

The Current State of Bullet-Lead Evidence*

D.H. Kaye**

For more than forty years, analyses of the composition of bullet lead have been used to identify the source of bullets.¹ It is a method of last resort. In most cases, markings from a bullet's manufacture and its passage through the gun are much easier to obtain, and these marks can help associate the bullet with a particular gun or set of guns.² When the bullet is too deformed or a bullet fragment is too small, however, striations and toolmarks are not available. Nevertheless, the lead alloy used in bullets contains a variety of trace elements, and the concentrations of these elements vary across production runs. If unused cartridges can be found in the possession of a suspect, similarities in these concentrations might forge a link between the crime-scene bullet and other bullets owned by the suspect.

The theory sounds simple, but its implementation is not. The elemental concentrations in samples of lead removed from the crime-scene bullet and the bullets found in the defendant's possession error must be ascertained, and the errors in the measurement process must be small relative to the variations in the concentrations of the elements across bullets. In other words, suitably precise instrumentation to measure trace element concentration is required. The modern technique to accomplish this is known as inductively coupled plasma-optical emission spectroscopy. The precision of ICP-OES as applied to bullet lead can be examined in laboratory studies.

Once the pertinent elemental concentrations have been ascertained, the differences in the samples being compared must be interpreted. Are the concentrations sufficiently similar to indicate that the bullets come from the same production run? How firm is this inference? That is, would bullets

*© 2006 D.H. Kaye

**Regents' Professor and Fellow, Center for the Study of Law, Science, and Technology, Arizona State University. This article is adapted from 4 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY ch. 37 (David L. Faigman et al. eds., 2006–2007 ed., forthcoming).

¹NATIONAL RESEARCH COUNCIL, COMMITTEE ON SCIENTIFIC ASSESSMENT OF BULLET LEAD ELEMENTAL COMPOSITION COMPARISON, FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE 1 (2004) [hereinafter cited as NRC REPORT]; FBI National Press Office, National Academy of Sciences Releases FBI-commissioned Study on Bullet Lead Analysis, Feb. 10, 2004, available at <http://www.fbi.gov/pressrel/pressrel04/bullet021004.htm> (“[s]ince the early 1980's this technique has been used in approximately 2,500 cases and in less than 20% of those cases, bullet lead compositional analysis was presented as circumstantial evidence in court. “).

²See, e.g., 4 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY ch. 36 (David L. Faigman et al, eds., 2005).

manufactured at different times or places also be this similar? Analyzing bullets known to be made at various times and places can help answer these scientific questions.

Finally, even if one can be reasonably confident that the bullets were manufactured at the same time and place, would they end up in the same location in the distribution chain and hence in defendant's possession? Or do bullets from different production runs wind up in the same boxes?

Despite a glaring lack of data on the manufacturing process and the geographical and temporal distribution of bullets, expert witnesses have been willing to testify that two bullets come from "the same manufacturer's source (melt) of lead"³ or even "the same box of ammunition."⁴ As the defense bar finally began to recognize the vulnerability of this kind of testimony,⁵ the FBI, as the sole agency performing bullet-lead ICP-OES, decided to address such issues. It commissioned a study by the National Research Council (NRC) of this compositional analysis of bullet lead (CABL).⁶ The report concluded that the ICP-OES method of determining the concentrations of elements in bullet lead is valid and reliable, but that research does *not* support testimony that two bullets originated from the same manufacturer, from the same melt of lead, from the same box of bullets, or on the same date.⁷

When the report was released, the FBI responded by reaffirming its view that the testimony of its agents was scientifically sound. It announced that "[t]he basis of bullet lead compositional analysis is supported by approximately 50 peer-reviewed articles found in scientific publications beginning in the early 1970's. Published research and validation studies have continued to

³NRC REPORT, *supra* note 1, at 93.

⁴*Id.* at 92; *see also* Commonwealth v. Daye, No. 11238-11246, 2005 WL 1971027, 19 Mass. L. Rptr. 674 (Mass. Super. Ct., Aug. 3, 2005) ("Special Agent Riley testified as an expert as to how, thorough testing of the bullets, he was able to determine that one of the bullets found in Patricia Paglia's body and the bullet found in the basement of the Rye, N.H. home were 'from the same box of ammunition, or from another box of ammunition that was produced at the same place on or about the same date.'").

⁵*See* Mark Hansen, *Bullet Proof*, A.B.A.J., Sept. 2004, at 58.

⁶The FBI requested "an impartial scientific assessment of the soundness of the scientific principles underlying CABL to determine the optimum manner for conducting the examination and to establish scientifically valid conclusions that can be reached using the examination." NRC REPORT, *supra* note 1, at ix.

⁷Previous commentary to the same effect includes Edward J. Imwinkelried & William A. Tobin, *Comparative Bullet Lead Analysis (CBLA) Evidence: Valid Inference or Ipse Dixit?*, 28 OKLA. CITY U. L. REV. 43 (2003).

demonstrate the usefulness of the measurement of trace elements within bullet lead.”⁸ “The science.” it insisted, “has continually withstood legal challenges in federal, state, and local criminal courts.”⁹ With the NRC committee’s handwriting on the wall and indications of growing skepticism from the courts, however, the FBI announced on September 1, 2005, that “after extensive study and consideration, it will no longer conduct the examination of bullet lead.”¹⁰

Despite the discontinuance of the CABL testing in the United States, the scientific and evidentiary status of CABL remains significant, for two reasons. First, the issue continues to arise in post-conviction proceedings, and second, the recommendations of the NRC committee as to the type of testimony that should be permitted has implications for other forensic individualization techniques.¹¹ Consequently, this article describes the developing case law on CABL and the scientific and statistical issues that have engendered controversy.

I. CASES BEFORE AND AFTER

Prior to the NRC report, numerous courts readily admitted testimony associating a bullet in a suspect’s possession with a bullet used in a crime.¹² For some thirty years, however, neither bar nor bench nor forensic scientists seriously probed the claims of FBI analysts that they could trace bullets to the same box, production run, or the like. In recent years, this situation began to

⁸FBI National Press Office, *supra* note 1.

⁹*Id.*

¹⁰FBI National Press Office, FBI Laboratory Announces Discontinuation of Bullet Lead Examinations, Sept. 1, 2005, available at http://www.fbi.gov/pressrel/pressrel05/bullet_lead_analysis.htm.

¹¹See D.H. KAYE ET AL., THE NEW WIGMORE ON EVIDENCE: EXPERT EVIDENCE ch. 12 (2004).

¹²E.g., United States v. Davis, 103 F.3d 660 (8th Cir. 1996); Commonwealth v. Daye, 587 N.E.2d 194 (Mass. 1992); State v. Grube, 883 P.2d 1069 (Idaho 1994); State v. Ware, 338 N.W.2d 707 (Iowa 1983); Ragland v. Commonwealth, No. 2002-SC-0388-MR, 2004 WL 2623926 (Ky. Nov. 18, 2004), reheard, 191 S.W.3d 569 (Ky. 2006); Mathes v. City of Omaha, 576 N.W.2d 181 (Neb. 1998); State v. Noel, 723 A.2d 602 (N.J. 1999); State v. King, 546 S.E.2d 575, 584 (N.C. 2001); State v. Krummacher, 523 P.2d 1009, 1017 (Or. 1974); State v. Robertson, 130 S.W.3d 842, 850 (Tenn. Ct. Crim. App. 2003); cf. People ex rel. S.G., 91 P.3d 443, 451-52 (Colo. Ct. App. 2004) (bullet matching evidence used to support a verdict despite other perjury of an expert witness).

change. In *United States v. Mikos*,¹³ a federal district court issued an unreported opinion finding that

the required standard of scientific reliability is met only as to the proposed opinion testimony that the elements composition of the bullets recovered from the body is indistinguishable from the composition of the bullets found in the Defendant's car. There is no body of data to corroborate the government's expert's further opinion that from this finding it follows that the bullets must or even likely came from the same batch or melt.¹⁴

Accordingly, the court allowed the FBI agent to testify to the chemical similarities in the bullets, but it excluded testimony on “how probable is it that bullets having compositions so nearly identical as to be deemed ‘analytically indistinguishable’ came from the same ‘source.’”¹⁵

In *State v. Behn*,¹⁶ the same FBI agent whose testimony was constrained in *Mikos* freely “opined that the lead in the fragments recovered from the decedent and the lead in bullets the defendant possessed were analytically indistinguishable, that both the lead fragments recovered from the decedent's body and the defendant's bullets came from the same source of lead, and both the fragments recovered from the decedent's body and the defendant's bullets came from the same box or boxes and were packaged on the same date by the manufacturer.”¹⁷ The defendant, who was convicted of murder and armed robbery, sought postconviction relief based on “scientific developments which took place after his trial.”¹⁸ Pointing to newer studies that contradicted the claim that matching bullets necessarily originated from the same “source of lead,” a state intermediate appellate court concluded that “the expert testimony was based on erroneous scientific foundations and its admission met the

¹³No. 02 CR 137, 2003 WL 22922197 (N.D. Ill. Dec. 9, 2003).

¹⁴*Id.* at *6. Why the opinion is unpublished is mysterious. It breaks new legal ground with reasoning that applies to a substantial number of other cases.

¹⁵*Id.* at *3.

¹⁶868 A.2d 329 (N.J. Super. Ct. App. Div. 2005).

¹⁷*Id.* at 335.

¹⁸*Id.* at 331.

requirements for granting a new trial on the ground of newly discovered evidence.”¹⁹

In the wake of the NRC report, judicial rejection of same-source testimony has only accelerated. *Ragland v. Commonwealth*²⁰ starkly reveals the potential impact of the report. Late in 2004, the Supreme Court of Kentucky penned an opinion rejecting an attack on bullet-lead testimony.²¹ Shane Layton Ragland was convicted for murdering a fellow student at the University of Kentucky. The murder occurred ten years earlier, when the victim was celebrating his twenty-first birthday with some friends on the front porch of his residence. A rifle shot rang out. No one saw who fired it or from where it was fired. But there was no doubt as to where it landed. Fragments of the fatal bullet were recovered from the student’s skull during the postmortem examination.

The police investigation fizzled until a break came in 2000. Ragland’s ex-girlfriend informed the police officers that back in 1995, Ragland told her that he did the shooting because the victim had caused his college fraternity to blackball him. Ragland, she said, even showed her the rifle and later said that he had hidden it in his mother’s home. A search of the home revealed a .243 caliber Wetherby Vanguard rifle with three unspent .243 caliber bullets in the chamber. A search of Ragland’s father’s home revealed an ammunition box containing seventeen unspent .243 caliber bullets. A label on the box indicated the Winchester Ammunition Company had manufactured the bullets on April 28, 1994.

An FBI forensic scientist conducted an analysis of the three bullets found in the Wetherby Vanguard rifle, of more bullets found in the ammunition box, and of the fragment of the bullet from the skull. According to the Kentucky Supreme Court’s 2004 opinion, “[s]he testified at trial that one of the bullets recovered from the rifle and nine of the bullets found in the ammunition box were ‘analytically indistinguishable’ in composition from the bullet that killed [the student], a finding she described as ‘consistent with’ the bullets having originated from the same source of molten lead.”²²

Ragland had moved to exclude this testimony under the standard for admitting scientific evidence adopted by the United States Supreme Court in

¹⁹*Id.* at 331–32. *But cf.* *Commonwealth v. Fisher*, 870 A.2d 864 (Pa. 2005) (holding that a press release describing the NRC report did not provide a basis for questioning a capital verdict that was unavailable to a petitioner for post-conviction relief at the time of the trial).

²⁰191 S.W.3d 569 (Ky. 2006).

²¹*Ragland v. Commonwealth*, No. 2002-SC-0388-MR, 2004 WL 2623926 (Ky. Nov. 18, 2004).

²²*Id.* at *2.

Daubert v. Merrell Dow Pharmaceuticals, Inc.,²³ and incorporated into Kentucky law by that state's courts.²⁴ The trial court admitted the evidence, the jury convicted, and Ragland was sentenced to thirty years in prison.

In its 2004 opinion, the state supreme court minimized the FBI analyst's testimony. It noted that she "never opined that the analytically indistinguishable bullets did originate from the same source," but merely stated that the bullets were "consistent with' their having the same source."²⁵ The court insisted that this modest assertion satisfied *Daubert*. It reasoned as follows:

The test of admissibility is not whether a particular scientific opinion has the best foundation or whether it is demonstrably correct. Rather, the test is whether the particular opinion is based on valid reasoning and reliable methodology. . . . [¶] Again, [the expert] did not testify that the bullets must have come from the same batch of molten lead but only that their metallurgical composition was consistent with having come from the same batch. . . . [¶] Other jurisdictions have admitted similar evidence of comparative bullet lead analysis. . . . [The expert] testified that the analysis has been subjected to peer review in a number of scientific journals. We conclude that there was substantial evidence to support the trial court's finding that the methodology used to determine the metallurgical composition of lead bullets and [the analyst's] reasoning that the fact that two or more bullets have an analytically indistinguishable metallurgical composition is consistent with their having come from the same source were both scientifically reliable. . . .²⁶

In the end, the supreme court reversed the conviction, but on unrelated grounds.

²³509 U.S. 579 (1993). *Daubert* requires the proponent of scientific evidence to show that the theory or technique is scientifically valid. *See generally, e.g.,* KAYE ET AL., *supra* note 11, ch. 6.

²⁴*Mitchell v. Commonwealth*, 908 S.W.2d 100, 101–02 (Ky. 1995), *overruled in part on other grounds by* *Fugate v. Commonwealth*, 993 S.W.2d 931, 937 (Ky. 1999).

²⁵*Ragland v. Commonwealth*, No. 2002-SC-0388-MR, 2004 WL 2623926, at *12 (Ky. Nov. 18, 2004).

²⁶*Id.*

Two years later the same court issued a second opinion, having decided to rehear the case.²⁷ This time, the supreme court deemed the testimony to be much less innocuous. The court observed that the expert “never opined that the analytically indistinguishable bullets did originate from the same batch of molten lead, though she strongly suggested as much throughout her testimony”²⁸ For example, the new opinion revealed that she also testified that “we have seen that bullets that come from the same source of lead will have the same composition, and bullets from different sources of lead have different composition,” that “you expect to find bullets of the same composition in a given box or other boxes, but, you know, it's the same type of ammunition that's produced at the same time. That's the—that's where you expect to find compositional similarities.”²⁹ Quoting extensively from the NRC report, the supreme court held that such innuendo was inadmissible under *Daubert*.³⁰

Ragland is not the only post-NRC-report case to condemn testimony that cannot be squared with the committee's recommendations. In *Clemons v. State*,³¹ Maryland's highest court examined testimony that “[w]here the bullet and the cartridges are analytically indistinguishable . . . [i]t means . . . it came from the same smelt of lead.”³² Based on its review of the literature, especially the NRC report, the court held that the introduction of this testimony violated the general-acceptance standard for admitting scientific evidence in that state.³³

II. THE FUTURE

²⁷*Ragland v. Commonwealth*, 191 S.W.3d 569 (Ky. 2006).

²⁸*Id.* at 576.

²⁹*Id.*

³⁰*Id.* at 580.

³¹896 A.2d 1059, 1079 (Md. 2006).

³²*Id.* at 1067. On the voir dire, the FBI agent had stated:

We can't tell them apart. That tells us that they were manufactured or they were likely manufactured in the same pot of lead at a bullet manufacturer. So out of the whole population of nine billion or so cartridges that are produced here in the United States, we can narrow it down to tens of thousands of bullets being produced that would have the same composition.

Id. at 1065 n.7.

³³For a description of this standard, see generally KAYE ET AL., *supra* note 11, ch. 5.

The original opinion in *Ragland* may well be the zenith for efforts to harmonize the admission of CABL testimony that posits a common source for matching bullet-lead samples with the scientific-validity or general-acceptance standards for scientific evidence. The NRC report makes it extremely difficult for courts to discern scientific acceptance or validity in the once-typical statements about the common origin of matching bullet-lead samples. But what if the expert merely reports (1) that the bullets match (are “analytically indistinguishable,” “consistent with a common source,” or the like), and (2) that this enhances the probability the bullets came from the same source (or that it is more likely that bullets from a common source would match than it is that bullets from different sources would match)?

In other contexts, some commentators have maintained that such testimony should be admissible even though it is not possible to quantify the probabilities involved.³⁴ For example, the jury can benefit from learning that two sets of fingerprints or handwriting match even if the available data do not allow an expert to specify the probability of a coincidental match.

At first glance, one would think that the FBI’s decision to cease CABL testing renders this question purely academic in the context of bullet-lead evidence. Because no state laboratories use CABL, there will be no *new* cases in which CABL analysts can strive to keep their testimony within the bounds prescribed by the NRC committee. Nonetheless, a few existing cases may remain on direct appeal, and, as *State v. Behn* shows, new trials of old cases may be ordered. Consequently, whether there is any room for testimony about CABL testing is not without practical import, and the question of whether all the NRC recommendations for presenting such testimony should be translated into rules of evidence may yet arise.

A few predictions and suggestions therefore may be appropriate. Generally speaking, uncontradicted descriptions of scientific knowledge in the report will be given great weight, as they were in *Ragland*. But not all the NRC recommendations are scientific in nature. Several are judgments about how the legal system should be structured to provide information on scientific findings to jurors. For instance, the committee wrote that the possibility of coincidental matches “should be acknowledged in the laboratory report and by the expert witness on direct examination.”³⁵ Although including such boilerplate warnings is desirable, a rule of evidence that regards the failure to provide them as grounds for excluding the testimony may strike some courts as excessive. Certainly, a transgression of this proposed rule will not lead to reversal if the expert discloses the possibility on cross-examination. Because not all departures

³⁴*See id.* at ch. 12.

³⁵NRC REPORT, *supra* note 1, at 107.

from the report's recommendations for the presentation of CABL evidence would constitute violations of the scientific-validity or general-acceptance standards for scientific evidence, courts may yet have to ask whether testimony that was relatively restrained but that cannot be squared with some aspects of NRC report's legal analysis was erroneous.³⁶

Such determinations will require some appreciation of the logic underlying the committee's legal recommendations. The inescapable fact is that the situation with CABL (and other trace evidence from manufactured items such as paint or textiles) is more complicated than that of trace evidence such as DNA or toolmarks. The later evidence characterizes individuals or individual items. For example, the meaning of a "common source" for two samples of handwriting, or two bullets with the same striations produced by the passage through the barrel of a rifle, is clear. The same person might have written both documents, and the same rifle might have produced the marks on the bullets. But the meaning and significance of a "common source" for two bullets is much less clear. Is it a single box of ammunition? A huge melt of lead? Something in between? The NRC committee sought to finesse this problem by introducing the idea of a "compositionally indistinguishable volume of lead" (CIVL). It suggested that very limited testimony about the possibility that two bullets came from the same CIVL might be legally acceptable.³⁷

³⁶In post-conviction proceedings, the issue might be whether testimony that did not conform to the NRC report's prescriptions violated the right to due process of law rather than a state rule of evidence.

³⁷The NRC Report enumerates the following findings and recommendations:

Finding: CABL is sufficiently reliable to support testimony that bullets from the same compositionally indistinguishable volume of lead (CIVL) are more likely to be analytically indistinguishable than bullets from different CIVLs. An examiner may also testify that having CABL evidence that two bullets are analytically indistinguishable increases the probability that two bullets came from the same CIVL, versus no evidence of match status.

Recommendation: Interpretation and testimony of examiners should be limited as described above and assessed regularly.

Finding: Although it has been demonstrated that there are a large number of different compositionally indistinguishable volumes of lead (CIVLs), there is evidence that bullets from different CIVLs can sometimes coincidentally be analytically indistinguishable.

Recommendation: The possible existence of coincidentally indistinguishable CIVLs should be acknowledged in the laboratory report and by the expert witness on direct examination.

Finding: The available data do not support any statement that a crime bullet came from, or

To evaluate whether the committee’s solution is satisfactory, we must consider the reasoning that undergirds CABL. We address three topics: first, measuring the elemental concentrations in bullet-lead samples; second, deciding that the measurements are so similar that the samples match; and third, drawing inferences from the degree of matching to the source of the bullets. As we have seen, it is the last issue that has led to holdings that CABL findings are inadmissible under the normal rules for scientific evidence.

A. Measuring Concentrations of Elements

The first step in CABL is measuring the concentrations of the various trace elements in bullet lead. Just as the weight of an object cannot be ascertained down to last decimal point, such measurements are subject to variability, often called “measurement error.”³⁸ There is little controversy about the NRC committee’s finding that “[t]he current analytical technology used by the FBI—inductively coupled plasma-optical emission spectroscopy (ICP-OES)—is appropriate and is currently the best available technology for the

is likely to have come from, a particular box of ammunition, and references to “boxes” of ammunition in any form is seriously misleading under Federal Rule of Evidence 403. Testimony that the crime bullet came from the defendant’s box or from a box manufactured at the same time is also objectionable because it may be understood as implying a substantial probability that the bullet came from defendant’s box.

Finding: Compositional analysis of bullet lead data alone do not permit any definitive statement concerning the date of bullet manufacture.

Finding: Detailed patterns of distribution of ammunition are unknown, and as a result, an expert should not testify as to the probability that a crime scene bullet came from the defendant. Geographic distribution data on bullets and ammunition are needed before such testimony can be given.

Recommendation: The conclusions in laboratory reports should be expanded to include the limitations of compositional analysis of bullet lead evidence. In particular, a further explanatory comment should accompany the laboratory conclusions to portray the limitations of the evidence. Moreover, a section of the laboratory report translating the technical conclusions into language that a jury could understand would greatly facilitate the proper use of this evidence in the criminal justice system. Finally, measurement data (means and standard deviations) for all of the crime scene bullets and those deemed to match should be included.

NRC REPORT, *supra* note 1, at 107–08 (footnotes omitted).

³⁸*See, e.g.*, 1 MODERN SCIENTIFIC EVIDENCE, *supra* note 2, §§ 5:17–5:18 .

application.”³⁹ In addition, it seems clear that the error in the measurements of elemental concentrations is sufficiently small relative to the variations in these concentrations across bullets to permit useful comparisons. The report proposes certain operational improvements and protocols for making the necessary measurements.⁴⁰ To the extent that these measures represent a minimal level

³⁹NRC REPORT, *supra* note 1, at 23.

⁴⁰*Id.* at 23–24:

Recommendation: The FBI Laboratory’s analytical protocol should be revised to contain all details of the inductively coupled plasma-optical emission spectroscopy (ICP-OES) procedure and to provide a better basis for the statistics of bullet comparison. Revisions should include:

1. Determining and documenting the precision and accuracy of the ICP-OES method and the concentration range of all seven elements to which the method is applicable.
2. Adding data on the correlation of older neutron activation analysis and more recent ICP-OES results and any additional data that address the accuracy or precision of the method.
3. Writing and documenting the unwritten standard practice for the order of sample analysis.
4. Modifying and validating the digestion procedure to assure that all of the alloying elements and impurities in all samples (soft lead and hard lead) are dissolved without loss.
5. Using a more formal control-chart system to track trends in the procedure’s variability.
6. Defining a mechanism for validation and documentation of future changes.

Recommendation: The FBI should continue to measure the seven elements As, Sb, Sn, Cu, Bi, Ag, and Cd as stated in the current analytical protocol.

Recommendation: A formal and documented comprehensive proficiency test of each examiner needs to be developed by the FBI. This proficiency testing should ensure the ability of the analyst to distinguish bullet fragments that are compositionally indistinguishable from fragments with similar but analytically distinguishable composition. Testing could be internal or external (for example, conducted by the National Institute of Standards and Technology), and test results should be maintained and provided as appropriate. Proficiency should be tested regularly.

Recommendation: The FBI should publish the details of its CABL procedure and the research and data that support it in a peer-reviewed journal or at a minimum make its analytical protocol available through some other public venue.

Recommendation: Because an important source of measurement variation in quality-assurance environments may be the analyst who makes the actual measurements, measurement repeatability (consistency of measurements made by the same analyst) and

for forensic scientific practices, the use of procedures that are less well tailored to produce accurate results could be problematic under the scientific-validity or general-acceptance standards.

B. *Declaring a Match*

To declare that measurements from one sample is essentially identical to those from another sample, the FBI employed a number of dubious statistical procedures. These involved three seven-element measurements on three fragments from the two samples to be compared.⁴¹ Thus, there were nine seven-element measurements per sample. The FBI had three rules for declaring matches: (1) For each element, use a within-bullet standard deviation (SD) of each bullet, form a two-SD window about each bullet's mean, and declare a match if the two windows overlap (for every element).⁴² (2) Look for an overlap between the two bullets in the range of the replicate measurements on each bullet.⁴³ (3) "Chaining"—a rather peculiar method that seems comparable to saying that if A is a friend of B, and if B is a friend of C, then A is a friend of C.⁴⁴ The NRC Report proposes statistically superior approaches — essentially, computing a distance function for differences in the seven-element measurements or computing a series of *t*-statistics and using one or the other of these quantities to define a match.⁴⁵

reproducibility (consistency of measurements made by different analysts) need to be quantified through Gage R & R studies. Such studies should be conducted for the FBI comparison procedures.

Recommendation: The FBI's documented analytical protocol should be applied to all samples and should be followed by all examiners for every case.

⁴¹*Id.* at 29.

⁴²*Id.* at 28–30.

⁴³*Id.* at 31.

⁴⁴*Id.* at 31–35.

⁴⁵*Id.* at 35–70. Specifically, the committee proffered the following recommendations with respect to declaring matches:

Recommendation: The committee recommends that the FBI estimate within-bullet standard deviations on separate elements and correlations for element pairs, when used for comparisons among bullets, through use of pooling over bullets that have been analyzed with the same ICP-OES measurement technique. The use of pooled within-bullet standard deviations and correlations is strongly preferable to the use of within-bullet standard

Some data indicate, however, that the first of the existing FBI procedures works well enough. Specifically, the FBI laboratory used data from 1,837 bullet-lead samples to determine how well CABL matching works with samples known to come from different manufacturers or production sources.⁴⁶ If matches routinely arise when one sample is compared to another, then the procedure has a high rate of false positives. If the rate is low, then the procedure rarely associates two bullets that have different origins. Since every sample was compared to all the other 1,836 samples for each measured element, a total of 1,686,366 pairwise comparisons were made. For each element that was successfully measured in both samples, the concentration was declared to match if the two-standard-deviation for the means of the triplicate measurements was

deviations that are calculated only from the two bullets being compared. Further, estimated standard deviations should be charted regularly to ensure the stability of the measurement process; only standard deviations within control-chart limits are eligible for use in pooled estimates.

Recommendation: The committee recommends that the FBI use either the T2 test statistic or the successive t-test statistics procedure in place of the 2-SD overlap, range overlap, and chaining procedures. The tests should use pooled standard deviations and correlations, which can be calculated from the relevant bullets that have been analyzed by the FBI Laboratory. Changes in the analytical method (protocol, instrumentation, and technique) will be reflected in the standard deviations and correlations, so it is important to monitor these statistics for trends and, if necessary, to recalculate the pooled statistics.

Recommendation: To confirm the accuracy of the values used to assess the measurement uncertainty (within-bullet standard deviation) in each element, the committee recommends that a detailed statistical investigation using the FBI's historical dataset of over 71,000 bullets be conducted. To confirm the relative accuracy of the committee's recommended approaches to those used by the FBI, the cases that match using the committee's recommended approaches should be compared with those obtained with the FBI approaches, and causes of discrepancies between the two approaches—such as excessively wide intervals from larger-than-expected estimates of the standard deviation, data from specific time periods, or examiners—should be identified. As the FBI adds new bullet data to its 71,000+ data set, it should note matches for future review in the data set, and the statistical procedures used to assess match status.

Recommendation: The FBI's statistical protocol should be properly documented and followed by all examiners in every case.

Recommendation: Interpretation and testimony of examiners should be limited as described above, and assessed regularly.

⁴⁶Robert D. Koons & JoAnn Buscaglia, *Forensic Significance of Bullet Lead Compositions*, 50 J FORENSIC SCI. 1 (2005).

satisfied. The total number of indistinguishable sample pairs was 674, for a relative frequency of 0.04%.⁴⁷ The researchers interpreted such findings as demonstrating that “[c]ompositional comparison of bullet lead provides a reliable, highly significant point of evidentiary comparison of potential sources of crime-related bullets.”⁴⁸

Although the small percentage of false matches in the FBI study is reassuring, the subset of bullet lead samples constructed for the study does not estimate the probability of a false match in practice. The report notes that the subset “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.”⁴⁹ Suppose that the bullets from one large manufacturer are more variable than most of those in the FBI sample (which was intentionally designed to represent many sources). These bullets would be easier to distinguish, and the 0.04% figure would overestimate the chance of a false positive in case work. Conversely, if the bullets from a dominant manufacturer were less variable than those in the FBI sample, then a larger error rate would be expected in practice. In short, the FBI study indicates that CABL can exclude bullets originating from different sources, but it does not supply an error rate that would apply in a particular case.⁵⁰

⁴⁷This false-positive rate includes samples for which not all seven elemental concentrations could be measured. For bullets in which all seven elements were determined, the match frequency was 0.014%. In addition, “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.” *Id.* at 9.

⁴⁸*Id.* at 1.

⁴⁹*Id.* at 3.

⁵⁰Indeed, the authors write that:

We believe that it is neither possible nor appropriate to calculate reliable probabilities of chance occurrence of indistinguishable items of nonbiological 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

Moreover, the premise that two samples either do or do not match is itself an oversimplification.⁵¹ The elemental concentrations are continuous quantities, and the degree to which they match affects the probative value of the evidence.⁵² Further research on ways to express the implications of various degrees of similarities in the measurements from the two samples might suggest a more refined approach than the traditional effort to force the results into two somewhat arbitrary categories (match or nonmatch).⁵³ This problem of interpretation is the subject of the remaining section.

C. *Inferring an Association from a Match*⁵⁴

mean that the significance of two samples having indistinguishable properties is low, only that it cannot be calculated precisely.

Id. at 10.

⁵¹See, in this context, D.H. Kaye, *The NRC Bullet-Lead Report: Should Science Committees Make Legal Findings?*, 46 JURIMETRICS J. 91 (2005). The exposition of this point in note 26 of this article is marred by the mistaken claim that the distance statistic described there is normally distributed with the specified mean and standard deviation. I am grateful to Joseph Gastwirth for calling this to my attention, but I do not believe that a more accurate formulation would lead to different conclusion.

⁵²This point was well ventilated in connection with the presentation of similarities and differences in the molecular weight measurements obtained from gel electrophoresis of DNA fragments in forensic settings. See authorities cited, *id.* at 96 n.25.

⁵³The prospects for satisfactory modeling may be dim. FBI scientists have reported that:

[I]n 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.

Koons & Buscaglia, *supra* note 46, at 2. The article offers no support for its claim that “most U.S. courts” have rejected “approaches based on likelihood statistics,” and we know of no such reported opinions.

⁵⁴Parts of this section are taken, without further attribution, from Kaye, *supra* note 51.

If measurements are reasonably precise and a matching rule is generally accurate in classifying samples from the same source as matches and those from different sources as nonmatches, then finding a match indicates that the two items originated from the same source. As previously noted, however, the meaning of “source” in this context is unclear, and the question of how the laboratory findings of similarity in elemental concentrations should be explained to the jury is not a strictly scientific question.

The National Research Council report argues that analytical chemists or technicians can provide relevant evidence by testifying that two bullets originate from the same “compositionally indistinguishable volume of lead,” or “CIVL.”⁵⁵ The committee offers the concept of a CIVL as a more precise term than “source,” which might be a billet (made by remelting large blocks of smelted lead), a wire (extruded from a billet or mix of billets), or something else.⁵⁶ Yet, inasmuch as the boundaries of a CIVL never are known or measured, this concept is hardly free from ambiguity. A CIVL can range from “[a]t the very least, . . . several wires”⁵⁷ to an entire “vat of molten lead whose composition is not altered during the pouring of bullets.”⁵⁸ By definition, a CIVL is “produced during one production run at one point in time [and] is at least as large as the sample taken for analysis.”⁵⁹ Citing a “review of the literature and discussions with manufacturers,” the committee found “that the size of a CIVL ranges from 70 lbs in a billet to 200,000 lbs in a melt. That is equivalent to 12,000 to 35 million 40-grain, .22 caliber longrifle bullets from a CIVL compared with a total of 9 billion bullets produced each year.”⁶⁰

The issue that matters to the jury, however, is not whether two bullets came from the same CIVL, but whether they both came from ammunition in the defendant’s possession. If the elemental measurements do not affect the probability that the crime-scene bullets came from the defendant, then the

⁵⁵NRC REPORT, *supra* note 1, at 85 (defining the term).

⁵⁶*Id.* at 82–84 (decrying “[t]he poor definition and understanding of the term ‘source’” and defining these terms).

⁵⁷*Id.* at 85.

⁵⁸*Id.*

⁵⁹*Id.* (emphasis omitted).

⁶⁰*Id.* at 107; *see also id.* at 167. Evidently, the committee believed that testifying experts should present similarly broad ranges for the meaning of “source.” It recommended that “[e]xpert witnesses should define the range of ‘compositionally indistinguishable volumes of lead’ (CIVL) that could make up the source of analytically indistinguishable bullets, because of variability in the bullet manufacturing process.” *Id.* at 106.

laboratory findings are irrelevant—regardless of what one concludes about the “CIVLity” of the bullets being compared.⁶¹ Nevertheless, the NRC report focuses almost entirely on “CIVLity.” It asserts that if two bullets that have the approximately same elemental concentrations, then they are more likely to be from the same CIVL than from different CIVLs.⁶²

At first glance, this claim seems too obvious to question. By definition, a CIVL is so homogeneous that differences within it are undetectable. But how much overlap is there among different CIVLs? The report states that “[t]he frequency of coincidentally identical CIVLs is unknown.”⁶³ Apparently, the production process could be characterized by a few large, “coincidentally identical CIVLs” along with thousands of smaller, highly variable ones.

Even if it were clear that small differences in the elemental concentrations of two bullets implied a common CIVL origin, would a small value for these differences then warrant an inference to the conclusion that the two bullets once were in the defendant’s possession? Is it safe to move from “same CIVL” to “same box” or even “same or nearby boxes”? The answer depends in part on how bullets are packed and shipped — on how they move from the production line, to the wholesaler, to the retailer, to the customer. The mixing of bullets from different CIVLs can lead to bullets from many CIVLs being in many boxes, frustrating any assumption that two bullets are more likely to end up in the same box if they come from the same CIVL than if they come from different CIVLs.

Even if a match in the concentrations is highly probative of CIVLity, it may not be probative as to the proposition that is of interest to the jury — whether the two bullets were near neighbors at the end of the production and distribution chain. Research into the likelihood that bullets from one CIVL segregate in the same boxes would seem necessary to establish that the bullets came from the box or boxes of ammunition obtained by the defendant. Suppose, for instance, that the crime-scene bullet is definitely from a given CIVL but that a great many boxes (presumably possessed by a great many gun owners) also contain such bullets. Then it is not clear that CIVLity makes it even slightly more probable that the defendant’s collection of ammunition is the source of the crime-scene bullet. This possibility is consistent with the NRC report’s observations that “[i]t is routinely found that a single box contains multiple

⁶¹See William C. Thompson, *Analyzing the Relevance and Admissibility of Bullet-lead Evidence: Did the NRC Report Miss the Target?*, 46 JURIMETRICS J. 65 (2005).

⁶²NRC REPORT, *supra* note 1, at 97–99.

⁶³*Id.* at 99.

distinct compositional groupings—as many as 14,⁶⁴ and that “distribution can lead to varied scenarios regarding retail dispersion of bullets from a distinct compositional group.”⁶⁵ As such, “[a] conclusion that two bullets came from the same melt does not justify an expert in further testifying that this fact increases the odds that the crime bullet came from the defendant.”⁶⁶

In short, the committee finesses the scientific validity or general acceptance problem by having the expert truncate his testimony at the CIVL stage. Under its recommendations, the analyst may testify that the laboratory tests make it more probable that certain bullets originated from the same CIVL, but the expert cannot testify to what follows from this fact.⁶⁷

Whether a solid base of scientific studies is needed to admit this constrained testimony depends on whether a reasonable judge or juror can conclude, without the benefit of expert testimony, that CIVLity actually shifts the probability of the more interesting proposition that the defendant had access to the crime-scene bullet. This is really a question of relevance, or “fit” in the language of *Daubert*.⁶⁸ The fact of CIVLity is relevant if, in general, when bullets come from the same CIVL, it is more likely that they will wind up in the possession of the defendant. If this is something that is reasonable to assume (without empirical research), then the testimony is at least logically relevant. And, in that event, admissibility turns on the balance between probative value and the usual counterweights of unfair prejudice, time-consumption, and confusion.⁶⁹ The danger is that the jury will overweight the evidence, thinking that the link between CIVLity and the defendant is stronger than it is. The matter can be put into perspective with expert testimony from the defense on how little is known about this link, but whether the game is worth the candle is open to question. When all is said and done, the jury will have heard somewhat intimidating and complex testimony about spectroscopic measurements and the process by which bullets are manufactured and

⁶⁴*Id.* at 84.

⁶⁵*Id.*

⁶⁶*Id.* at 102.

⁶⁷The report recognizes that “[t]he large number of bullets made from a single melt and the absence of information on the geographic distribution of such bullets precludes such testimony *as a matter of expertise*,” but it leaves open the possibility that “[s]uch an inference is a matter for the jury.” *Id.* (footnotes omitted).

⁶⁸See KAYE ET AL., *supra* note 11, § __.

⁶⁹See, e.g., FED. R. EVID. 403; *cf.* FED. R. EVID. 702 (requiring expert evidence to “assist the jury”).

distributed. The match, if established, will incriminate the defendant—but to an unknown degree. Until there is reason to believe that CIVLity is strongly indicative of a defendant’s association with both bullet-lead samples, the wiser course may be to exclude even the NRC report’s sanitized version of CABL testimony.