthropod succession patterns in exposed carrion on the island of O'ahu, Hawaiian Islands, USA. J Med Entomol 1986; 23(5):520–31.
14. Tullis K, Goff ML. Arthropod succession in exposed carrion in a tropical rainforest on O'ahu Island, Hawai'i. J Med Entomol 1987;24(3):332– 9.
15. Bourel B, Hubert N, Hedouin V, Gosset D. Entomologie médico-légale appliquée à un cas de momification. Ann Soc Entomol Fr 2000;36 (3): 287–90.
16. Oliva A. Insects of forensic significance in Argentina. Forensic Sci Int 2001;120:145–54.
17. Covell CV Jr. A field guide to the moths of eastern North America. The Peterson Field Guide Series No. 30. Boston, MA: Houghton Mifflin Co, 1984.
18. Rust MK. Pest notes: clothes moths. University of California Division of Agriculture and Natural Resources Publication 7435, 2000.

Additional information and reprint requests:
Sibyl Rae Bucheli, Ph.D.
Sam Houston State University
P.O. Box 2116
Huntsville
TX 77341
E-mail: bucheli@shsu.edu

Recognition∕Appreciation of Guest Reviewers—2009

The success of the *Journal of Forensic Sciences (JFS)* depends on both the quality of the manuscripts submitted and the peer reviews performed by both the Editorial Board Members and the Guest Reviewers. I offer my thanks to all who have assisted *JFS* in the review process. In appreciation of all the efforts of the Guest Reviewers in 2009, I have listed them in this issue of *JFS.*

I would like to recognize Ms. Kim Waggoner of the FBI Laboratory who has coordinated the many reviews performed by that organization throughout 2009.

Colin Aitken	Michael Coble	James Hamby	Laura Liptai
Ivo Alberink	Kim Collins	Randy Hanzlick	Dianne Little
Peter Alexander	Jeffrey Comparin	Sally An Harbison	David Lucy
Marc Allard	Anthony Costantino	Peter Harrington	Niels Lynnerup
John Allison	Heather Coyle	Edward Harris	Lewis Maddox
Gail Anderson	Rhonda Craig	Richard Harruff	Abdul Malik
Jan Andrasko	Christian Crowder	Amy Hart	Craig Mallak
Roman Aranda IV	James Curran	Kristen Hartnett	Mary Manhein
Melanie Archer	Mark Dadum	Lynne Herold	Pierre Margot
Dana Austin	John Daugman	Carolyn Hill	Craig Marhefka
Alon Avidan	Jan De Ceuster	Madeline Hinkes	Mark Marpet
Benjamin Bachrach	Patrick De Smet	Thomas Holland	Kenneth Marr
Robert Barsley	Dororthy Dean	Robert Hoppa	Lance Martini
Eric Bartelink	James Defrancesco	Frank Horvath	Janet Masson
Scott Batterman	Stephen deRoux	Marilyn Huestis	Carl McClary
Andy Becue	Betty LDesPortes	Lee Hulse-Smith	R. Blake McConnell
Lynne Bell	Robert Drew	Benjamin Jones	Kevin McElfresh
William Bernet	Olaf Drummer	Moshe Kam	Sue McGlashan
Herman Bernitz	J. Christopher Dudar	Sree Kanthaswamy	James McNamara
Arthur Berrier	David Duewer	Kathleen Kasper	Marilyn Menotti-Raymond
Stephen Billick	Johan Duflou	Ronald Kelly	Richard Merritt
Martha Blake	George Duncan	Joseph Kenan	Owen Middleton
John Bond	Tom Duncan	Jules Kieser	Erin Miller
Andy Bowen	Brian Eckenrode	Erin Kimmerle	Harry Mincer
Joseph Bozenko	Mahmoud Elsohly	Pascal Kintz	Gordon Miskelly
Charles Brenner	Martin Evison	Lawrence Kobilinsky	El Molto
Scott Bresler	Scott Fairgrieve	Sandra Koch	Stephen Montgomery
Vaughn Bryant	Todd Fenton	Lisa Kohler	David Morello
John Buckleton	Shari Forbes	Jason Kokoszka	Bruno Morgan
Stephen Bunch	Laura Fulginiti	Debra Komar	Robert Muehlberger
JoAnn Buscaglia	Laura Gahn	Lyle Konigsberg	Dawn Mulhern
Mary Bush	Zeno Geradts	Robert Koons	Douglas Murphy
Peter Bush	Eugene Giles	Gerry LaPorte	Niamh NicDaeid
John Butler	Peter Gill	Douglas Lacy	Ronald Nichols
Roger Byard	Gregory Golden	Glenn Langenburg	Janice Nicklas
James Caruso	John Goodpaster	Elizabeth Laposata	Rachel Norris
Eoghan Casey	Eleanor Graham	Sarah Lathrop	Di Nunno Nunzio
Amarjit Chahal	Catherine Grgicak	Wendy Lavezzi	William Oliver
Ranajit Chakraborty	Rod Gullberg	Jane Lewis	William R. Oliver
Christophe Champod	Wendy Gunther	Linda Lewis	Jimmie Oxley
Shawn Christ	Luke Haag	Simon Lewis	Carla Oz
Stephen Cina	Brad Hall	Adrian Linacre	Frank Padula
Karl Citek	Joy Halverson	Sandra Lines	Peter Palmer

Letter to the Editor—Quality Assurance in Disaster Victim Identification (DVI) Exercises

Disaster victim identification (DVI) exercises involve the coordinated identification of human remains following a mass disaster. The number of victims involved can vary from under 10 in a light aircraft crash to thousands, as in the World Trade Center attack of 2001. Bodies may be intact or fragmented, fresh or putrefied, and there are a number of issues that arise in terms of jurisdiction, recording of data, and security.

One of the recurring problems in DVI incidents that involve multiple agencies is the potential for misidentification because of failure to establish clear practice guidelines and lines of authority at the earliest possible opportunity, inadequate or idiosyncratic initial examinations, failure to follow protocols, and failure to ensure adequate quality assurance reviews (1). Each step of the identification process, from locating bodies at the scene to formal acceptance of identifications at Reconciliation Panel Meetings requires an internal audit. Each body or part needs to have a unique identifier, and this should be able to be tracked through the entire process of fingerprint, pathology, dental, and DNA examinations.

In our experience, this is not always the case, and a large amount of time in the later stages of large DVI exercises is spent in re-examining and sampling bodies and parts that may have been examined multiple times before—a very time-consuming and costly process. A way to highlight the effectiveness of a DVI exercise would be to focus on the number of cases/specimens that had to be re-examined to correct failures in procedures. This "Correction of Failures Index" could be measured quite simply as the total number of cases/specimens that had to undergo re-examination as a percentage of the total number of cases/specimens examined.

An exercise where most cases have had to be re-examined would give a CF index close to 100% indicating inefficient practice, whereas an exercise with minimal errors requiring very few re-examinations would be close to 0%. The lower the index, the more effective has been the exercise. The CF index could provide a simple way to assess the overall effectiveness of an exercise and also to monitor activity during specific times of the process, to determine when errors were most likely to occur, or be discovered. Examining teams with high CF indices could be identified early in the process and steps taken to focus on specific issues that would require correction or modification.

The desirable range and application of an effective CF index will only be established by trialling this in the field. As each disaster will be different, it is quite likely that the denominator may also be different, representing complete bodies in one instance or fragmented body parts in another.

Reference

1. Byard RW, Winskog C. Potential problems arising during international disaster victim identification (DVI) exercises. Forensic Sci Med Pathol 2010;6:1–2.

Roger W. Byard,[1] M.D. and Calle Winskog,[2] M.D.
[1]Discipline of Pathology, Level 3 Medical School North Building, The University of Adelaide, Frome Road, Adelaide, SA 5005, Australia.
[2]Forensic Science SA, 21 Divett Place, Adelaide, SA 5000, Australia.
E-mail: byard.roger@saugov.sa.gov.au

BOOK REVIEW

Jens Amendt,[1] *Ph.D.*

Review of: *Forensic Entomology—The Utility of Arthropods in Legal Investigations*

Forensic entomology, the use of insects and other arthropods in legal matters, is still on the move. Its lively state is illustrated by the growing number of articles in all relevant forensic and medical journals in the past decade. It is therefore hardly surprising that there is a new edition of one of the first pioneers of this recent development after almost 10 years. Probably one of the most popular (but maybe also most difficult to meet) claims of the readership is, that a second edition should be better then the first one. Here comes the good message: It is better, definitely. Why? There are simple formal reasons like the much better print quality of the pictures, their improved embedding in the text, and the fact that they are all in color now. Formerly comprising 14 chapters we now find 21, leading to about 250 more pages. So, first of all there is simply more information in a much better design. But there are also reasons of quality, which make this book better than its precursor. Many of the chapters are mini-reviews about the state of art, most of them offering a good introduction in the topics described.

A general section about entomology and insects of forensic interest, introduced by a survey of the historical and recent development of forensic entomology, leading to its current status, is followed by instructions for the appropriate collection of entomological evidence. Here, the reader finds much general information about entomology, definitely a mandatory part for a forensic scientist who needs an introduction into this theme for the first time. However, I doubt the necessity of the excessive illustration as a guideline for instructions for the sampling of insects. Aerial netting of flying insects or labeling a vial does not need to be illustrated by 5–10 pictures each. The detailed and helpful chapter about the laboratory rearing of insects finishes this practical part of the book, which introduces the basic tools of forensic entomology.

The following 17 chapters present the science behind the maggot mystery, dealing e.g., with insect succession on corpses, introducing different habitats like soil and aquatic environments, or discussing the fragile technique of estimating the postmortem interval by the means of necrophagous insects. The new chapters especially enrich this second edition, such as the introduction into forensic acarology and a survey about the interactions of soil and necrophagous insects.

A chapter about ecology theory does not hesitate to name some inconvenient truths, e.g., the unsatisfying situation that there are many studies which describe patterns in insect succession, but do not provide any insights in the mechanisms that shape the structure of insect communities on cadavers over time and space. Unfortunately, this lack of knowledge hampered an appropriate use of succession data in forensic entomology, and this chapter offers some inspiring ideas for future research to correct this situation.

The new edition also presents an excellent key for the blow flies of America, North of Mexico, and a helpful introduction into forensic meteorology. Some of the topics, which were covered already in the first edition, are now improved. Here, I would highlight as an example the very well-written chapter about molecular methods in forensic entomology. However, there are also chapters that are more or less a copy of the contributions of the first edition. For example, the reader will find statements like that one author "*recently*" published some interesting results, but in fact "*recently*" means "*1999*," i.e., 2 years before the first edition was released. This might be just bothersome, but would be really annoying if important developments in the recent past are not considered and discussed. There are several recent articles and studies, trying to improve our ability in estimating the age of immature insects, but they do not appear in any of the chapters dealing with insect development and calculation of intervals. One of those chapters does not even give any reference younger then 2001, and this is disappointing. So while it was a pleasure to read the majority of the contributions, a few chapters may need refreshment in a possible third edition.

In summary, the book offers information and background for many different groups of readers like forensic pathologists and forensic scientists in general, for the crime scene investigator as well as for the entomologist, for the beginner as well as for the advanced researcher. It is an excellent start to familiarize oneself with forensic entomology, but also offering at the same time original information which you will not find elsewhere.

[1]Institute of Forensic Medicine, 60596 Frankfurt, Germany.

BOOK REVIEW

Joseph P. Bono,[1] *M.A.*

Review of: *Challenges to Fingerprints*

REFERENCE: Haber L, Haber NH. Challenges to fingerprints. Tucson, AZ: Lawyers and Judges Publishing Company, 2009, 217 pp.

The evaluation of friction ridge pattern individualization (FRPI) has been a much discussed topic during the past year. The National Research Council of the National Academies report "Strengthening Forensic Science in the United States: A Path Forward" has raised the level of awareness regarding what needs to be "strengthened" in this forensic discipline. This book addresses some of the same issues where FRPI must be strengthened. The text should have been more properly entitled "Challenges to Friction Ridge Pattern Individualization." Fingerprint analyses are but a subset of FRPI.

Before beginning a review, there are some basic corrections that must be addressed:

- The American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) is the correct accrediting body for forensic science laboratories. The authors refer to ASCLD, which is not an accrediting body, and is in fact a separate organization.
- The Federal Bureau of Investigation (FBI) is NOT housed within the NIJ (National Institute of Justice). The FBI is a part of the Department of Justice.
- The NIJ does not "overview" FRPI. It does, however, provide funding for laboratories and research.

Basic misstatements like these impact the credibility of the authors of any authoritative text. One requirement for an authoritative text is that the authors be "authorities." The academic degrees of the authors (both have a Ph.D. degree) appear on the cover; however, there are no statements specifying their field(s) of expertise. Aside from taking some basic courses, real world laboratory training and experience required of most FRPI examiners seems absent from their credentials. The authors also state "that [FRPI examiners] rarely, if ever make an erroneous identification in court." They then proceed in the next 192 pages to delineate errors in every phase of a FRPI. Page after page is replete with what is done improperly in the laboratory. This reviewer could not find any examples of what constitutes an inherently validated protocol currently in use. The only review of a profession, which is more questionable than one in which the author(s) claims that everything is "right," is one which the author(s) claims that everything is "wrong." The "everything is wrong" approach is quite evident in this text.

Chapter 3 contains some basic descriptions of images FRPI examiners evaluate in the laboratory. The text does not reference the source(s) of this information; however, the narrative and photos could have come from any introductory text describing FRPI and they do appear to be accurate.

The authors' reference to the term "error rate" (ER) appears in almost every chapter. The authors dictate to the reader that the absence of this factor from a report or testimony is grounds for challenging a FRPI analysis. There are plenty of reasons for challenging any kind of forensic analysis. This certainly applies to FRPI, and many of those challenges are justified. What is not justified is repeatedly challenging someone with an undefined term. The authors inform the reader that a "computation of an error rate for fingerprint conclusions based on court data cannot be performed." Without defining the numerator or denominator required for any "rate" computation, they emphasize that a calculation or ER for FRPI casework is not available. This reviewer believes that such a calculation based on proficiency tests cannot be used to extrapolate this "rate." The question then becomes: What is this nebulous term which has taken on a life of its own in the literature by those whose arguments are based on a philosophical perspective rather than on empirical data and images?

The authors are correct to point out that examiners should not claim that "zero errors" are the norm in FRPIs. It would be more valuable for the discipline if the authors had included a statement that errors do occur and provide suggestions for eliminating these errors. Then again, the authors are not examiners and they might not have suggestions for how to improve this forensic science discipline.

In Chapter 6, the authors discuss the Automated Fingerprint Identification System (AFIS) and the issues inherent to this system. The authors rightfully "ensure" (they mistakenly and repeatedly use the word "insure") the reader that AFIS does not make "identifications." Then, they follow with: "AFIS is already being used as an identification system for tenprints." This simply is not true. These kinds of contradictory statements appear throughout the text. An AFIS search will provide a list of friction ridge patterns with a prioritized listing of algorithm "matches." This search does not identify anyone. That is the examiner's job.

This text is devoid of substance for anyone seeking solutions for developing thresholds for the individualization of a friction ridge pattern to a particular person or for ways to "strengthen" FRPI. Most interested parties already understand the challenges; the real value in any text that evaluates this discipline will lie in providing solutions. This reviewer cannot recommend the text because it is a restatement of what most already know about what must be strengthened, and it contributes very little to finding the solutions to the "challenges."

[1]Adjunct Instructor, Forensic and Investigative Sciences Program, Indiana University, Purdue University, Indianapolis, IN.

BOOK REVIEW

Mark D. Semon,[1] *Ph.D.*

Review of: *Mathematical Methods for Accident Reconstruction*

The extensive notes I took while reading this book are best organized by separating them into three categories: The Good, The Bad, and The Ugly.

The Good

Chapters 9–13 contain a wealth of information distilled from many sources. Chapter 9 presents the construction standards for various classes of roadways and also has excellent discussions of the concepts of "design speed" and "sight distance." Chapter 10 discusses train accidents, and Chapter 11 discusses accidents involving larger vehicles, such as tractor-trailers, and accidents involving off-road equipment, such as cranes and forklifts. Chapter 12 presents a nice discussion of "perception-reaction" times and their role in formulating the standards of road design. Chapter 13 presents the most common mathematical methods used to solve systems of linear equations, analyze data sets, and perform numerical differentiation and integration. Some well-known software packages for accident reconstruction are also discussed in this chapter as is some of the equipment used by reconstructionists, such as accelerometers and "black boxes." Chapters 9–13 are well written, clearly presented, and will save forensic examiners the time that would have been required to find this information on their own.

The Bad

At the beginning of the book, on page xiii, the authors define the symbols to be used and note that "In all cases, bold lettering represents vector quantities, and nonbold lettering represents scalar quantities." Unfortunately, the use of bold and nonbold symbols is wildly erratic. For example, on page 11, velocity (a vector) is set equal to the time rate of change of the x-position (a scalar), on page 145 the variable t for time is in boldface, two equations on page 17 have scalars equal to vectors, and on page 91 the "total vector coefficient of kinetic friction" is represented by a nonbold symbol. Also, equations on both pages 17 and 32 show two vectors "multiplying" each other with no indication of whether the multiplication is a scalar or vector product. The lack of attention to the use of bold and nonbold symbols makes the equations and derivations at best hard to follow and at worst incorrect.

There are many mistakes that could have been corrected by a knowledgeable proof reader. For example, equation (2.14) is missing a t, equation 2.45 has w_0 instead of $(w_0)^2$, the quantity d_s in equations (2.77) and (2.78) should be ds, equation (2.85) is missing two minus signs, and equation (9.4) should have a t_i^2 instead of t_i. In Table 4.2, the coefficient of friction for tires on new concrete is quoted as being between 0.08 and 1.2 for speeds less than 30 mph—the 0.08 should be 0.8. In equation (14.15), the numerator is written as ∇v instead of Δv and, when cylindrical coordinates are introduced on page 8, the angle in the *x–y* plane is symbolized by θ in the text but by φ in Figure 2.2.

The heading for Section 19 of Chapter 2 is "Dissipation and Conservation of Forces." However, it is *energy* that is dissipated or conserved, not *forces*. The section itself discusses *dissipative* and *conservative* forces, which is certainly a correct way to describe forces. Perhaps the chapter heading was meant to be "Dissipation and Conservation of Energy" or "Dissipative and Conservative Forces." Unfortunately, as it stands, the chapter heading is incorrect.

Most of the references in the Bibliography are relatively old. For example, all of the references on crush damage were published before 1993. I was surprised that the Bibliography did not include the classic series on accident reconstruction written by Baker and Fricke (from the Northwestern University Traffic Institute). There is a reference on page 136 to R. Limpert but, because there are two publications by Limpert in the Bibliography, we do not know the one to which the reference refers. Also, the Bibliography lists the text by Sears and Zemansky twice and fails to note that neither Sears nor Zemansky has been listed as an author since 1999.

It is not unusual to find one symbol being used to represent different quantities. For example, at the beginning of Chapter 4, the letter s represents a distance, then halfway through the chapter s represents a speed, and then five equations later s represents a distance again. Sometimes symbols are used before they are defined or are never defined. For example, on page 77 the symbol Δs appears in an equation without being defined and then is never used again.

There are very few worked, numerical examples in the book, and the first one does not appear until page 119.

The Ugly

Some of the explanations of the fundamental laws of physics are incorrect. On page 21, the text says that Newton's "second law is a direct consequence of Newton's first law." This is incorrect—just because a body free from all external influences moves with a constant velocity does not mean that its change in velocity is proportional to the net force acting on it or is inversely proportional to its mass. (For example, in principle, there is no reason why the change

[1]Department of Physics and Astronomy, Bates College, Lewiston, ME 04240.

in velocity could not be proportional to the mass squared.) In fact, if Newton's second law was a direct consequence of the first law, then it would not be referred to as a law, it would be a corollary to Newton's first law.

An example used on page 22 to illustrate Newton's third law is incorrect. The text says that for a body sitting on a table, the Newton's third law pair is the force of the Earth on the body (acting down) and the force of the table on the body (acting up). The actual third law pair is the force of the Earth on the body and the force of the body on the Earth. On page 23, the text says that "Newton's third law describes the equilibrium of a particle or an object." This statement is incorrect because each force in a Newton's third law pair acts on a different body, whereas the equilibrium of a body is determined by all the forces acting on it alone. Finally, starting on page 78, the text includes mv^2/r as one of the forces acting on a body rather than correctly treating it as a particular form of ma (i.e., as a term appearing on the right-hand side of $\mathbf{F_{net}} = m\mathbf{a}$).

Some derivations contain errors, and some examples are worked incorrectly. For example, in the derivation on page 206, the cosine in equation 11.3 should be a sine, and there is a minus sign error in equation (4.11). In the example on page 260, a falling human is modeled as a cylinder tipping over. As the cylinder is only tipping, it is just rotating about a fixed point, but the text says it is rotating and translating. Consequently, the final answer derived in the text for the speed at which the top of the cylinder hits the ground is incorrect. There also is an implicit assumption in this model that should have been stated, namely, that the neck is held rigid throughout the whole process.

Conclusion

Chapters 9–13 provide excellent reference material that is very well presented. Other than that, because of all the notational and conceptual errors it contains, I am sorry to say that I cannot recommend the book.

BOOK REVIEW

John A. Williams,[1] *Ph.D.*

Review of: *Recovery, Analysis, and Identification of Commingled Human Remains*

REFERENCE: Adams BJ, Byrd JE. Recovery, analysis, and identification of commingled human remains. Totowa: Humana Press, 2008, 374 pp.

Mass disasters involving the commingling of human remains such as at the World Trade Center (WTC) and incidents of genocide are on the increase. Forensic anthropologists and other death investigators face the real challenges brought on by these events. Bioarchaeologists have been aware of the enormity of the task of reuniting commingled human remains for some time. Only now, this task has gone from an academic exercise involving nameless prehistoric individuals to one in which lives may be altered based on the choices made to assess commingled remains.

The two authors/editors of this text themselves have some degree of experience with the effects of commingling, Adams in his role as a forensic anthropologist with the Office of the Chief Medical Examiner (OCME) of New York and Byrd with the Joint POW/MIA Accounting Command (JPAC).

There are 18 chapters and 36 contributors in this volume. Contributors run the gamut from those assisting medical examiners, to biarchaeologists, to those in the human rights arena. The chapters fall into one of two categories, narratives and case studies including recovery scenarios, and methodological applications. Both chapter formats are scattered through the text making for a somewhat disjointed chapter sequence. Content varies from nonforensic applications like the Roman age tombs described by Ubelaker and Rife to the recent situations at the TriState Crematorium and the WTC.

Several chapters stood out. Mundorf working for the New York OCME described anthropologist-oriented triage experiences with the WTC, AA Flight 597, and the State Island Ferry crash. The distinction made between Type 1 (recovery induced) and Type 2 (disaster induced) commingling was a strong point in how triage progresses in different mass fatality/commingling scenarios. Byrd describes a very simple, effective method of sorting using bone metrics and regression formulas. The technique is relatively easy to employ and requires a very minimum of equipment. I was disappointed, however, that a spatial analysis program described by Tuller and colleagues used to document mass graves in Serbia was not made available for use by other researchers/practitioners. The final two chapters are devoted to larger issues of mass fatality management. An easily overlooked aspect of commingled remains is data control. As practitioners at the processing level anthropologists, odontologists, and pathologists can get caught up in the activities of the morgue and lose sight of their mission to identify. To accomplish both these efficiently and reliably, data input must be rigorously attended to. Hennessey in the final chapter describes the workflow of data, how and where problems occur, and how to head them off.

The target audience of this text is not clear. For the anthropologist who has worked with commingled remains many of the chapters, while of interest, do not provide specific applications for practical use or are too shallow to allow the reader to draw upon. Still other chapters simply cover information that is common knowledge. Viner's discussion of the use of radiology in mass fatality events is a case in point. For the anthropologist who has not had exposure to commingled human remains or for a death investigator outside of the field, this book does serve as a good overview of the handling of commingled remains, the variety of circumstances that may be encountered, and some of the methods that can be utilized.

[1]Department of Anthropology and Sociology, Western Carolina University, Cullowhee, NC.