Editor's note: Any and all future citations of the above referenced paper should read Seward GH. Practical implications of charge transport model for electrostatic detection apparatus (ESDA). [published erratum appears in J Forensic Sci 2000;45(2)] J Forensic Sci 1999;44:832–6.

Commentary on Wu AHB, Hill DW, Crouch D, Hodnett CN, McCurdy HH. Minimal standards for the performance and interpretation of toxicology tests in legal proceedings. J Forensic Sci 1999;44(3):516–522

Sir:

The article of Wu et al. is a thought-provoking discussion of a number of relevant points concerning interpretation of toxicological testing results. The authors make the statement that there is no published conversion factor relating concentrations of 11-nor-delta-9-tetrahydrocannabinol-9-carboxylic acid (THCA) in serum to those in whole blood. While it is inconsequential to the authors' conclusions, that is not quite accurate. The data of Hanson et al. (1) quite clearly show that, in a series of nearly 50 subjects, the blood/serum concentration ratios for both delta-9-tetrahydrocannabinol (THC) and THCA are the same and that they average 0.57 (range, 0.50–0.67). I apologize for not stating this more explicitly in the 1983 article.

Randall C. Baselt, Ph.D.
Chemical Toxicology Institute
of National Medical Services
P.O. Box 8209
Foster City, CA 94404

Reference

1. Hanson VW, Buonarati MH, Baselt RC, Wade NA, Yep C, Biasotti AA, Reeve VC, Wong AS, Orbanowsky MW. Comparison of ^{3}H- and ^{125}I-radioimmunoassay and gas chromatography/mass spectrometry for the determination of Δ^9-tetrahydrocannabinol and cannabinoids in blood and serum. J Anal Tox 1983;7(2):96–102.

Commentary on Keto RO. Analysis and comparison of bullet leads by inductively-coupled plasma mass spectrometry. J Forensic Sci 1999;44(5):1020–6

Sir:

It appears to me that there is more information in Keto's (1) bullet lead impurity data than the author supposes.

Keto (1, pp. 1024–25) computed the equivalent of the scalar products of 1,770 pairs of bullet impurity concentration profiles, considered as 8-dimensional vectors. He is thus able to show that sample bullets of the same brand tend to resemble one another more often than they resemble bullets of another brand, sometimes even to the exclusion of other brands.

However, citing data insufficiency, he feels unable to assign a complete set of probabilities of brand membership to each of the concentration profiles he has available (1, Table 4). He claims only that his data "suggests that when two element signatures match, it is unlikely that the bullets originated from different sources," and that "[g]iven a sufficient database, [the scalar product] could be a useful tool in establishing the 'rarity' or 'commonality' of a specific elemental signature, and the probability of a random match [between bullets] could be estimated."

I decided to see whether a Bayesian (2,3) treatment of Keto's data might yield useful brand membership probabilities, and this appears to be the case.

By means of a multivariate Bayesian analysis of the data in Keto's Table 4, I computed brand membership probabilities. Because I lacked a separate test set, I used Keto's sample bullets both collectively, as the parametric data set, and individually, as the test set. The mutual independence of the concentration data for different elements permitted me to do this. Keto (1, p 1023) states that "[s]catter plots of each element against each of the other elements showed no visual correlations, either linear or non-linear."

Because of software limitations, I limited my analysis to ten of Keto's 12 bullet brands, ranging alphabetically from Defence through Toledo. I did, however, use all eight of Keto's element concentrations for each bullet.

I compiled, for each of 50 bullets, a probability distribution over ten bullet brands, as a function of that bullet's concentration profile. For the sake of brevity, and because the probability for the "correct brand", even when low, generally dominates the other nine values, Table 1 displays only "correct brand" assignment probabilities. Note that the table's probability scale runs from 0.50 to 1.00. (The complete parametric data set and the complete set of brand probability distributions are available on request. In only one case out of the 50 was there some ambiguity about the correct brand.)

With Table 1 in hand, one can now consider the question of decision threshold. A juryman may want a defendant's ammunition connected to the crime with a probability greater than 0.999 (odds of ~1,000 to 1), in order to vote "guilty." A prosecutor may want a probability greater than 0.85 in order to bring a case to trial. A police officer may feel that 0.75 is enough to justify arrest, and that 0.60 or more indicates "prime suspect." Assuming all this, Table 1 suggests that a Bayesian comparison of a crime scene bullet with the perpetrator's ammunition would exceed the "prime suspect" threshold about 96% of the time, that it would exceed the arrest threshold about 90% of the time, and that it would exceed the prosecution threshold about 78% of the time. As for the juryman, the bullet-brand evidence may not be quite enough, by itself, to support a "guilty" vote. The highest brand probability value I obtained was 0.998.

In closing, I point out that the issue of bullet source identification is not necessarily related to brand differences. Conceivably, several suspects may each possess a box of ammunition of the same brand (which is stamped on the case heads), each box being the result of a different production "run", with a more or less distinct set of bullet lead impurity profiles. Or so we must hope.

TABLE 1—*Distribution of 50 "test" bullets by the probability which was computed for the correct brand.*

	Correct Brand Probabilities									
Probability Range	0.50–0.55	0.55–0.60	0.60–0.65	0.65–0.70	0.70–0.75	0.75–0.80	0.80–0.85	0.85–0.90	0.90–0.95	0.95–1.0
Number of Bullets	1	1	0	2	1	3	3	4	12	23

References

1. Keto RO. Analysis and comparison of bullet leads by inductively-coupled plasma mass spectrometry. J Forensic Sci 1999;44(5):1020–6.
2. Box GEP, Tiao GC. Bayesian inference in statistical analysis. Wiley, 1992.
3. Schmitt SA. Measuring uncertainty: An elementary introduction to Bayesian statistics. Addison-Wesley, 1969.

Donald I. Promish, M.S.
68 Richardson St.
Burlington VT 05401-5002

Author's Response

Sir:

Mr. Promish appears to have missed one of the major points put forth in my article. Many of the bullets within a box do not resemble others from the same box any more than they resemble bullets from other manufacturers. This was illustrated graphically in Fig. 2 on page 1024. Calculating an "average" elemental composition for a box, and then showing that individual bullets from that box more closely resemble that average than they do averages of other boxes is meaningless. What matters, in a real case scenario, is whether known and questioned bullets match.

The purpose of bullet lead analysis, in most situations, is not to brand identify unknown bullets, as suggested in the final paragraph of Mr. Promish's letter. This is because the trace elemental composition of bullets from a single manufacturer can vary almost continuously over time, depending on the source for the lead. The analyst's opinion is usually limited to stating that the suspect bullet is consistent with (or could have come from) the same source as the known bullet. This is the same standard that applies to other forms of trace evidence as well, such as paint chips, glass fragments, and fibers, in the absence of a physical match. The "scalar products," or correlation values, were calculated in an effort to quantitate the quality of the match between two bullets, and arrive at a means of unambiguously distinguishing between known matching and nonmatching bullet leads.

While statistical approaches to data interpretation may be useful, they are, as stated in the final paragraph of my paper, "not a substitute for direct comparison of the raw elemental data when formulating an opinion as to the similarity of two bullets". A 95% probability that two bullets match means little when a look at the raw data shows that they could not be from the same melt.

I appreciate Mr. Promish's enthusiasm, and applaud his volunteering his probabilistic approach. More of this type of thinking needs to be applied to the forensic sciences. However, it must be applied with caution. I question the advisability of using the probability of a bullet match as a decision threshold for arrest, prosecution, or finding of guilt. Such decisions can only be based on a much broader scope of evidence, which could include bullet comparison. The bullet analyst cannot be expected to give a qualified opinion as to the certainty of a match; his findings must be either positive or negative to be of use.

Raymond O. Keto, M.F.S.
National Laboratory Center
Bureau of Alcohol, Tobacco, and Firearms
Rockville, MD 20850

Verbal Conventions for Handwriting Opinions

Sir:

A paper of mine in Science and Justice (1) on reporting conventions was recently the subject of a spirited debate on Docexam_L, the forensic document examiners e-mail forum (membership enquiries to andersonc@docexam.com.au). The rough and tumble of an e-mail discussion is all very well but it seems to me that if the core issues are to be resolved then the appropriate forum lies within the pages of a peer-reviewed journal. Nothing of what I say here is original and the subject has been covered in greater detail elsewhere but it seems to me to be appropriate that I should state my case in the journal that carried the letter that announced the reporting convention with which I take issue.

The letter from McAlexander, Beck and Dick (2) is to be applauded for its motivation. It promotes the idea that there is a need for standardization of terminology among experts when they express opinions. It also argues convincingly for the need to consider handwriting evidence probabilistically. Rightly, the authors pointed to the weaknesses of phrases which appear, regrettably, to be in widespread use in the forensic science world: I refer, in particular to the use of "could have" and "consistent with." I agree with McAlexander et al. that these phrases should have no place in any convention for expressing the weight of an item of scientific evidence.

The letter described a reporting convention which became the subject of ASTM standard E 1658–96 (3). My copy is headed "Standard Terminology for Expressing Conclusions of Forensic Document" (sic).

Laudable though the attempt at standardization undoubtedly is, I must point out that there is a serious problem with the manner in which the convention uses the notion of probability. In my opinion, the reporting convention is incompatible with a logical approach to evidence interpretation. Whether or not one agrees with me that this is a problem depends on whether or not one wishes to view handwriting comparison as having a sicentific rationale. If is scientific, it has to be logical; it follows that if probability is to be invoked, then the laws of probability cannot be violated.

Probabilistic thinking in relation to forensic science has, until comparatively recently, been seen to be something that evolved in the 1970's, when the paper by Finkelstein and Fairley (4) was an important milestone—though that, in turn, had evolved to some extent from lines of reasoning followed by Mosteller and Wallace (5) in considering the authorship of *The Federalist* papers. However, recent research at the University of Lausanne (6) has pinpointed the work of Poincaré, Darboux, and Appell as, apparently, the first example of what we now call the Bayesian view of forensic evidence. It is particularly germane that the reasoning of Poincaré and his colleagues was concerned with a critical review of Bertillon's evidence in a notorious handwriting case: the trial of Dreyfus for treason. In modern parlance, we would say that Bertillon committed what Thompson and Schumann (7) called the "prosecutor's fallacy". Poincaré and his colleagues pointed out the error.

It is not necessary for me to explain the Bayesian view here because of the extensive body of literature that now exists in the forensic sphere. Useful introductions to the ideas are provided by Robertson and Vignaux (8) and Aitken and Stoney (9). The key principles that emerge from this view include, first, the notion that the forensic scientist should always consider (at least) two propositions that, in the adversary system of justice, will represent the defence and prosecution positions. Next, the fundamental principle is that the scientist must address questions of the kind "what is the probability of the *evidence* given the proposition?". Questions of the kind "what is the probability of the proposition given the evidence?" are the province of the jurors, who will not only take into