Robert D. Koons,[1] *Ph.D. and JoAnn Buscaglia,*[1] *Ph.D.*

Forensic Significance of Bullet Lead Compositions*

ABSTRACT: The concentrations of seven elements in projectile lead specimens received as evidence were used to assess the frequency of the occurrence of two unrelated samples having indistinguishable compositions. A set of data from 1837 samples was selected for this study from a sampling of 23,054 lead bullets and shot pellets received as evidence in the FBI Laboratory over the period 1989 through 2002. The method used for selection of samples from case submissions ensured that no two samples of the same general type from the same case were included and that no bias was introduced concerning representation of manufacturers or production sources. A total of 1,686,366 pairwise lead sample comparisons were made using the concentrations of the elements Sb, Cu, As, Ag, Bi, Sn, and Cd using a match criterion of two times the sum of the standard deviations of the paired samples. Of the 1837 samples, 1397 samples, or 76%, are distinguishable from every other sample in this study. The total number of indistinguishable sample pairs is 674, for a frequency of 1 out of every 2502 comparisons. The frequency of occurrence of matching samples decreases as the number of measured elements is increased and as the precision of the measurements improves. For bullets in which all seven elements were determined, the match frequency is 1 in 7284. Compositional comparison of bullet lead provides a reliable, highly significant point of evidentiary comparison of potential sources of crime-related bullets.

KEYWORDS: forensic science, criminalistics, trace evidence, bullet, lead analysis, elemental analysis, ICP-AES

Evaluation of the evidential significance of bullets is typically done by physical intercomparison of bullets and comparison of surface markings on fired bullets with markings produced by a gun barrel. However, in cases where a firearm is not recovered or cannot be associated with a subject, or a fired bullet is too mutilated for physical comparison, two or more bullets may be compared using the concentrations of selected elements in the lead component (1–2). The basis for the forensic compositional comparison of bullets is the same as for other forms of class evidence; namely, the ranges of compositions of bullets from a single classification or production unit (variously called batch, melt, pour, run, etc.) are much smaller than is the diversity of compositions across multiple sources. Analytical methods that have been used to determine element concentrations in bullet lead include neutron activation analysis (NAA) (1,3–9), atomic absorption spectrophotometry (9–11), spark source mass spectrometry (12–13), inductively coupled plasma-atomic emission spectrometry (ICP-AES) (8,14), and inductively coupled plasma-mass spectrometry (ICP-MS) (15–17). Spark source-optical emission spectrometry (SS-OES) is also widely used for quality control purposes within the lead manufacturing industry, but it is not generally suitable for forensic comparison of bullet leads because of its relatively poor precision. Currently, the procedures used most frequently for this examination employ acid dissolution of the lead and analysis of the resulting solution by ICP-AES (14) or ICP-MS (15). The FBI Laboratory

uses an ICP-AES protocol (2,8,14,18) to determine the concentrations of seven elements, arsenic (As), antimony (Sb), bismuth (Bi), cadmium (Cd), copper (Cu), silver (Ag), and tin (Sn) in lead specimens. Although compositional analysis for comparative bullet lead examination has been widely utilized by forensic chemists and accepted by courts throughout the world for over 35 years, it has not been without challenges. The analytical procedures that are used for lead composition determinations have generally been accepted as reliable when appropriate quality control procedures are followed. However, as with other forms of trace evidence, the major points of debate, both scientific and legal, are in the assessment of the significance of the results of lead compositional comparisons (19).

Most published studies concerning compositional analysis of bullets have been concerned with development and reliability assessment of analytical procedures. However, these studies also provide some indication of the significance of results. It has been demonstrated in many studies and is widely accepted that considerable variability exists among lead sources despite manufacturers' efforts to control the concentrations of selected elements (2–4,8). For example, an ammunition manufacturer will typically specify a given level of antimony in order to control the hardness of a lead alloy that is to be used for a particular product line. A smelter is tasked with supplying the ammunition manufacturer with a lead product that falls within the specified antimony tolerances. However, the acceptable antimony range is much greater than both the variations that occur within an individual product from the smelter and the analytical variations in ICP-AES measurements. Therefore, differences in the concentrations of alloying elements (which include tin and arsenic, as well as antimony for some products) between two projectiles made from the same nominal lead alloy are readily measurable by the precise analytical techniques used by forensic laboratories.

The concentrations of trace elements also display measurable differences between manufacturing production runs of a given lead product (18). These differences are the result of variations in the

levels of trace constituents in the raw materials and the mixing of recycled lead scrap from multiple sources during processing of the lead by smelters and bullet manufacturers. The variations in trace and alloying elements provide a basis for discrimination between melts of lead, even those having the same alloy specification. The forensic utility of compositional comparison of bullet leads lies in the fact that the lack of strict control by the manufacturers results in a large number of distinguishable combinations of these element concentrations.

The variations in the concentrations of six elements within individual bullets, among bullets in a box of cartridges, among boxes packaged on the same and on different dates, and among boxes from four U.S. ammunition manufacturers have been determined using precise analytical methods of NAA and ICP-AES (8). In most boxes of cartridges, the lead components of the bullets form more than one distinguishable compositional pattern and a given compositional profile is repeated in boxes of cartridges produced by a single manufacturer over some period of time. Element concentrations determined in samples from two secondary smelters and an ammunition manufacturer indicate that the multiple compositions within a box of cartridges result from the intermingling of bullets from several melts during the bullet processing, jacketing, cartridge loading and packaging operations (18). Bullets having a given compositional profile may be packaged in multiple boxes because the number of bullets produced from a particular melt of lead greatly exceeds the number of cartridges in a box. Boxes containing the same compositional profile are generally produced within a short period of time, such as one packaging run. However, in some instances, bullets have been stored for longer than one year prior to being loaded into cartridges and placed into boxes (8).

The published studies to date describe the distribution of bullet lead compositions within and among boxes and production runs of cartridges, but they have not definitively answered the question posed to the forensic scientist, namely what significance can be placed on a finding of two lead projectiles having indistinguishable compositions. The information needed to assess this significance is the repeat rates of compositional patterns among bullets, i.e., the frequency of occurrence of any given pattern in a population of bullets. This information has been difficult to obtain because it requires the compilation of consistent analytical data for a large, statistically representative sample population. Two approaches have typically been presented for determining frequency of occurrence statistics. The first involves constructing a probability model based on the manufacturing process. For example, in a Bayesian approach to assessing significance of evidence, the probability that crime-scene and subject-associated bullets are indistinguishable when there is no true association of sources (the denominator of a likelihood ratio) can *theoretically* be calculated. However, to make this calculation, accurate information must be known concerning factors including, but not limited to, intra- and inter-batch variabilities, production volumes, product distribution and use, and geographic and temporal changes in the distribution of bullet compositional distributions. This information is impossible to obtain for any but the simplest of case scenarios. As a result, approaches based on likelihood statistics have been rejected by most U.S. courts and have been abandoned by most forensic scientists for comparisons of all forms of non-biological trace evidence. In this study, we have taken a second approach, that is, to calculate the frequency of indistinguishable compositional patterns in a set of bullet compositional data that was selected from casework submissions so as to represent separate production sources. While this approach will not provide a precise frequency of occurrence statistic for repeat patterns in the general bullet population, it does assist in the

evaluation of the evidentiary value of associations based on compositional bullet lead comparisons.

Experimental Details

Source of Compositional Data

The lead specimens selected for this study were either received as evidence by the FBI Laboratory or were exemplars from known sources used for purposes of evidence evaluation during the period 1989–2002. During the period 1989 to March 1994, element concentrations in bullet lead were determined using both ICP-AES and NAA procedures. After 1994, all measurements were made using ICP-AES only. Detailed descriptions of the sample digestion and ICP-AES protocols used for data collection have been reported previously (8,14). Several specific details, including dilution factors, analyte lists, and analytical wavelengths have changed over the 13-year period of data collection. Whenever any changes to the analytical procedures were made, the consistency of results was confirmed in order to meet the validation requirements in effect at that time. A brief, general description of the analytical procedure follows. Each lead specimen is examined under a stereomicroscope and any visible contamination or original surface material is removed. Three analytical samples, each weighing 50 to 75 mg, are cut from the interior portion of the lead (avoiding destruction of rifling marks) using a separate, clean scalpel. In some instances, the evidentiary specimen is too small to allow the collection of three fragments and a smaller number of analytical samples are used. Each fragment is flattened between two sheets of polyethylene and the resulting disk is weighed to the nearest 0.01 mg and placed in a 15-mL polypropylene tube. The samples are dissolved by heating in 3.2 M (1:5 v/v) reagent grade nitric acid with 10–20 μL of reagent grade hydrofluoric acid added to stabilize soluble tin and antimony species. Sets of digestion blanks and standard reference materials SRM2416, Bullet Lead; SRM2415, Battery Lead; and SRM2417, Lead Base Alloy (National Institute of Standards and Technology (NIST), Gaithersburg, MD) are prepared along with each batch of samples to serve as quality checks and to assure long-term consistency of the results. Multielement calibration of the ICP-AES instrument is accomplished using two standard solutions made by digestion of SRM 2416 and a 99.9999% pure lead standard with addition of an appropriate aliquot of a certified 1000 μg/mL cadmium standard solution. Triplicate samples consisting of the appropriate mixture of HNO_3 and HF are also prepared and analyzed as digestion blanks with each set of samples. All samples, blanks, and standards are loaded into an autosampler for automatic introduction into an ICP-AES instrument for analysis. An analytical run consists of a set of calibration standards, followed by one replicate of each of the three standard reference materials, the digestion blank, and the individual samples. This process of calibration followed by SRM, blank, and sample analysis is repeated three times to provide the triplicate measurements for each sample in a set. For large sample sets, recalibration is performed after approximately every 75 analytical samples. For sample sets containing fewer than about 20 samples, a single calibration run is made for the three replicates. The number of elements determined increased from four in 1989 when NAA data were also being obtained to seven (arsenic, antimony, bismuth, cadmium, copper, silver, and tin) in all samples analyzed after 1996. Once it is verified that the results for all SRM samples agree within the FBI Laboratory's tolerance limits with the values listed on the NIST Certificates of Analysis, the results for all analytical samples and SRMs are added to the FBI Laboratory's historical data collection. At the initiation of this study, element concentrations measured in replicate samples from each of 23,054

TABLE 1—*Element concentrations in SRM 2416, Bullet Lead Standard Reference Material (NIST). Normalization factors were utilized to adjust the mean results for SRM 2416 for each analytical set to match the values shown. All element concentrations are certified by NIST, except Cd, which is only recommended. The concentration shown for Cd was determined by repeated analysis in the FBI Laboratory and is within the recommended range by NIST.*

Element	Concentration (%)
Ag	0.0044
As	0.056
Bi	0.1
Cd	0.00007
Cu	0.065
Sb	0.79
Sn	0.09

TABLE 2—*Minimum standard deviation values for the test set. Any measured standard deviation that was less than a listed value was increased to that value.*

Element	Minimum Standard Deviation (%)
Ag	0.00002
As	0.0002
Bi	0.0001
Cd	0.00001
Cu	0.00005
Sb	0.0002
Sn	0.0002

lead projectiles and replicate samples from each of 2,994 standard reference material runs were contained in the data collection.

Formation of the Test Subset for a Statistical Study

For this study, a test set was formed consisting of the mean and standard deviation values for each measured element in representative lead specimens from the FBI Laboratory's historical data collection. The aim in forming the test subset was to select the maximum number of samples originating from different production sources, while representing each production source only once per case submission. A five-step process was used to reduce the number of samples in the historical data collection to form the test dataset. First, the element concentration measurements were normalized to assure consistency and comparability of results among different analytical runs. Each set of lead specimens was normalized by multiplying the results for each element by the factor that transforms the mean concentration of that element determined for the SRM2416 samples in that set to the value shown in Table 1. The initial data set only included sample sets where the SRM results met the FBI Laboratory's quality control criteria. As a result, the normalization factors were typically between 0.9 and 1.1. Second, those evidentiary specimens for which less than three fragments had been analyzed were excluded from the data set. Third, to eliminate within-case samples that could have come from the same production source, data from only one bullet was retained in the test set in those instances in which multiple bullets could have originated from the same source. To do this, evidentiary submissions were considered one case at a time, and no cross-case comparisons were made during the sample selection process. For each case, one specimen from each combination of bullet caliber, style, nominal alloy class (e.g., 1.5% Sb), and manufacturer, where known, was selected and that data was placed into the test data set. In cases where two or more bullets had the same characteristics, one specimen was randomly selected for the test set (using a random number table) from those containing the maximum number of elements measured. The test set includes case-related and research specimens of both fired and unfired bullets and shot pellets from which three replicate samples were analyzed. Fourth, for those samples for which the three replicate concentration measurements for an element were so close to the same value that the calculated standard deviation was better than could be expected from the ICP-AES procedure, the measured standard deviation was increased to the minimum method precision. These minimum standard deviation values for each element, shown in Table 2, were selected as representative of the best precision values typically obtained over the course of the data collection period. The use of these minimum values is needed

to make a proper statistical comparison between specimens and to be conservative in assigning a method discrimination capability. Fifth, NAA values were used for some samples. Prior to 1995, arsenic was not determined using the ICP-AES procedure, because the FBI Laboratory still used the NAA method. For 320 samples, analyzed prior to 1995, NAA results for arsenic were included in the test set. During this period, copper was not determined for all samples by ICP-AES, again because the NAA data were available. For 18 samples, copper concentrations determined by NAA were included in the test set. For both arsenic and copper, the accuracy and precision of NAA and ICP-AES data have been shown to be comparable (8,14).

This method of sample selection ensures that the test set does not include data from two specimens from the same production source in a given case. The test set in this study, therefore, should represent an unbiased sampling of the FBI Laboratory's bullet lead historical data collection in the sense that only one randomly-selected specimen was included to represent each production source of lead in a case. It should be noted that, because no cross-case comparisons were made in the selection of samples, there might be some samples from the same production source in the data set if they occurred in separate case submissions. However, for the purposes of this study, the presence of multiple specimens with the same composition will be considered to be a coincidental occurrence. The final test set used for this study consists of the mean and standard deviations for all elements that were determined in 1837 specimens from 1009 cases and research studies. This method of sample selection yields one sample from each of what are assumed to be 1837 separate production sources of bullets. It does not, nor is it meant to, represent any production volumes, or geographic or temporal distributions of bullets, such as would be required for calculations of frequency of occurrence in a general bullet population.

Results and Discussion

Descriptive Statistics

The distributions of the 1837 samples in the test set among those manufacturer, caliber, and type categories defined by the FBI are shown in Table 3. Specimens having their manufacturer or caliber indicated as "Not Specified or Unknown" and those having their type indicated as "Fragment of Bullet or Shot" are fired bullets or shot pellets whose physical characteristics could not be determined unequivocally. The manner of sample selection assures that the test set is a fair representation of the types and compositions of ammunition examined by the FBI Laboratory over the 1989–2002 time period, with the exception that small caliber shot whose individual pellets are not large enough to obtain triplicate analytical samples is not included. The test set was formed in order to test the frequency of coincidental repeat occurrences of

TABLE 3—*Numbers of specimens in the test set in three designated descriptive categories.*

(a) Distribution by manufacturing source

3-D (remanufactured)	1	Geco (Germany)	6	PMC (Korea)	61
Activ	137	Gevelot (Canada)	1	Reload	56
Armco (China)	2	Handcast	19	Remanufactured	1
Black Hills	1	Hornady	3	Remington – Peters	353
Brazil	4	Indonesia	1	Russia	2
Britain	1	Israel	7	RWS/Geco/Rottweil (Germany)	2
CAC	2	Italy	10	S&W	16
CCI (Omark-CCI), (Speer)	3	Lapua (Finland)	2	Sako	1
China	260	MagTech (Brazil)	3	Sears	2
Cor-Bon	1	Master Cartridge	3	Sellier & Bellot	4
Czechoslovakia	10	Midway (remanufactured)	1	Sierra	1
Denmark	6	Norinco (China)	1	SNG	1
Dominion (Canada)	1	Norma (Sweden)	1	Taipan	1
Eldorado (PMC)	19	Nosler for Federal	2	U.S. Military	6
Federal	3	Not Specified or Unknown	357	U.S. Military – Reload	1
Finland	1	Other American	21	Winchester – Western	420
Fiocchi (Italy)	7	Other Foreign Made	34	Yugoslavia	5

(b) Distribution by caliber

0.25	1	.300 Win-Mag	2	.44 REM MAG	11
0.32	10	.30-30 Winchester	19	.44 S&W Special	4
0.38	5	.308 Winchester	1	.45 Auto	41
0.44	1	.32 Auto	19	10 mm Auto	1
0.45	15	.32 S&W	5	7.62 × 39 mm	10
.22 (rimfire)	59	.32 S&W Long	29	9 mm	5
.22 Hornet	4	.357 MAG	99	9 mm Luger	74
.22 Long	28	.38, .357 & 9 mm	107	9 mm Luger +P	2
.22 Long Rifle	287	.38 S&W	7	9 mm Luger +P+	1
.22 Short	20	.38 Special	259	Bird Shot	1
.22 Win. Magnum	3	.38 Special +P	20	Bird Shot-No. 4	8
.222 Remington	3	.380 Auto	22	Bird Shot-No. 5	2
.223 Remington	24	.380 Auto, .38 Auto, .38 Super Auto & 9 mm	161	Bird Shot-No. 6	28
.243 Winchester	2	.40 caliber	1	Bird Shot-No. 8	1
.25 Auto	42	.40 S&W	13	Not Specified or Unknown – Formed Lead	268
.30 & 8 mm (7.62 mm)	33	.41 MAG	2	Other Handgun	20
.30-.06 Springfield	6	.44 MAG	19	Other Rifle	32

(c) Distribution by projectile type

3-D (remanufactured)	1	Geco (Germany)	6	PMC (Korea)	61
Activ	137	Gevelot (Canada)	1	Reload	56
Bronze or Silver Tip	9	Cast Core	1	Hi-Shok	2
Buck Shot	2	Coated Lead Bullet	23	Hydrashok	22
Buck Shot-0	1	Coated Lead Bullet-Hollow Point	100	Lead other than Bullet or Shot	1
Buck Shot-00	36	Coated Lead Bullet-Round Nose	132	Metal Case Profile	1
Buck Shot-000	4	Coated Lead Bullet-Semi-Wadcutter	5	Metal Point	4
Buck Shot-No. 1	2	Coated Lead Bullet-Truncated Cone	3	Not Specified or Unknown	147
Buck Shot-No. 3	2	Coated lead fragment	2	Partition Soft Point (Nosler)	2
Buck Shot-No. 4	1	Copper-Coated	3	Polymer Tip Boat Tail (Hornady)	1
Bullet	7	Expanding Point-coated	1	Rifled Slug	15
Bullet Core (and Core Fragment)	4	Exploding Bullet	1	Semi-Jacketed Bullet	50
Bullet-All Lead	46	Fragment of Bullet or Shot	57	Semi-Jacketed Bullet-Hollow Point (Full-Semi)	199
Bullet-All Lead – Truncated Cone	2	Full Metal Case (& TMJ)	391	Semi-Jacketed Bullet-Soft Point	70
Bullet-All Lead-Hollow Point	23	Full Metal Case-Hollow Point (& TMJ)	17	Silvertip	10
Bullet-All Lead-Round Nose	302	Gas Check	1	Special Coating (Nyclad, Teflon, Plastic)	6
Bullet-All Lead-Semi-Wadcutter	102	Gold Dot	5	Supreme – Ballistic Silvertip	1
Bullet-All Lead-Wadcutter	19	Golden Saber	1	Totally Encapsulated Metal Case (TEMC)	1

compositional profiles without consideration of manufacturer, date of production, and type of bullet. However, it is not an appropriate population suitable for frequency-of-occurrence studies in case evidence interpretation scenarios, because it was not constructed to represent appropriate weighting of production volumes, geographic distributions of bullets, rate of depletion of bullets in the general population, or other factors that may be significant in a particular case.

The ranges of concentrations of all measured elements in the test set samples are shown in Table 4. Each element was present at levels below the detection limit in some samples, so all ranges shown are from not detected (n.d.) up to the maximum measured value.

TABLE 4—*Concentration ranges and number of specimens in the test set analyzed for each element.*

Element	Concentration Range (%)	Number of Specimens
Ag	n.d. – 0.0199	1837
As	n.d. – 1.46	1750
Bi	n.d. – 0.170	1829
Cd	n.d. – 0.0479	857
Cu	n.d. – 0.834	1826
Sb	n.d. – 10.8	1837
Sn	n.d. – 5.41	1387

n.d. indicates concentration below the method detection limit.

Since the specimens in the test set are a representative subset of all evidentiary bullets on which the FBI Laboratory performed elemental analysis since 1989, these ranges reflect the overall compositions of the lead portion of projectiles involved in a wide variety of criminal activities or investigations. For each element, the number of specimens for which that analyte was determined is shown in the third column of Table 4. As discussed previously, all determinations were by ICP-AES, except 320 for arsenic and 18 for copper, which were by NAA. Antimony and silver are the only elements whose concentrations were determined in all samples. Bismuth and copper were determined in all but a few of the samples analyzed during the early part of the study period. The number of arsenic, tin, and cadmium determinations reflect the dates in which they were added to the analytical protocol. All seven elements have been determined in all samples analyzed since 1997.

The distribution of samples within these overall ranges determines the capability of discrimination among sources by each element. Antimony is intentionally controlled or specified by the manufacturers in many cases, so it has been suggested that it might have relatively poor source-discrimination capability (19). This element has been selected for further discussion. The distribution of mean antimony concentrations for the test set samples is shown in Fig. 1. To construct this figure, the samples were ordered from low to high antimony and their concentrations plotted as a function of sample order. Flat portions of the curve in Fig. 1 represent compositions that are common to many samples. The approximately 200 samples shown with antimony concentrations of less than 0.25% are soft leads, which are characteristic of .38, .357, and .45 caliber, and 9 mm full metal jacketed and semijacketed bullets. The large number of samples in the 0.75–1.25% range consist predominantly of .22 caliber bullets. Even though antimony is controlled at nominal levels by the manufacturers, the range of values observed for all products shown in Fig. 1 is quite large. When measurements are made using an analytical method having good precision, such as ICP-AES, a large number of subgroups can be distinguished within the full compositional range. Methods with relatively poor precision, such as the SS-OES method of Randich et al. (19) in their lead smelter study, are less discriminating.

The distributions of the other six elements are also shown in Figure 1. For ease of visualization, the concentrations of As and Bi are plotted on a logarithmic vertical scale; all others are linear. Although there are some similarities among them, each of these elements exhibits its own characteristic distribution among the samples in the test set. Several elements, notably Cd and to a lesser extent Sn and As are present at low concentrations in most samples, but occur in much higher levels in a few samples. Arsenic concentrations are distributed quite evenly over a logarithmic range from 0.0001 to 0.1%. The samples having As concentrations greater than 0.5% are birdshot pellets. If plotted on a linear vertical axis, the As distribution would appear visually similar to that of Sn. Both Sn and Cu

plots display minor plateaus, much smaller than those discussed for Sb. This might reflect the presence in the data set of a few products for which Sn has been intentionally added by the manufacturers. The concentrations of Ag are approximately normally distributed about a median value of 0.0022% (22 ppm) and those of Bi are approximately log-normally distributed about a median value of 0.0103% (103 ppm). The other elements do not display normal distributions either on a concentration or log concentration scale. As discussed for Sb, a wide range of compositions and correspondingly good discrimination capability are demonstrated for each element.

The precision measure associated with the mean concentration of each element in each lead projectile is represented by the standard deviation of the three replicate sample determinations. As discussed previously, in order to be conservative in estimating discrimination capabilities, the calculated standard deviations were replaced with the values shown in Table 2 for those element-sample combinations whose measured standard deviations were less than the analytical precision. The relationships between the precision, as percent relative standard deviation (RSD; standard deviation/mean × 100%) and the mean concentration are shown for each element in all projectiles in Fig. 2. For each element, the RSDs decrease rapidly with increasing element concentration at low element concentrations and then level off to ranges of 1 to 3% RSD at higher concentrations. This relationship is well known to be typical of that observed for many analytical procedures. A significant number of samples have element concentrations near the detection limits with correspondingly high RSDs. However, the ICP-AES results shown in Fig. 2 exhibit excellent precision across the concentration ranges of the majority of lead sources. The RSD values for each element taken in combination with the range across samples shown in Fig. 1 determine the discrimination capability of the analytical method. The data shown in Figs. 1 and 2 provide a strong visual demonstration that the discrimination capability of ICP-AES measurements of element concentrations in actual case-related bullet lead specimens is excellent.

A few data points, most notably for the element copper shown in Fig. 2, display RSD values that are greater than would be predicted from the general pattern of decreasing RSD with increasing concentration. Those bullets that have greater than normal measured variation in copper concentrations are scattered throughout the upper right hand portion of the copper plot. All but one of the samples with high copper RSDs are jacketed or coated bullets or are fragments of unknown bullet type. The high variation in measured copper concentrations for these samples may result from incomplete removal of copper-based jacket or coating during sample preparation. This higher than usual variation must be taken into account when comparing any of these samples with another bullet. In case work, if the measured results have better precision than the data used in this study, then the actual discrimination capability will be better than the results of this study. Alternatively, if a particular item of evidence yields poor precisions, then the discrimination capability for that evidentiary item would be worse than that calculated in this study. Hence, it is a necessary practice to use the measured precision value for each specimen in forming comparison criteria between bullets in casework.

The discrimination capability of an analytical protocol is dependent upon the number of independent element concentrations measured. The number of elements determined in each sample was three to seven, depending upon the protocol in effect at the time of sample analysis. A histogram of the number of samples analyzed as a function of the number of elements determined is shown in Fig. 3. In most of the samples (including all but one of the samples received after August 1994), the concentrations of five or more
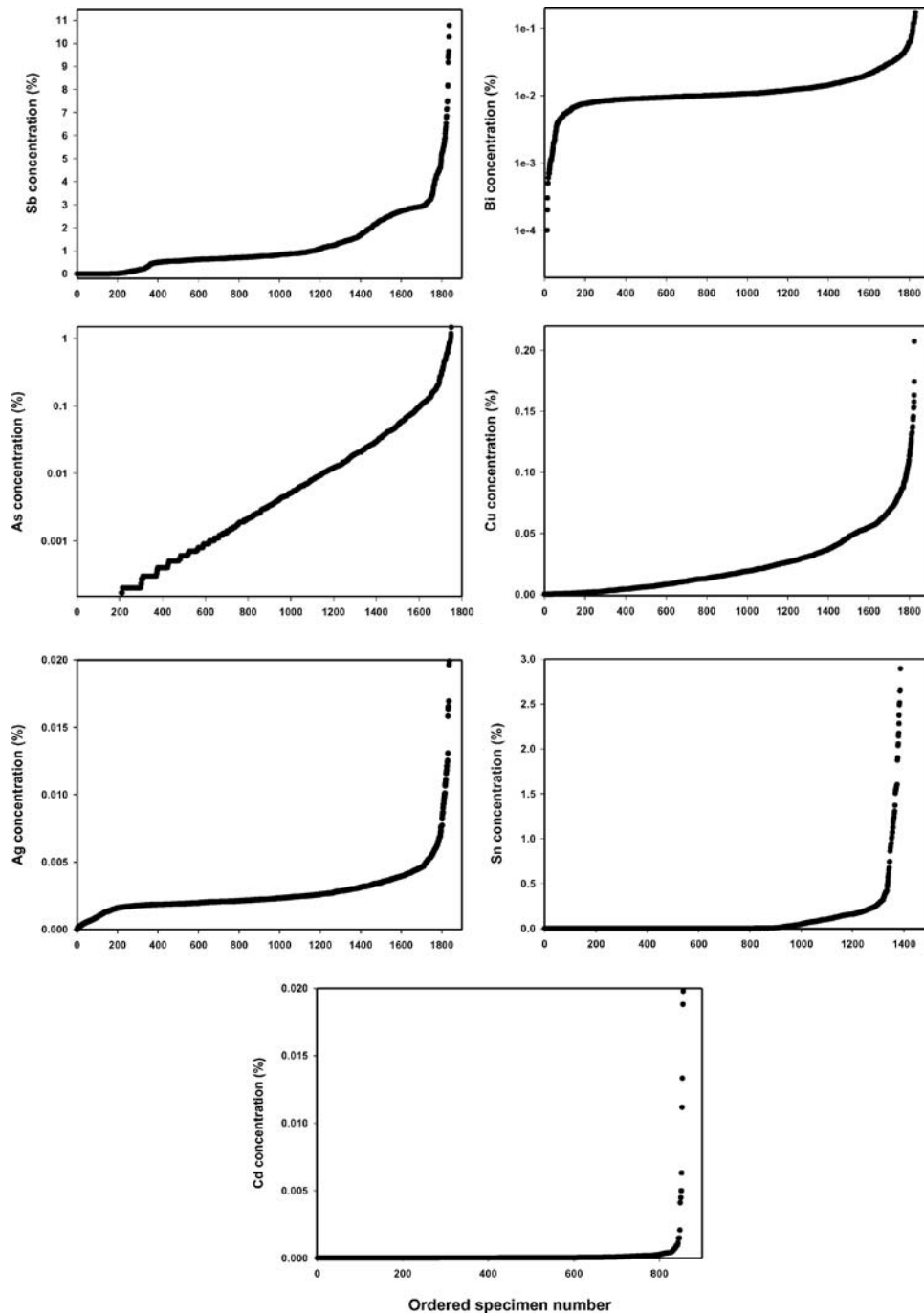
FIG. 1—*Distribution of mean concentrations of seven elements among the 1837 lead samples. Element concentrations are plotted in increasing order within each graph. Vertical scales for As and Bi are logarithmic, others are linear.*

elements were determined. In comparing two samples, only those elements that were determined in both samples can be considered. For this test set, most sample-pair comparisons consist of five or more elements. In some lead sources, not all of the elements measured were present at detectable concentrations. Figure 4 shows the number of samples as a function of the number of elements found to be present above the detection limits. For purposes of this figure and for ease of discussion, an element concentration is considered to be greater than the detection limit when the mean concentration minus two times its associated standard deviation is greater than zero. As shown in Fig. 4, four to six elements were determined to be present at concentrations greater than the detection limit in most samples.

Comparison of Figs. 3 and 4 indicates that most samples contain one or two elements at concentrations below the detection limits. In comparing two samples, elements present at levels below the detection limit in one or both samples are still compared, because they represent valid analytical characteristics of those samples.

Comparison of Samples

Each of the 1837 samples in the test set was compared with every other sample for each measured element. The total number of sample comparisons taking 1837 samples two at a time is 1,686,366. For each element that was determined in both samples, the concentrations were declared to be indistinguishable when the difference
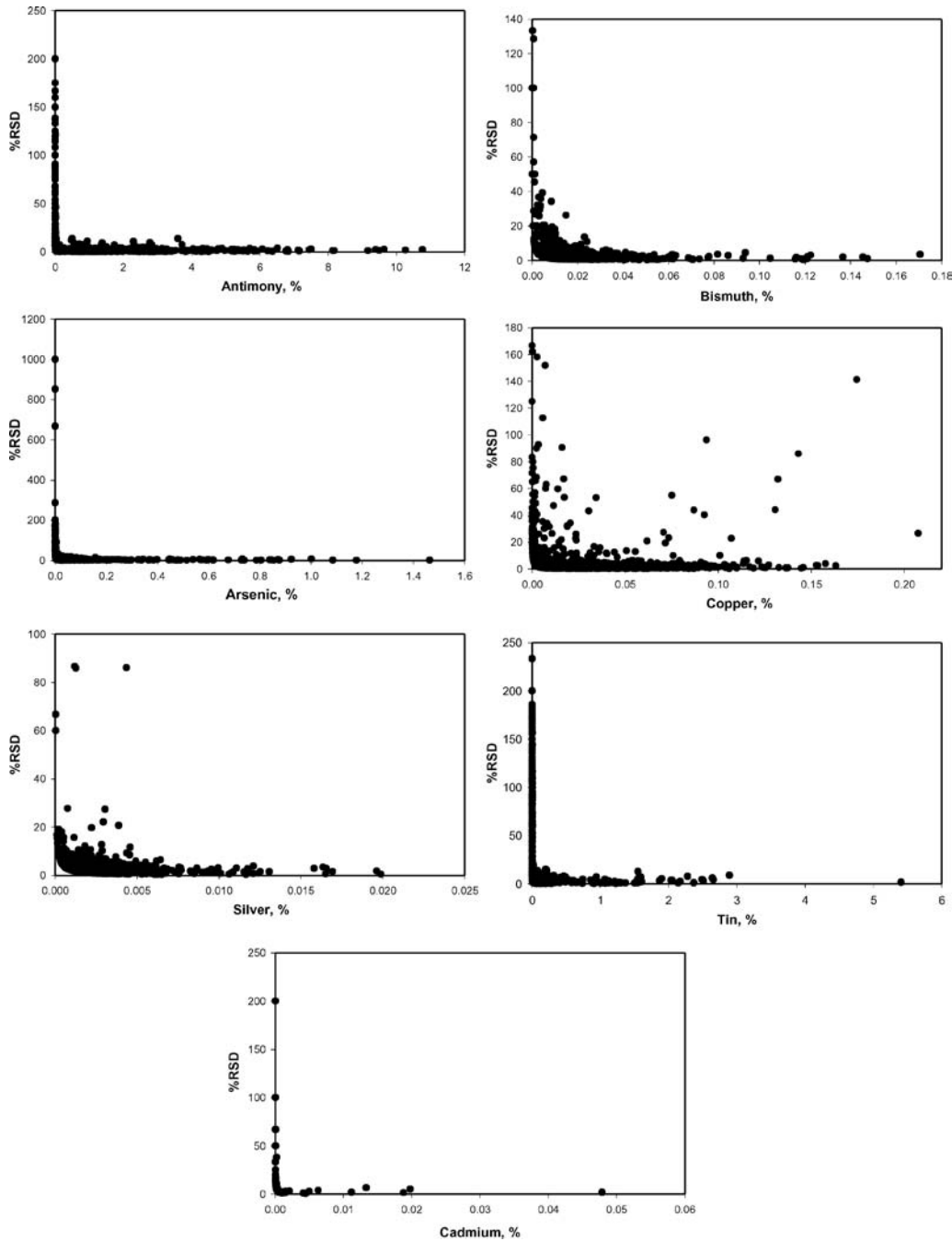
FIG. 2—*Relationships between analytical precision and mean concentrations for seven elements.*

between the mean values was less than two times the sum of the standard deviations of the two means ($2\Sigma s$). This process was repeated for each element and the number of elements that were significantly different between the two samples of each pair was determined. The **$2\Sigma s$** discrimination criterion was selected for this study because the FBI Laboratory uses it in casework during the study period; however, this process can be applied to any appropriate discrimination criterion. For each element, the number of sample pairs that were compared and the number that are indistinguishable are shown in Table 5. The frequency with which each element distinguishes among these samples is a relative measure of its discrimination capability. The order of decreasing discrimination capability is Sb, Cu > Ag > As, Bi > Sn > Cd, with each of the first five listed elements separately providing discrimination

between any randomly selected pair of samples more than 75% of the time. The elements Sb and Cu are good discriminators for all alloys. All other elements listed provide better discrimination capability for alloys that contain relatively high Sb concentrations than they do for unhardened lead alloys. These results indicate that, despite attempts by lead smelters and ammunition manufacturers to control the concentrations of some of the elements listed in Table 5 to within relatively narrow ranges, the variations remaining in the levels of these elements still provide good source discrimination. Of the elements measured, Cd provides the poorest degree of discrimination. However, even Cd alone discriminates between two sources of lead in this study 44% of the time.

It should be noted that the order of discrimination capability of the elements determined in this study differs from that described by
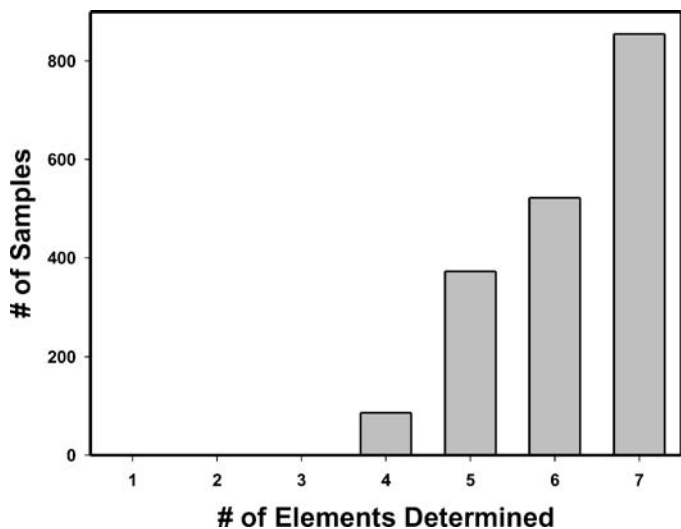
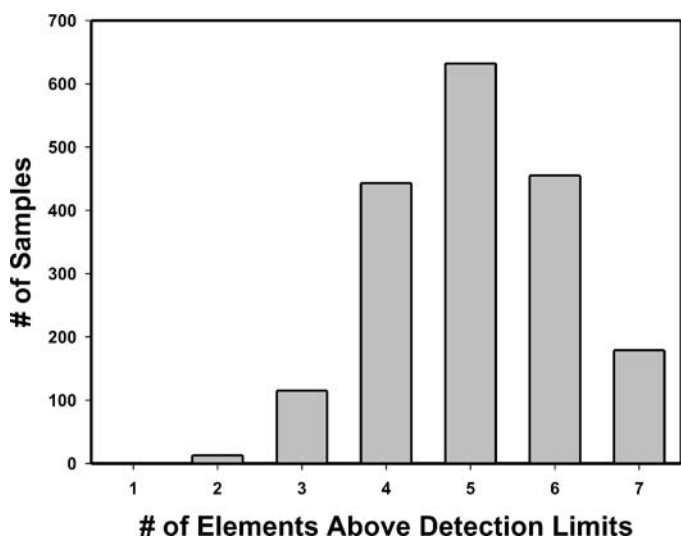FIG. 3—*Histogram of the number of samples analyzed as a function of the number of elements determined.*

FIG. 4—*Histogram of the number of samples analyzed as a function of the number of elements determined to be present at concentrations above the detection limits.*

TABLE 5—*Results of lead sample comparison considering each element separately.*

Element	Number of Comparison Pairs	Number of Indistinguishable pairs	Percent Indistinguishable
Sb	1,686,366	90,385	5.4
Cu	1,666,225	110,456	6.6
Ag	1,686,366	250,315	14.8
As	1,530,375	323,751	21.2
Bi	1,671,706	374,623	22.4
Sn	961,191	375,930	39.1
Cd	366,796	211,488	57.7

Peele et al. (8), which listed the order of decreasing discrimination capability as Cu, As > Sb > Bi, Ag. The different element orders in the two studies can be attributed to differences in bullet composition datasets tested. Peele et al. (8) studied sixteen boxes of .38 caliber lead round nose bullets produced by four manufacturers over a short production interval to determine variability within and among boxes

TABLE 6—*Pearson product moment correlation coefficients for element concentrations in 1837 bullet lead samples. Not all elements were determined in each sample.*

	Sb	Sn	Bi	Cu	Ag	Cd
As	0.371	0.291	0.0428	0.022	0.020	0.024
Sb		0.560	0.195	0.096	0.258	0.043
Sn			0.185	0.142	0.252	0.104
Bi				0.009	0.531	−0.001
Cu					0.092	−0.014
Ag						−0.021

and manufacturers. Because the samples in the present study reflect a wider range of bullet compositions in the overall bullet population, the results from this study will yield a better assessment of element discrimination capability order for distinguishing among all bullet sources.

Multivariate comparison of samples has the greatest forensic significance when the variables are independent. In fact, throughout the forensic literature, multivariate match probabilities have been calculated by multiplying individual variable probabilities when it can be demonstrated that the variables are independent. To test for the independence of element concentrations, the Pearson linear correlation coefficients between all pairs of elements were calculated. These correlation coefficients are shown in Table 6. Many of the values shown are near zero, indicating an almost total lack of linear correlation between most pairs of elements over the concentration ranges of all lead sources in this study. Lack of significant correlations between these same elements in a collection of bullet leads was previously reported by Keto (17). The concentrations of all combinations of the elements, As, Sb, and Sn exhibit small, positive correlations. The fact that the positive correlations between the concentrations of these three elements are so small is slightly surprising since they are often added in combination to increase the hardness of the lead or the sphericity of shot pellets. The maximum correlation coefficient (0.56) is between Sb and Sn. If only those bullets containing less than 0.5% of both Sb and Sn are considered, the correlation between the concentrations of these elements is reduced to 0.05. The higher degree of correlation between Sb and Sn for all bullets is likely caused, in part, by the presence of several handcast bullets, probably made from tire weights or solder containing percent levels of both Sb and Sn. The only other pair of elements with a correlation coefficient greater than 0.4 is Bi and Ag, which has a positive correlation coefficient of 0.53. These two elements are not typically controlled during either secondary smelting or bullet manufacturing. Therefore, any correlation that may exist between their concentrations probably results from covariation of their concentrations in the raw materials used to make lead alloys or consistent changes over time in their concentrations in battery lead, which is recycled into lead for bullets at secondary smelters. Visual inspection of bivariate plots and calculation of second order correlation coefficients indicate that nonlinear correlations between the concentrations of elements measured in this study also do not exist to a significant degree.

It is instructive to make the assumption that the seven element concentrations are independent and multiply together the seven individual element match frequencies derived from the discrimination percentages. The product of the seven percentages in the right-hand column of Table 5 (as fractions) results in an overall match frequency value of 1 in 176,956 (5.7×10^{-6}). This figure can be considered as an estimate of the maximum potential discrimination capability (lowest match frequency) for up to seven elements. The calculated figure should probably be considered to

TABLE 7—*Results of pairwise comparisons of 1837 lead projectiles.*

Number of elements Compared	Number of Comparisons	Number of Indistinguishable pairs	Percent Indistinguishable
2	30	0	0.000
3	7173	38	0.530
4	157041	142	0.090
5	577681	231	0.040
6	580210	213	0.037
7	364231	50	0.014
Total	1686366	674	0.040

be accurate to no better than an order of magnitude. Any correlation that exists between element concentrations will reduce the potential discrimination. At a minimum, the magnitude of the calculated frequency indicates that ICP-AES analysis of bullet lead can provide highly discriminating points for comparison of projectile leads. The match frequency can be calculated for any combination of measured elements. For example, for the four elements, Sb, Cu, Ag, and As the expected minimum match frequency is 1.1×10^{-4}. The discrimination frequency numbers shown in Table 5 and any derived results are only appropriate for the ICP-AES data in this study. Other analytical methods have different discrimination capabilities depending upon their analytical precisions. Therefore, these results cannot be directly applied to results obtained by other methods, particularly those with poorer analytical precisions.

A more straightforward method of measuring the frequency with which samples cannot be distinguished is by direct comparison of specimens using all measured elements. In trace evidence comparisons, two samples are said to be indistinguishable when none of the measured variables differ significantly between the two samples. In this study, the number of elements compared in a pair of bullet lead samples is the number of elements measured in common to both samples of the pair, which may be less than the number of elements determined in one or both of the samples. Of the 1,686,366 pairwise comparisons that were made in this study, the number of pairs that are indistinguishable using all measured elements is 674, for an overall match frequency of 1 in 2502 (4.0×10^{-4}). The first and perhaps most significant point to note with regard to these comparisons is that 1397 of the 1837 samples, or 76%, are distinguishable from every other sample in this study. This result again demonstrates the considerable evidentiary value of bullet lead compositional analysis. Further, many of the 674 matching pairs consist of two bullets that can be discriminated by obvious differences in their caliber or some other physical characteristic. In cases where this information is available, the combination of compositional and physical comparison of bullets will provide greater discrimination capability than the match frequencies determined in this study.

The distributions of the 674 indistinguishable sample pairs among numbers of elements compared are listed in Table 7. As expected, the frequency of indistinguishable pairs decreases as the number of elements compared increases. The overall improvement in discrimination capability obtained when the number of elements compared is increased from 4 to 7 is approximately 7-fold. Values of random match frequencies in the range of 0.014% to 0.04% for 5 to 7 element matches of bullet lead are consistent with our expectations based on evaluation of element concentration data for a variety of materials. The match frequencies actually measured are all at least an order of magnitude lower than the maximum values calculated previously, reflecting the lack of complete independence of element concentrations.

Although only 242 of the 1837 samples contain less than 0.05% antimony, they account for approximately half (335) of the matching pairs. The higher frequency of matching compositions among these unalloyed leads than in the harder leads is a reflection of the lower discrimination capability of antimony at low concentrations. Also, many of the leads that contain no added Sb, also contain low concentrations of other elements, such as As, which are impurities in the Sb alloy added to the lead. Another factor that must be considered when evaluating the frequency of indistinguishable sample pairs is that several of the samples in this study have many matches. This results either because these samples have wide ranges in the concentrations of one element or because a key discriminating element was not determined. For example, in this study there are nine samples that are each indistinguishable from 20 or more other samples, with a worst case of two samples each having 32 indistinguishable partners. All of these samples exhibit poor precision in their copper concentrations or levels of several elements that are below the detection limits. In case work, the significance of two samples having indistinguishable compositions should be given less weight when they have low antimony concentrations, or several elements are present below their detection limits or have poor precisions, than for cases where all seven elements are present and measured with good precision.

In this study, match frequencies have been calculated by comparing each bullet to every other bullet without regard to physical characteristics. This is the appropriate approach to use in cases in which the recovered bullet is too damaged or is a fragment too small for useful physical comparison. In cases where the type or manufacturer of the questioned bullet can be determined, it can be argued that the appropriate population of bullets for probability assessment is some subset of the complete dataset. The effects of subsetting can be illustrated by testing the statement made by Randich et al. (19) that compositional analysis does little to discriminate among sources of 0.7–0.75% Sb .22 caliber bullets. They base this conclusion on their belief that for nearly all .22 caliber bullets in this narrow Sb concentration range, the other elements have limited concentration ranges because they are controlled by the lead smelters. In the present study, there are 401 samples of .22 caliber bullets listed in Table 3. Of the 80,200 pairs that can be made from these 401 samples, the number of matching pairs is 51, for a match frequency of 1 in 1573. This result is within the range of values shown in Table 7 indicating that the frequency of matches in the subset of .22 caliber bullets is no different than the overall match frequency for all bullets despite the perception that the range of each element is more limited than it is across all sources. In general, it is not productive to divide the test database or the larger general historical data collection into smaller groups based on caliber, bullet style, or manufacturer for frequency calculations, because the number of specimens within each category is too small to obtain statistically meaningful results.

Conclusions and Significance

The purpose of this study was to provide quantitative data that can be used in a general sense to assess the significance of the results of bullet lead compositional comparisons. It is not current practice, nor do we recommend the calculation of probabilities to determine significance of a match in casework. However, we can use the information gained in this study to make several general statements about the likelihood of random, cross-source matches of bullet compositions. Comparison of the concentrations of the selected elements in bullet and shot lead can provide strong circumstantial evidence for use in forensics. The significance of an association

increases with increasing number of measured elements, rarity of composition, and improved precision of the analytical results. For a representative sampling of all bullets involved in FBI case work over the 1989 to 2002 period in the U.S., the probability of two unrelated bullets having indistinguishable compositions is on the order of 10^{-3} to 10^{-5}. Matches for soft lead are not as significant as those for hardened lead, because some of the comparison elements are more likely to be present at concentrations too low for determination with good precision. In cases where multiple compositions occur in crime scene samples and the same compositions occur in subject-associated samples, the value of the association can be very strong. That is, given that the probability that a recovered bullet from a crime scene and another truly unrelated bullet associated with a subject having indistinguishable compositions is low, the likelihood that two or more matching pairs each with different compositions would occur coincidentally is extremely low (i.e., approximately equal to the product of the two individual probabilities).

During the time period coincident with this study and following its completion, several other related publications have appeared. Principal among these are the publication of the results of a study of lead smelter samples by Randich et al. (19) and the release of a report of the National Research Council (NRC) Committee on Scientific Assessment of Bullet Lead Elemental Composition Comparison (20). These publications address aspects of either the FBI's method of comparing bullet lead compositional data or utilize the 1837 sample data set of this study. A full discussion of the many issues raised concerning this data set and its utilization is beyond the scope of this paper. However, a few important comments are provided herein concerning the study reported by Randich et al. and the NRC Report.

The conclusions reached in this study are in sharp contrast with the principal statements made by Randich et al. (19) in their study. For their discussion, they utilized SS-OES measurements made of small lead disks by lead smelters and extrapolated the results so as to represent lots containing up to 100,000 kg of lead. They used compositional data from 122 production run samples of nominal 0.7–0.75% Sb lead alloys refined in two smelters for sale to ammunition manufacturers for use in .22 caliber lead round-nosed bullets. Their data did not include any precision measures, but it is evident that it has considerably poorer precision than the ICP-AES data used in our study, as demonstrated by their wide variability reported for multiple measures of the same lot and their statement that RSDs are in the 5–10% range. Without stating the criteria used to determine whether or not composition measurements were distinguishable, they concluded that bullet lead compositional comparison has little significance. In fact, careful evaluation of their data supports the general conclusions of our study, despite their poorer analytical precisions and their definition of a large smelter production run with large compositional variability as a single source. They report only 13 matching pairs among the 7381 pairwise combinations of their 122 samples. Considering that their samples consisted of only one product line and two smelters, this match frequency of 1 in 568 is in surprisingly good agreement with the results of our study and in sharp contrast with their own general conclusions.

In 2003, the FBI Laboratory contracted with the NRC to evaluate all aspects of the compositional comparison of bullet lead and its presentation in court. The NRC convened a Committee on Scientific Assessment of Bullet Lead Elemental Composition Comparison, which met four times, interviewed members of the relevant manufacturing and forensic communities, researched FBI cases and evaluated testimony, and performed its own research studies concerning significance of matching compositions. At the request of the Committee, we provided them with the data developed in the study reported herein. The Committee produced a final report in February 2004, a large portion of which contains a statistical analysis of our 1837 sample data set. The NRC Committee made an effort to utilize the data set to determine the likelihood of matching compositions among randomly selected bullets in the general population, or what we have referred to as a frequency of occurrence study. The Committee commented that the samples in this data set were not selected in a manner to represent any larger population of bullets, such as those available to a subject in a criminal investigation. As a result, they felt that the bullets in our study are likely to be more dissimilar than a randomly selected set of samples that could include a greater number of sets of bullets from the same manufacturing source. As stated previously, the samples in the 1837 data set were selected to represent as many independent sources as possible with equal weighting of data from one specimen each, making our data set not optimum for the NRC Committee's intended use. The NRC also suggested that the FBI consider alternatives to the **2Σs** method of declaring two bullets to be analytically indistinguishable. The data set used in this study would be applicable to evaluate coincidental match rates using other comparison criteria.

Some participants in the legal process have expressed interest in having probability statistics calculated to help them assess the significance of the evidence. In part, that is because they are conditioned to expect this data from its use in interpretation of DNA profiles. We believe that it is neither possible nor appropriate to calculate reliable probabilities of chance occurrence of indistinguishable items of non-biological trace evidence. This is particularly true for manufactured items and for comparison of highly discriminating variables such as elemental concentrations. Reasons for this include the lack of databases that are both of sufficient size and representative of the distributions of the measured variables in the evidentiary material. For items whose measured characteristics in the population are either geographically or temporally variable, such databases are impossible to obtain. This fact has proven to be one of the greatest drawbacks preventing the adoption of probabilistic methods of evaluating evidentiary significance. However, the inability to acquire databases appropriate for precise frequency of occurrence calculations does not mean that the significance of two samples having indistinguishable properties is low, only that it cannot be calculated precisely. The inability to calculate precise and accurate significance statistics applies generally to all comparative characteristics of trace evidence, except possibly those whose distributions are controlled by time-invariant genetic rules (e.g., biomarkers), or whose characteristics are deemed to be unique (e.g., fingerprints, toolmarks). Despite the lack of precise significance statistics, comparative trace evidence examinations continue to provide highly significant conclusions and meaningful assistance to all participants in the legal process.

References

1. Lukens HR, Schlesinger HL, Guinn VP, Hackleman RP. Forensic neutron activation analysis of bullet-lead specimens. Washington DC: US Atomic Energy Commission Report GA-10141, 1970.
2. Peters CA. The basis for comparative bullet lead analysis. Forensic Sci Comm. 2002;4(3) (http://www.fbi.gov/hq/lab/fsc/backissu/july2002/peters.htm).
3. Lukens HR, Guinn VP. Comparison of bullet lead specimens by nondestructive neutron activation analysis. J Forensic Sci 1971;16:301–8. [PubMed]
4. Guinn VP. NAA of bullet-lead evidence specimens in criminal cases. J Radioanal Chem 1982;72(1–2):645–63.
5. Desai HB, Parthasarathy R. A radiochemical neutron activation analysis method for the determination of tin, arsenic, copper and antimony for the forensic comparison of bullet lead specimens. J Radioanal Chem 1983;77(1):235–40.

6. Brandone A, Piancone GF. Characterisation of firearms and bullets by instrumental neutron activation analysis. Intl J Appl Radioactive Isotopes 1984;35(5):359–64.

7. Guy RD, Pate BD. Studies of the trace element content of bullet lead and jacket material. J Forensic Sci 1973;18:87–92, also J Radioanal Chem 15:135–42. [PubMed]

8. Peele ER, Havekost DG, Peters CA, Riley JP, Halberstam RC, Koons RD. Comparison of bullets using the elemental composition of the lead component. Proceedings of the International Symposium on the Forensic Aspects of Trace Evidence; 1991 June 24–28; Quantico, VA. Washington DC: US Government Printing Office, 57–68.

9. Krishnan SS. Trace element analysis by atomic absorption spectrometry and neutron activation analysis in the investigation of shooting cases. Canad Soc Forensic Sci J 1973;6:55–77.

10. Gillespie KA, Krishnan SS. Analysis of lead shot—a comparison of analyses using atomic absorption spectrophotometry and neutron activation analysis. Canad Soc Forensic Sci J 1969;2:95–103.

11. Brunelle RL, Hoffman CM, Snow KB. Comparison of elemental composition of pistol bullets by atomic absorption: preliminary study. J Assoc Off Anal Chem 1970;53:470–4.

12. Haney MA, Gallagher JF. Differentiation of bullets by spark source mass spectrometry. J Forensic Sci 1975;20:484–500. [PubMed]

13. Haney MA, Gallagher JF. Elemental analysis of bullet lead by spark source mass spectrometry. Anal Chem 1975;47:62–5. [PubMed]

14. Peters CA, Havekost DG, Koons RD. Multielement analysis of bullet lead by inductively coupled plasma—atomic emission spectrometry. Crime Lab Digest 1988;15:33–8.

15. Suzuki Y, Marumo Y. Determination of trace impurities in lead shotgun pellets by ICP-MS. Anal Sci 1996;12:129–32.

16. Dufosse T, Touron P. Comparison of bullet alloys by chemical analysis: use of the ICP-MS method. Forensic Sci Intl 1998;91:197–206.

17. Keto RO. Analysis and comparison of bullet leads by inductively-coupled plasma mass spectrometry. J Forensic Sci 1999;44:1020–6.

18. Koons RD, Grant DM. Compositional variation in bullet lead manufacture. J Forensic Sci 2002;47:950–8. [PubMed]

19. Randich E, Duerfeldt W, McLendon W, Tobin W. A metallurgical review of the interpretation of bullet lead compositional analysis. Forensic Sci Intl 2002;127:174–91.

20. National Research Council (U.S.). Committee on Scientific Assessment of Bullet Lead Elemental Composition Comparison. Forensic Analysis: Weighing Bullet Lead Evidence. Washington, DC: National Academy Press, 2004.