Comparative Bullet Lead Analysis: A Case Study in Flawed Forensics

If this technique had not been used for years to send people to prison, no reasonable scholars of forensic evidence would consider it ready for court.¹

— William Thompson, Professor of Law and Criminology, UC Irvine

Expert testimony about bullet lead comparisons has been admitted in criminal trials without significant challenge for close to 40 years. Yet recent studies have revealed serious flaws in the underlying theory and assumptions, forensic practice, and conclusions of bullet lead experts. Emerging evidence from the studies confirm, as Professor Thompson has observed, that the practice of comparative bullet lead analysis (CBLA) is not ready for judicial prime time.

CBLA is now under attack. The criticisms are based on reviews of FBI Laboratory analytical and forensic practices, studies of metallurgical phenomena and data associated with lead smelting and bullet manufacturing industry practices, scientific reports and publications, and over 50 trial transcripts. They reveal that the foundational premises of CBLA evidence have been based on flawed logic and unsupported subjective belief for decades. CBLA evidence has survived in the courtroom for decades because the defense bar was at a hopeless disadvantage as to specialized analytical equipment (a nuclear reactor was required for the first several decades) and availability of expertise (the FBI Laboratory is the only laboratory in the nation offering CBLA).

Accordingly, the practice was never effectively challenged. Because of the absence of scientific and legal oversight, scientific reliability of the underlying theory, hypotheses, assumptions and forensic practice have never been proven. Several independent recent studies now cast serious doubt as to the validity of CBLA practice as it has been proffered in courts for almost 40 years.

Since the FBI is the only forensic laboratory in the United States offering the comparison of bullet lead compositions as evidence of guilt, they commissioned a study by the National Research Council of the National Academies (NRCNA herein), as a consequence of challenges to CBLA by the author and colleagues. The NRCNA released its report on February 10, 2004.

By William A. Tobin
The report did not directly address the scientific validity of the underlying theory or major premises and assumptions of CBLA, in part because there have never been any comprehensive and meaningful studies of the hypotheses and assumptions of CBLA (and thus, there exists no body of data for the committee to study), and in part because the Committee (on Scientific Assessment of Bullet Lead Elemental Composition Comparison at the NRCNA) was constrained by charter to the Statement of Tasks furnished by the entity commissioning the study (the FBI Laboratory). However, even though the Committee was inappropriately constituted for the metallurgy-based challenge, it confirmed major flaws in the remaining aspects of the forensic practice, to include the basis by which a compositional “match” between and among bullets is claimed. Accordingly, the report severely restricts experts’ allowable conclusions in the courtroom for all future testimonies relating to compositional comparisons of bullets.

This article is not intended as a stand-alone explanation of the actual forensic practice of comparing bullet lead compositions nor of the various processes of lead smelting or bullet manufacturing. Information on the processes can be obtained by review of several of the references noted in this article or by direct contact with the author.

Bullet-Making 101

Very basically, bullet lead originates as a large “pot” of molten lead refined from recycled automotive batteries. At the secondary refinery converting the used batteries, only a small fraction of the recovered and refined lead is sent to bullet manufacturers; the overwhelming majority of it is used for the manufacture of new automotive batteries. The lead intended for the new automotive batteries has very stringent compositional specifications. In large part because the fraction of refined lead intended for bullets is so small, the secondary refiner does not “un-engineer” the tight battery lead composition specifications for the 5 percent, more or less, of the lead going for bullets.

The refined lead for bullets is then shipped by the secondary refiner to bullet manufacturers in the form of ingots or billets (not bullets) to eventually be extruded or cast into bullets. The bullets are then packaged in boxes stamped with packing codes (sometimes called “lot codes”) and generally shipped to jobbers and wholesalers. For .22 caliber bullets, the caliber that drives the bullet lead industry 24/7, Wal-Mart and K-Mart purchase and market more than half the total produced, a significant factor for FRE 403 considerations. (And similar state rules of evidence). These considerations will be discussed later.

Comparative Bullet Lead Analysis (CBLA) 101

When forensic bullet and/or bullet fragment specimens from crime scenes and suspects are analyzed and compared as to elemental composition, 6 or 7 specific elements in the lead matrix (but not the lead) are compared. When each of the 6 or 7 analytes of a “questioned” bullet is within a certain statistical range of the same (corresponding) element in a “known” bullet, the CBLA expert witness has historically testified that the bullets must, therefore, have a common source as to original pot of molten lead and, therefore, must have been produced by the same manufacturer on the same day. Many times the witnesses have even concluded that the bullets came from the same box of bullets.

Phases of Forensic Practice

It is easiest to understand the forensic practice of comparing bullet lead compositions by considering it as a three-phase process:

Analytical phase: In the first phase, the bullet samples are analyzed for elemental composition (bullets are not pure lead) with high-technology analytical instrumentation. The instrumentation for the technique of choice used for the first (approximately) 25 years, neutron activation analysis or NAA, was a nuclear reactor. Since approximately 1995, the technique of choice became inductively coupled plasma — atomic (or optical) emission spectrometry (abbreviated ICP), which no longer required access to a nuclear reactor.

Grouping phase: In the second phase of the procedure, the elemental composition numbers generated during the first (analytical) phase are “grouped” according to similarity of compositional presence (amount). Compositions similar to a crime scene bullet(s) are put in one group and considered “analytically indistinguishable”; compositions considered dissimilar are placed in different groups and considered “analytically distinguishable.”

Inference phase: In the third phase, the expert witness draws a conclusion as to alleged probative significance of finding “analytically indistinguishable” (similar) compositions in both crime scene and “known” bullet samples.

In our challenge to CBLA practice, we generally assume the competent conduct of the first phase of the examinations (analytical chemistry). With regard to the second phase (where bullet compositions are collated into groups by compositional similarity, groups within which bullets are statistically considered “analytically indistinguishable”), although we found the process subjective in practice and without written protocol until only relatively recently, we have not generally challenged the Phase II activities except in case-specific instances. Those case-specific instances include where examiner bias was evident or when a process of statistical data treatment known as “chaining” was used to claim incriminating associations. Inasmuch as the recent NRCNA report now condemns the practice of “chaining” for bullet lead comparisons, it is doubtful that data chaining will be used for future cases and will not be discussed here. However, cases in the appellate or post-conviction relief stages should be reviewed for possible use of data-chaining.

It is the third phase of the practice, the inference phase, that is most objectionable and, unfortunately, the most directly incriminating. In this phase, the expert witness has most frequently concluded that, because the crime scene and known bullets are allegedly “analytically indistinguishable” (compositonally similar), the bullets have a common origin as to “molten source.” In the courtroom, the expert witness states that the compared bullets “were made by the same manufacturer, on the same day, from the same molten source of lead,” and often that the bullets “originated from the same box of bullets.” In other words, because specimens are claimed to be similar in composition, they, therefore, must have come from the same molten source. The witnesses have asserted as universal assumptions that all bullets from the same molten source have the same composition, and that all bullets from different molten sources have different compositions.

The assumptions that are necessary to support such conclusions and the assertions posited by witnesses as universal assumptions are logical and readily understood. The very first assumption required is that the tiny fragment under analysis accurately represents the composition of the source from which it originated. This assumption is, therefore, that the forensic specimen is a representative sample.

The second assumption is that the source from which the fragment or sample originated is compositionally uniform or homogeneous; in other words, that the source is the same in composition from top to bottom, left to right, inside center
to external surface. This assumption is one of homogeneity and has been expressly stated by CBLA witnesses in the majority of the testimonies reviewed.

The third assumption is that no two molten sources are ever produced with the same composition. This assumption has also been expressly stated in many testimonies reviewed and in others indirectly. In Government’s Memorandum in Opposition to Defendant’s Motion In Limine on Proposed Bullet Lead Expert Opposition to Defendant’s Motion Under Rule 702 of The Federal Rules of Evidence at 17 in the high-visibility matter of U.S. v. Mikos, by the government’s own admission:

This premise is an important one for comparative bullet lead analysis because if most or all sources of lead had the same elemental composition, then a match between bullets would have little significance.

The “same composition = same batch of molten lead” theory of CBLA underlying the conclusions typically rendered in criminal trials requires that all three assumptions be valid.

In an elegantly simple expression of summary logic, Judge Ronald A. Guzman stated in Mikos:

This is only true if all bullets which come from the same batch have the same composition and if bullets from other batches do not.

Glaring Flaws in Logic and Foundational Validity

The major flaw of CBLA is one of unjustifiable extrapolation from the generally accepted portion of the practice (the analytical phase) to objectionable inference, and is readily understandable by analogy to blood testing. Let’s assume for a moment that samples of blood are extracted from both the reader and the author. The samples are analyzed and compared with regard to the analytes iron, calcium, potassium, sodium, HDL and LDL. The quantitative presence (amount) of each analyte is found to be “similar” (within the precision of the analytical equipment used) between the reader’s and author’s blood samples such that the samples are considered “analytically indistinguishable.”

The analytical (first) phase of this process is “generally accepted” in the medical and scientific communities and would not be challenged (in our comparison to CBLA). However, a conclusion that, because the blood samples are of similar composition, therefore the reader and author have similar origins as to source (parents) is clearly an unjustifiable extrapolation. The generally accepted practice of analyzing blood constituents has been contorted into scientifically unreliable and unfounded inference. That is exactly what has occurred in courtrooms for almost four decades with regard to compositional comparisons of bullet lead. When challenged as to the reliability of CBLA, practice advocates have typically addressed the general acceptance of the instrumentation and analytical procedure rather than scientific acceptability of the inference(s) drawn, and courts have admitted the practice as generally accepted in the scientific community, eventually allowing it to become a venerable forensic practice.

Longevity of Frye and Daubert Admissibility

There are a confluence of circumstances that have facilitated longevity of CBLA admissibility under Frye,4 Daubert5 and other evidence admissibility criteria, which have entrenched the practice in judicial admissibility. To increase chances for successful challenge of the practice under Daubert, it is beneficial to understand these circumstances with respect to each of the Daubert factors expressed in the stare decisis, advisory committee notes and legal treaties to date.

If a judge were to apply each of the factors enumerated by Justice Blackmun in the Daubert decision to CBLA in order to decide whether the majority of factors cut in favor of or against admitting the testimony, the judge might well conclude CBLA testimony is inadmissible. Although the underlying theory and premise of CBLA, sample representativeness, compositional uniformity of sources, and compositional uniqueness of sources, are testable, there has been no comprehensive or meaningful testing to validate them. Worse, recent testing yields results at odds with the premises, and there is no meaningful peer reviewed and refereed literature supporting CBLA theory. Further, “general acceptance” is limited to a very small group (2 to 4 at present) of forensic analytical chemists employed by the same institution.

Ironically, and unknown to the judicial community, there did not even exist general acceptance within the very small community of bullet lead experts. One of the early researchers himself, Dr. Vincent P. Guinn, objected to the manner in which the FBI Laboratory was unjustifiably extrapolating with invalid inference the practice he pioneered.

The challenge to CBLA raised by the author and colleagues was vindicated in almost every aspect by the recently released NRCNA report of the National Academies of Science (NAS). The report culminated an approximate 12-month study by the Committee on Scientific Assessment of Bullet Lead Elemental Composition Comparison (“the Committee” herein). As observed by John Thornton, Professor Emeritus of Forensic Science at UC Berkeley, the NAS report “…diminished the probative value of the technique in a substantial way.” Professor Thornton described the bullet manufacturing variations cited by the report “a real harpoon in the validity of the method.”

“If you were told that the perpetrator had brown hair, would that be relevant? Yes, but it doesn’t get you very far,” observed Professor David L. Faigman, UC Hastings College of Law in San Francisco. “It doesn’t mean that it ought to be admitted, because it may not have enough relevance to offset the possibility that it might confuse a jury and waste a court’s time.”

However damaging the NAS report is to the status quo practice of comparative bullet lead analysis, there remain several findings and observations of the report that are too limited in scope and/or subject to misinterpretation, misrepresentation or outright spin. Notwithstanding the claim by the FBI following release of the NAS report that the practice “…will still be useful in linking individuals to a crime as well as to exclude others”, it will probably become apparent to the reader that forensic worth as evidence of guilt is questionable at the present time. Significant probative value is likely many years away, if it even exists, and only after comprehensive and meaningful research and study.

Origin and Evolution

During the 1960s, when expanded applications for nuclear technology were being explored, researchers at Gulf General Atomic evaluated the possibility of analyzing the compositional (constituent) elements in bullet lead for forensic use and believed that it had promise. The FBI Laboratory adopted the technology and began offering it as a forensic service.

For the first approximately 25 years of the practice, the FBI analyzed only three elements in the lead matrix, antimony, arsenic and copper. In their 1970 report, however, Gulf General Atomic reported that analyzing only those elements in the lead was inadequate to uniquely characterize a source of lead. However, in the many hundreds of courtroom testimonies that followed, transcript reviews reveal no indication that courts were ever apprised of
NOTICE OF AUCTION

Buy a Piece of Cyberspace and Help NACDL’s Foundation — Domain Names Available —

The Foundation for Criminal Justice — NACDL’s charitable supporting organization — is auctioning four domain names in support of NACDL. This is your chance to acquire an eye-catching Web address, and to support the charitable purposes of the Foundation.

www.defenseattorneys.com
minimum bid: $1,500

www.legalcenter.com
minimum bid: $1,000

www.legaldefense.com
minimum bid: $1,500

www.constitutionallaw.com
minimum bid: $1,000

Initial bids will be accepted until Friday, July 23, 2004, either by letter or via e-mail. Additional verbal bids will be accepted on Friday, July 30, 2004, at NACDL’s Annual Meeting, in San Francisco, during that evening’s “Casino Cruise” on San Francisco Bay.

The highest bidder wins the auction, and the donations are tax deductible to the extent provided by law.

Mail all bids to: Foundation for Criminal Justice, c/o Domain Name Auction, 1150 18th Street, NW, Suite 950, Washington, DC 20036; or via e-mail to Ralph Grunewald, Foundation Manager, at ralph@nacdl.org

THE FOUNDATION FOR CRIMINAL JUSTICE
THE SUPPORTING FOUNDATION OF THE NATIONAL ASSOCIATION OF CRIMINAL DEFENSE LAWYERS

“ENSURE JUSTICE AND DUE PROCESS”
the limitation expressly indicated by the pioneering researchers. Accordingly, courts began to admit the then-novel theory and testing under Frye because of the apparent acceptance and forensic value of compositional comparisons, and of the widespread acceptance of the nuclear technology underlying the instrumentation used for the forensic analysis, some of the same technology that was responsible for providing electricity to millions of people and companies throughout the world for many years. But general acceptance in the scientific community of the instrumentation and analytical technique is one thing; concluding general acceptance as to the inferences drawn from the comparative analyses is quite another, as can be seen from our blood-testing analogy.

For these first 30 years of forensic practice, there were no significant public challenges to CBLA practice. There is no commercial or vital interest in comparing bullet leads outside the forensic community. But even within the forensic community, no forensic laboratory other than that of the FBI offered comparative bullet lead analysis as a forensic service. Worse, and probably the most formidable obstacle to scientific evaluation and peer review for the first 25 years or so, was the requirement of access to a nuclear reactor required for analysis by neutron activation.

These conditions posed realistically insurmountable obstacles to thorough scientific evaluation of the practice. In essence, the only “scientific community” with professional interest and appropriate instrumentation to conduct evaluations for purposes of hypothesis testing, theory validation and reliability were the handful of practitioners at the FBI Laboratory. Courts admitted and judicially perpetuated the practice under Frye more from a dearth of challenges than endorsement of practice theory. After so many years of judicial receptivity, the “theory” and testing were no longer “novel”, making laissez-faire judging the path of least resistance and challenges unbelievable.

In 1993, U.S. Supreme Court’s ruling in Daubert v. Merrell Dow Pharmaceuticals, Inc. became the evidentiary validating criterion for all federal cases. Because the Daubert decision was grounded on statutory construction rather than constitutional analysis, state courts remained free to interpret their validating statutes differently, and eighteen states have opted to continue to adhere to Frye. Notably, those states continuing to adhere to Frye include the most populous and litigious states of California, Florida, Illinois, New York, Pennsylvania and Washington. Accordingly, Frye still governs at most state trials.

### Foundational Adequacy Under Daubert

It is understood that evaluation of the adequacy of expert testimony validation must be relative in at least two respects: to the specific theory on which the expert proposes relying, and to the degree of definiteness of the opinion to which the expert contends testimony. Professor D. Michael Risinger, in one of the most influential and insightful contributions to the Daubert literature, argues that Daubert requires the trial judge to ask whether the available research data validates the specific theory on which the expert intends to rely. In Daubert, Justice Blackmun instructed judges to determine whether the foundation establishes that the expert can reliably perform the particular “task at hand.” The other two cases in the Supreme Court’s expert testimony trilogy, Joiner and Kumho, also contain language directing the judge to conduct a narrowly focused admisibility analysis. The question is not the “global” reliability of the expert’s field or discipline. In the context of CBLA, the question is not the validity of analytical chemistry or even of the use of the analytical technique (ICP) to assess elemental composition. The “task at hand” is deciding whether the crime scene bullet fragment and the bullets from a subject’s actual or constructive possession originate from the same manufacturer, molten source, batch or box. The implied theory is that the expert has developed a technique for enabling an accurate conclusion associating a fragment with its source by composition.

Before concluding that two bullets originate in the same source (the ultimate inference in the “task at hand”), the expert must make three assumptions: (1) that the small fragment or bullet accurately represents its source as to composition, (2) that each molten source is compositionally uniform (homogeneous) and, (3) that the composition of each molten source is unique or individual. Each assumption is an essential step in the chain of reasoning leading to the ultimate inference. If any one of the three assumptions is false, the chain fails and the inference as CBLA experts have presented it may not be drawn. Empirical studies of the author and colleagues, to include several bullet lead refiners, find that not just one of the links in the chain of assumptions are invalid, but that all three are invalid as universal assumptions.

From a scientific standpoint, the practice of CBLA fails all but one of the Daubert criteria enumerated in the legal literature and judicial rulings.

The first major flaw of theory reliability is that there has never been any meaningful and comprehensive research to validate the theory or three required and declared (by CBLA witnesses) underlying assumptions. Additionally, no blind or double-blind studies have ever been conducted. In reality, the assumptions have been based on nothing more than subjective belief and unsupported speculation. To practice advocates (who are not metallurgists, the appropriate discipline for the inference phase, but rather are analytical chemists, the appropriate discipline for the analytical phase), it seems that because the precision of their analytical equipment allows them to distinguish “molten source” constituents to the parts per million level and beyond, that each source must be distinguishable because no two sources could ever be produced with similar compositions, and that each source must be compositionally homogeneous. In short, there is simply no meaningful and comprehensive body of data that supports the [required] underlying assumptions.

Although the challenge to comparative bullet lead practice is only several years old, courts are beginning to recognize reliability deficiencies of the inference phase of the practice in their gate-keeping role under Daubert and are restricting testimonies to the observation of composition similarities only, consistent with witness credentials. In State of New Mexico v. Trujillo, the state court admitted testimony as to an alleged bullet lead composition “match”, but did not allow testimony regarding opinion as to forensic significance. And in the recent, high-visibility federal matter of U.S. v. Mikos, the Court specifically refers to the total void in supporting studies for the underlying premises of the practice.

We understand that the FBI Laboratory has performed comparative bullet lead analysis (CBLA) for many years. Furthermore, we understand that persons from the FBI Laboratory...have for years been allowed to testify at trials as to their opinions regarding the source of tested bullets based on CBLA. In our opinion, however, 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 cor-
Testability and Error Rate

Some legal scholars have referred to testability as probably the most important of the Daubert validation criteria. The underlying theory and assumptions of CBLA are certainly testable, as evidenced in recent research efforts by Koons and Grant (FBI researchers) and also by two earlier cursory projects of FBI examiners. However, in what should have been a red flag for CBLA practice advocates, Gulf Atomic and FBI researchers coincidentally encountered bullets likely from unrelated sources of similar composition in every study conducted.

In the most meaningful effort (Koons and Grant, 2002), an error rate of 25-33 percent was observed. The researchers rationalized that the indistinguishable compositions inadvertently encountered (which would have been “false positives” in a forensic setting) were of different calibers and would, therefore, not pose a problem for probative utility.28 However, the author has observed a bullet manufacturer’s operations where every bullet of every caliber produced by the company is made from the same pallet of ingots. Additionally, a Remington product specialist testified in at least one trial that Remington makes up to 15 different caliber bullets from the same source of lead.29 Significantly, one reason the forensic practice of comparing bullet lead compositions is generally requested and conducted is that, frequently, there is insufficient amount of the original projectile remaining in the victim’s body for the much-preferred conventional “ballistics” examinations (of matching gun barrel striations). Such circumstances dramatically increase the proportion of cases where only fragments of the original projectile are analyzed by CBLA. Thus, discrimination by caliber is most often not feasible in cases where bullet lead composition comparisons are requested.

Actual error rate of the practice has never been studied or evaluated, a fact confirmed by the recent NRCNA report. However, it is quite statistically and forensically significant that in each of the efforts by FBI researchers to study source homogeneity, the researchers coincidentally encountered unrelated bullets that would be considered “analytically indistinguishable” (compositionally similar), and thus ‘false positives’ in a criminal trial. Those coincidental encounters are quite likely indicative of an unacceptably high rate of error, and the encounters were not even confined to similar geographic regions, where one should expect very high “false positive” rates of error due to regional concentrations of similar composition bullets attributable to distribution practices. This consideration will be discussed later in this article relating to whether CBLA evidence is more probative than prejudicial (i.e., whether it passes muster under FRE 403 and similar state rules).

Although the recent report of the NRCNA asserts that the practice “has been tested”, it ignored the unexpected results of the only three “tests” conducted (two of which falsify the hypotheses of CBLA), and skirted the fact that the samples comprising the “testing” were woefully inadequate, statistically equivalent to several grains of sand from one of the world’s beaches. However, even had statistically appropriate sample sizes been involved and the test results favorable to CBLA theory, the finding of no coincidental encounters of similar compositions would not be dispositive of theory and/or assumption validity. As observed by Edward J. Imwinkelried, Professor of Law, University of California-Davis, attempts to disprove the hypothesis are more significant [than verification] in two respects. First, although a single outcome consistent with an hypothesis furnishes little proof of the truth of the hypothesis, a hypothesis phrased as a universal statement is disproved by even one singular inconsistent outcome. Second, even when there are an impressive number of consistent outcomes and no inconsistent outcomes, the hypothesis is not definitively confirmed because it is always possible that an empirical test will some day demonstrate the theory to be incorrect. The theoretical possibility of disproof remains.30

General Acceptance, Peer Review and Extrajudicial Use

The lynchpin criterion for admissibility under Frye was general acceptance in the community to which a forensic practice belongs. Although the Daubert decision invokes other factors in judicial gate-keeping assessments for evidentiary admissibility, general acceptance remains one of the considerations.

Since the FBI Laboratory is the only forensic laboratory in the United States that offers CBLA as evidence of guilt in criminal trials, the scientific community of practitioners is very small, today only two examiners. But even if the scientific community is defined as all forensic laboratories in the U.S., there are indications that the theory is not accepted even in forensic circles (as will be discussed shortly).

In the early CBLA cases, for approximately the first 25 years when prosecution experts relied on neutron activation analysis (NAA) for the compositional analyses, few, if any, American forensic laboratories had access to a nuclear reactor needed to conduct NAA tests. Thus, other forensic laboratories and independent scientists were not in a position to duplicate FBI CBLA tests to determine whether the underlying theory is scientifically valid, whether the three-phase technique is reliable, or to independently assess the validity of NAA CBLA tests and analyses to decide whether the procedure has an acceptable rate of error. In reality, there existed no effective peer review of the practice other than by the few practitioners reviewing each other’s work.

Another indicator of lack of acceptance even in the forensic community is that the instrumentation used for today’s CBLA practice (ICP) is ubiquitous in scientific and forensic laboratories throughout the United States according to CBLA
expert witnesses. There are many state crime laboratories with ICP capability, and the FBI Laboratory is not the only federal forensic laboratory in the United States. Another federal forensic facility is the National Laboratory Center of the Bureau of Alcohol, Tobacco, and Firearms (ATF) in Rockville, Maryland. At one time, the ATF laboratory explored the possibility of employing CBLA as a service for its investigative agents. Ultimately, ATF researcher(s) concluded that the technique is unreliable and discontinued research into the subject. In Europe, the Bundeskriminalamt (BKA, the German equivalent to the FBI) studied the practice for over 30 years and rejected its use in criminal trials as evidence of guilt, and even significantly curtails its use as an investigative tool, because of the unreliability of the practice. 

Additionally, it is doubtful that it can be said that the technique is generally accepted by scientists at non-forensic laboratories. Analytical chemists, or even other scientific professionals, outside the FBI Laboratory lack any practical incentive or motivation to learn, much less conduct the necessary research to validate, CBLA theory:

[W]hen new techniques or theories are espoused in the . . . field of corrosion, . . . there is widespread attention due to the financial interests of billion dollar industries such as oil and gas. [In contrast], [n]o such financial or even professional incentive exists for detailed scientific review of CBLA practice, outside the FBI Laboratory. 

Even though instrumentation has no longer been a barrier to entry since approximately 1990, no other laboratory other than the FBI Laboratory is known to offer CBLA as a forensic service even today. And while a very limited number of analytical chemists (and only at the FBI Laboratory) appear to accept the theory underlying CBLA, there is little or no evidence of acceptance in the two other highly relevant fields of metallurgy/material science and statistics. It is a mistake to inquire into a technique’s acceptance by surveying only experts who use the technique. Narrowly confining the inquiry to the experts employing the technique virtually ensures a finding of acceptance and admissibility. If the general acceptance standard is to have any teeth at all, the only sensible way to apply it is to expand the inquiry to canvass the sentiment in any group of experts whose education and training equip them to assess the validity of the theory.

Unjustifiable Extrapolation
In the G.E. v. Joiner part of the “Daubert trilogy”, the Joiner Court indicated that judges should bar expert testimony as unjustified extrapolation when they believe that there is “too great an analytical gap” between the available data and the expert’s inference. As mentioned in an earlier section, there are virtually no data of a scientifically meaningful nature supportive of CBLA assumptions and theory. In the Kumho portion of the Daubert trilogy, the Court stressed that before permitting an expert to draw an inference on the witness stand, the judge must ensure that the inference has a firmer basis than the expert’s mere “ipse dixit.” As we observed in Comparative Bullet Lead Analysis (CBLA) Evidence: Valid Inference or I�e Dixit?

Like the validity of fingerprint analysis, ultimately the validity of CBLA turns on the rationality of the inference drawn after the third, evaluative stage. Indeed, to a greater extent than fingerprint analysis, CBLA testimony poses the risk that the judge and the trier of fact will lose sight of the vital importance of that stage. In the two preliminary stages of CBLA, the expert employs instrumentation, either NAA or ICP. Those instruments are not only universally accepted; more to the point, they make exact measurements of the quantities of the chemical elements present. However, the very exactitude of those measurements can be beguiling. It can distract the judge and trier of fact and mislead them into thinking that the precision of the measurements in the earlier analyticic stages somehow guarantees the validity of the ultimate inference drawn in the later, evaluative stage.

There is no such guarantee. As this article has hopefully demonstrated, even if one posits the absolute accuracy of the measurements made by the NAA or ICP instrumentation, it is fallacious to leap to the conclusion that the CBLA analyst’s ultimate inference is valid. To draw that inference, the analyst makes the further assumptions of the representativeness of samples as well as the uniformity and uniqueness of molten sources. The uniformity and uniqueness assumptions are the Achilles’ heels of CBLA, since
the available research data calls into question the validity of those assumptions.

There is a lesson to be learned from the controversy over CB LA testimony: In trace evidence analysis, the courts must critically focus on the validity of the expert’s final inference. The essence of trace evidence analysis is not mechanical measurement or simple observation. Rather, the heart of the matter is the much subtler task of deciding which inferences may justifiably be drawn from the observations and measurements. The harsh reality is that even after making an exquisite measurement, a trace evidence analyst may draw an erroneous inference. Daubert condemns unsupported inferences to inadmissibility. 37

**Risk of Prejudice**

Although the NAS report suggests that the forensic practice of comparing bullet compositions may have probative value in some cases, assessment of each of the various Daubert validating criteria should raise questions as to admissibility of the practice in many, if not most, cases. But even if considered admissible under Daubert, additional concerns arise when weighing actual probabilistic evidence of identity that bullets “match” in composition against the risk of prejudice, particularly when bullet distributions are unknown and studies have never been conducted and published. As one judge observed,

> If the jury doesn’t know how many bullets are out there like this and how big that melt was, how can they possibly consider this testimony and give it accurate probative value? They are going to have to speculate as to how many people out there have got these bullets. I mean, it may be that everybody in Columbia [South Carolina] has got them. 38

Considerations affecting retail bullet distribution are myriad and apparently had never been researched for effect on probabilistic value of bullet composition “matches” until the author commenced such studies in early 2003. University of California researchers joined the study in late 2003, studies still underway as of this writing. Early results are even more surprising than expected and underscore concerns as to the value of bullet lead composition testimonies as evidence of guilt.

Particularly questionable, viewed in the light of the author’s bullet distribution study results, are attempts by proponents of the practice and the National Research Council to address “false positive probability” matters based on a nationwide collection of bullets. The author’s [yet unpublished] study in Juneau, Alaska, revealed that innocent customers had no choice but to purchase bullets of the same composition in many of the bullet brands and lines. And the author’s study in Fredericksburg, Virginia, revealed that similar composition bullets remain on the shelves for periods of many months, 39 at a minimum. Accordingly, a “false positive probability” based on a national collection of bullets that is offered to support or enhance probative value of bullet composition “matches” as evidence of guilt is not relevant except possibly in cases where a transient is suspected of committing a particular shooting. It is the author’s position that it is not particularly relevant to compare a bullet from a crime

---

"The Battle for Civil Rights: Exposing Police & Prosecutorial Misconduct in the Quest for Equal Justice"

Join NACDL in Atlanta for an unforgettable CLE experience in the home of Southern Hospitality. Atlanta is known the world over for its parks, culinary landmarks, shopping, museums, historic sites, and entertainment. Come see all the Peach State has to offer!

For more information visit www.nacdl.org/meetings
scene in Washington, D.C., with unfired bullets recovered from regions across the country from which it is not likely that the suspect (and innocent purchasers) would have purchased bullets of similar composition(s), absent case-specific reasons, in an effort to determine a “false positive probability”, the likelihood of wrongly implying association.

In fact, further undermining probative value of bullet composition “matches” is that such evidence is less probative than that of bullet caliber. At the facilities of at least two major bullet manufacturers, compositional distinction for bullet production plays no role for all or most of the bullet calibers. At one of the two manufacturers (American Ammunition), every bullet of every caliber under production was produced from the same pallet(s) of ingots during observations and interviews conducted by the author. At the second manufacturer (Remington), the product specialist has testified that up to fifteen different calibers are produced from the same source of bullet lead. Accordingly, in cases where evidence of caliber is also prof ered, it is quite likely that bullet composition “matches” constitute cumulative evidence that is almost certainly a waste of time, with the potential for prejudicial impact likely exceeding probative value. The latter is a significant consideration if the aura of infallibility of scientific expert witnesses, especially those from the FBI Laboratory, could result in “wink and a nod” juror reactions (“we know what the witness really means”) during deliberations, if testimony about a “match” of bullet compositions is admitted.

**National Research Council Report**

In an effort to avoid the vague and ambiguous term “source” as used by FBI Laboratory witnesses, the term “compositio nally indistinguishable volume of lead” (CIVL) was defined by the National Research Council as the volume of lead that is “produced during one production run at one point in time” [emphasis added]. The NRCNA report recommends that expert testimony be strictly limited to either or both of two conclusions: that bullets from the same CIVL are more like ly to be analytically indistinguishable than bullets from different CIVLs and/or that having two bullets that are analytically indistinguishable increases the probability that two bullets came from the same CIVL versus no evidence of match status (the underlined stipulation is quite important to note). Those inferences are most likely valid and supportable scientifically and statistically.

The concern here, however, is that the significance of those limitations is subtle and may well be lost in time. The recommended change in terminology of “source” to “CIVL” does not remove the major flaw of the practice: that there exist no meaningful data or studies of how often CIVLs are produced over the past four decades that are analytically indistinguishable. By the NRCNA definition of CIVL, there is a good possibility that another CIVL is produced the following day, week or month at a bullet manufacturer, and even at other bullet manufacturers. Thus, the purported increase in probative value of “more likely” or “increases the probability” that a “match” came from the same CIVL could be so miniscule as to be meaningless. The “increase in probability” (by finding that bullets “match”) could be so insignificant that it does not remove the high risk of prejudice outweighing any probative value such testimony might have.

Remember, a CIVL is not defined as the universe of lead volume with a specific composition, only the volume of lead that is “produced during one production run at one point in time.” If a CIVL had been defined as the entire universe of lead volume with a specific composition, evidence of a “match” would make it extremely likely (but not a certainty, given that statistical criteria for association and possible analytical or systematic error could be involved) that the bullets originated in a common CIVL. Thus, the NRCNA definition of CIVL does not remove concern for the phenomenon of “repeats,” or duplicitous lots produced with the same composition.

**Findings and Recommendations**

Some of the observations and findings of the Committee relating to the analytical methodology (Phase I of the three-phase forensic practice) are as follows (italics are those of the report; numbers in brackets indicate sections and page numbers in the NAS report):

1. Studies are needed to quantify measurement repeatability and reproducibility [ES-2; 2-9];
2. Current FBI procedure is not documented in a complete and detailed format that would allow other laboratories...to practice or even fully evaluate it [2-3];
3. Laboratory protocol should be revised; all details of the procedure should be expressly addressed in the recommended revised protocol [ES-2; 2-8];
4. Better statistical basis of bullet comparison is needed [ES-2; ES-3];
5. A formal and comprehensive proficiency testing of each examiner needs to be developed by the FBI [ES-2; 2-8];
6. The FBI should publish the details of its [CBLA] procedure and the research and data that supports it in a peer-reviewed journal or at a minimum makes [sic] its analytical protocol available through some other public venue [ES-2; 2-8];
7. Peer-reviewed articles are needed after the above remedial actions are implemented [ES-2; 2-8];
8. Revised procedures must be used consistently within the FBI Laboratory [ES-2; ES-3];
9. FBI’s documented analytical protocol should be applied to all samples and should be followed by all examiners for every case [ES-2; ES-3; 2-9].

The actual Council recommendation for this phase of the procedure states, “The FBI’s documented analytical protocol should be applied to all samples and should be followed by all examiners for every case.” The use of italics in the recommendation is indicative of the Council’s recognition of the subjective and arbitrary nature of historical CBLA practice.

Some of the observations and findings of the NAS Committee relating to the statistics for comparison (Phase II of the three-phase forensic practice) are as follows:

1. The practice of data chaining is objectionable and should be discontinued, as it leads to artificially large compositional groups of analytically indistinguishable bullets [ES-3];
2. Chaining is unlikely to serve the desired purposes of identifying matching bullets with any degree of reliability [3-17];
3. The largest source of error in the use of [CBLA] is the unknown variability within the population of bullets in the United States due to variations within and across manufacturing processes [3-22];
4. This variability is not sufficiently taken into account by the statistical methods currently in use in the analysis of [CBLA] data [3-23];
5. The FBI’s methods are not representative of current statistical practice [3-23];
6. The FBI Laboratory claim of false positive probability (likelihood of falsely claiming a “match”) of 1 in 2500 is not valid [ES-3]; [author’s comment: even if it were, FPP based on national database is not relevant];
7. Even if data chaining is eliminated, the current statistical protocol used by the FBI Laboratory for CBLA should be changed and documented [3-23; 3-24];
8. When statistical protocol is revised, it should be followed by all examiners in every case [3-24].

Note, again, the Committee’s use of
Some of the findings of the NAS Committee relating to the significance of the manufacturing process in interpretation of evidence are verbatim as follows:

1. The probability that a crime scene bullet which matches a suspect’s bullet actually came from the suspect might be vastly different in an isolated small town versus a major metropolitan area [ES-4];
2. Bullet distribution information does not exist or is considered proprietary [ES-4];
3. It is not possible to obtain accurate and easily understood probability estimates [that a crime scene bullet actually came from the suspect] that are directly applicable [in part because of an absence of information on distribution of bullets in the same geographic region] [ES-4];
4. It is unclear whether macro- and microscale inhomogeneities are present at some or all of the stages of lead and bullet production and if such inhomogeneities would affect CABL [4-9]; [Ed. note: i.e., there is no body of data to support the universal assumption and assertions of homogeneity or heterogeneity claimed to support CBLA inference for decades without foundation];
5. Variations among and within lead bullet manufacturers makes [sic] any modeling of the general manufacturing process unreliable and potentially misleading in CBLA comparisons [ES-4].

Note that the report reinforced objections as to probative significance based on unknown distributional patterns. The Committee confirmed that the likelihood that a crime scene bullet that matches a suspect’s bullet actually came from the suspect might be vastly different in an isolated small town versus a major metropolitan area, an objection not only of the author but also by the pioneer of the practice, Dr. Vincent Guinn. The Committee observed that bullet distribution information does not exist or is considered proprietary and, therefore, it is not possible to obtain accurate and easily understood probability estimates that are directly applicable to the process of assessing probative value.

Some of the observations and findings of the NAS Committee relating to the legal interpretation (Phase III of the three phase forensic practice) is as follows:

1. Available data do not support any statement that a crime bullet came from a particular box of ammunition [4-27];
2.References to “box” or “boxes” of ammunition in any form is seriously misleading under Federal Rule of Evidence 403 [4-27];
3. Compositional analysis of bullet lead data alone do [sic] not permit any definitive statement concerning the date of bullet manufacture [4-27];
4. The rate of laboratory error is unknown because the FBI Laboratory does not have a program of testing by an external agency that has been designed to assess the proficiency of its examiners [4-20];
5. The FBI’s internal testing program does not determine error rate [4-20];
6. The fact that courts have generally admitted this testimony is not the equivalent of scientific acceptance, owing to the paucity of published data, the lack of independent research, and the fact that defense lawyers have generally not challenged the technique [4-22];
7. 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 [4-28].

Finally (for this article), the report recommended that the possible existence of coincidentally indistinguishable CIVLs should be acknowledged in FBI Laboratory reports and by expert witnesses on direct examination, and that the frequency with which coincidentally identical CIVLs occur is unknown. The Committee expressly indicated that available data do not support any statement that a crime scene bullet came from a particular box or boxes of ammunition, that the rate of laboratory error is unknown because the FBI Laboratory does not have a program of testing by an external agency that has been designed to assess the proficiency of its examiners, and that the FBI's internal testing program does not determine rate of error.

Some of the reasons for Professor Thompson’s insightful observation should now be more readily understood. Indeed, “[i]f this technique had not been used for years to send people to prison, no reasonable scholars of forensic evidence would consider it ready for court.”

Notes:

You never stand alone.

NACDL’s Strike Force will review your case at no cost to you, if you have been:

- Subpoenaed for improperly representing a client
- Threatened with contempt
- Hit with an improper motion to disqualify you from a case

If you are interested, call NACDL’s Strike Force at 513-721-4876 (Cincinnati), 617-227-3700 (Boston), 973-701-7080 (Chatham, NJ), or 214-871-1122 (Dallas, TX) and tell them you are interested in having your case reviewed.

For immediate assistance call the lawyers’ Strike Force Circuit Coordinator nearest you.

**CIRCUIT COORDINATORS**

1st Circuit

- Martin S. Pinales
  - Cincinnati, OH (513-721-4876)
  - mpinales@icloud.com

2nd Circuit

- Ronald J. Meshbesher
  - Minneapolis, MN (612-340-7951)
  - meshbesher@meshbesher.com

3rd Circuit

- Burton H. Shostak
  - St. Louis, MO (314-725-3200)
  - bshostak@icloud.com

4th Circuit

- Richard A. Cremer
  - Roseburg, OR (541-672-1955)
  - rcremer@rosenet.net

5th Circuit

- Alan Zegas
  - Chantilly, VA (703-701-7080)
  - alanlative@icloud.com

6th Circuit

- Charles W. Daniels
  - Albuquerque, NM (505-842-9960)

7th Circuit

- Michael L. Stout
  - Roswell, NM (505-624-1471)

8th Circuit

- David A. Elden
  - Los Angeles, CA (310-203-9595)
  - edelen@heartlink.net

**STRIKE FORCE CO-CHAIRS**

- Martin S. Pinales (Cincinnati, OH)
- Susan W. Van Dusen (Miami, FL)
- Martin G. Weinberg (Boston, MA)
- Frank Jackson (Houston, TX)
- Edward J. Imwinkelried (Miami, FL)

**NACDL STRIKE FORCE**

How Probative Is Comparative Bullet Lead Analysis?


7. Pillar, supra.
8. Ibid.
18. Imwinkelried & Tobin, supra.
19. Ibid.
25. Risinger, supra note 86, at 773.
27. Mikos, supra.
29. Testimony of Paul Birch, Remington product specialist, in People v. Marlon Smith, 02CR3477 (Division 9), El Paso County District Court, Colorado (J. Patrick Kelly, J.).
32. Collaboration by author with Dr. Wilfried Stockelein, BKA-KT1, 5/2-9/03. Dr. Stockelein is in charge of physics and chemistry matters of the German BKA Federal Laboratory. Copies of e-mails available on request.
33. Tobin & Duerfeldt, supra note 6, at 29.
34. Giannelli & Imwinkelried, supra note 4, at § 1-5(G).
37. Imwinkelried & Tobin, supra, p.72.

About the Author

William A. Tobin retired in 1998 as a Supervisory Special Agent after 27 years of service with the Federal Bureau of Investigation. During 24 years as a forensic metallurgist at the FBI Laboratory (as the de facto Chief Forensic Metallurgist), he was called upon to render his expert opinion on many occasions. He is presently a forensic metallurgical and materials science consultant in criminal, civil and non-litigious cases.

William A. Tobin
2708 Gunstock Rd.
Bumpass, VA 23024-8882
804-448-3955
Fax 775-628-9472
E-mail wtabin@netxnet.net