J Forensic Sci, November 2007, Vol. 52, No. 6
doi: 10.1111/j.1556-4029.2007.00578.x
Available online at: www.blackwell-synergy.com

Commentary on: Nichols RG. Defending the scientific foundations of the firearms and tool mark identification discipline: responding to recent challenges. J Forensic Sci 2007; May;52(3):586–94.

Sir:

Mr. Nichols advises firearms and tool mark examiners to still the quibbling of attorneys by abandoning absolute identity conclusions and testifying that the likelihood that any tool besides the suspect tool could have produced the evidence tool mark(s) is so small that it can, for all practical purposes, be ignored. The problems with firearms and tool mark examiners' testimony are not linguistic, however, but scientific: the requisite empirical and statistical foundations have not been laid for *either* absolute *or* probabilistic identification conclusions.

The Association of Firearms and Toolmark Examiners (AFTE) Theory of Identification states that there is an exceedingly small likelihood that any tool besides the suspect tool produced the evidence tool mark(s) when the observed agreement between test and evidence tool marks is superior to that of the best known nonmatch and consonant with that of the best known match. Nichols acknowledges that "there is no universal agreement as to how much correspondence exceeds the best known nonmatching situation." He suggests, nonetheless, that in practice misidentifications will not result because of examiner underestimation of the resemblances between tool marks made by different tools. This optimistic suggestion is called into question by Masson's observation that as the number of tool marks inputted into the Integrated Ballistics System (IBIS) database grew, tool marks produced on bullets by different guns of the same caliber were found to be more similar than examiners had previously believed they could be (1). Similarly, in Biasotti et al.'s experience, many of the examiners' disagreements "stem from one examiner ascribing too much significance to a small amount of matching striae and not appreciating that such agreement is achievable in known nonmatch comparisons" (2).

Nichols states that "[w]hile oft criticized, the concept of 'I know a match when I see it' has its basis in [firearms and tool mark examiners' extensive] training." Both in testimony and in publications, however, Nichols has stated that "it's not surprising" and "not necessarily unexpected" for examiners to disagree about whether an inconclusive or an identification is the proper conclusion in a particular case (3–5). Indeed, he and his colleagues have stated that there is no standard for determining who is right in these situations and hence it is impossible to determine an actual error rate for casework (3,5). There is something basically wrong with a discipline that claims that its practitioners are trained to know it when they see it, even though different practitioners know mutually exclusive things when they see it.

The current system of proficiency testing tends to reinforce this situation. The Collaborative Testing Services, Inc. (CTS) tests are declared, rather than blind, and present examiners with simpler problems than they encounter in actual casework (2,5,6). Even so, Nichols and his colleagues are only able to report low error rates on the tests by not scoring errors when examiners reach inconclusive conclusions in regard to tool marks that were in fact made by the same or different tools. This scoring overestimates the competence both of examiners who never report exclusions when class characteristics match, and of examiners who react to their awareness of being tested by reporting inconclusives in situations where they would otherwise reach identifications (2,5,7). Peterson and Markham's classic study counted inconclusives as errors, and found error rates from 1978 to 1991 of 12% on the CTS firearms identification tests and 26% on the CTS tool mark identification tests (8).

Nichols downplays disagreements within his discipline by insisting that "CMS [Consecutive Matching Striae] is not a more objective way of performing examinations but simply a means by which an examiner can describe what he or she is observing in a striated tool mark comparison." He also describes CMS as an attempt "to standardize the concept of the best-known nonmatch discipline wide." If, as Nichols claims, CMS is not "a different method than has been practiced throughout the years," the CMS identification criterion must be such a malleable standard that when examiners disagree, as they do under the traditional approach, they each can manipulate CMS to show that they are right. To contribute to standardization, however, the CMS criterion must be inflexible enough to settle disagreements that arise under the traditional approach. Nichols fails to realize that unless CMS is more objective than the traditional "I know it when I see it" approach, there is no justification for using CMS to decide that some, but not other, examiners' conclusions are right.

Nichols claims to see a difference between DNA and firearms and tool mark identification that allows him to argue that neither tool databases nor calculations of frequencies are needed to lay a foundation for testimony that the likelihood that any tool besides the suspect tool made the evidence tool mark is vanishingly small. "[T]he characteristics being compared in DNA profiles are actually subclass characteristics. It is the relatively *rare* frequency with which a combination of subclass characteristics occur in a given population that allows DNA to approach, although not fully achieve, individualization. In firearms and tool mark identification, examiners are not basing their identifications on subclass characteristics or frequencies of combinations thereof, but individual characteristics." Nichols' characterization of DNA is, at best, highly controversial. More importantly, the contrast he draws between DNA and firearms and tool mark identification rests on a fundamental misunderstanding of the nature of tool marks. The relationship between the probabilistic nature of identity conclusions and the component nature of individual characteristics of tool marks was recognized as early as 1935: "It is probably true that no two firearms with the same class characteristics will produce the same signature, but it is likewise true that each element of a firearm's signature may be found in the signatures of other firearms .... An individual peculiarity of a firearm can, therefore, be established by elements of identity which form a *combination* the coexistence of which is highly improbable in the signature of other firearms with the same class characteristics" [Gunther and Gunther (9), quoted by Biasotti and Murdock (10)].

Nichols fails to realize that the impermanence of tool marks creates distinctive difficulties for the statistical treatment of firearms and tool mark identification. An individual's nuclear DNA remains the same over time, so any difference in the tested alleles on a suspect's and crime scene DNA automatically excludes the suspect as the source of the crime scene DNA. By contrast, differences between evidence and test tool marks *may* mean that the marks were made by different tools or *may* mean that a single tool

changed over time. Therefore, statistical assessments of the significance of differences between evidence from the suspect and the crime scene are needed for firearms and tool mark, but not nuclear DNA, identifications. Nichols dismisses this difficulty by suggesting that the impermanence of tool marks can cause missed identifications but not misidentifications. However, in the 1959 study which Nichols described in 1997 as "the most exhaustive statistical empirical study ever published," Biasotti found matches of 21–38% and 15–20% of the striae per land or groove impression on bullets respectively fired from the same and different .38 Special Smith & Wesson revolvers (11,12). This near-complete overlap in the amount of similarity in tool marks produced by the same and different guns strongly suggests that examiners can err by wrongly attributing differences between tool marks made by different tools to changes in the same tool over time.

Distinctive difficulties for firearms and tool mark identification also arise in regard to the interpretation of complete (or near total) similarity between evidence from a crime scene and a suspect. With the exception of identical twins, each individual's genome is unique. Therefore, a match on all sites of the genome would necessarily exclude anyone but the suspect or an identical twin as the source of crime scene DNA. Indeed, a match on all thirteen sites of the genome currently tested makes it virtually impossible that anyone but the suspect or an identical twin could be the source. By contrast, even a complete (or near total) match between the microscopic patterns on tool marks need not mean that they were produced by the same tool. Instead of being an individual characteristic unique to the tool marks produced by a single tool, the complete microscopic pattern might be common to all tool marks produced by tools in a batch—what firearms and tool mark examiners standardly refer to as a "subclass characteristic." The 19 studies that Nichols cites provide rough rules of thumb about circumstances in which subclass characteristics are or are not likely to occur. They do not, however, provide either strict rules for determining whether a microscopic pattern on a tool mark is an individual or a subclass characteristic or strict rules as to which tools or manufacturing processes do or do not produce tool marks with subclass characteristics. To avoid misidentifications based on confusing subclass characteristics shared by more than one tool with individual characteristics unique to one and only one tool, examiners need to rely on personal familiarity with types of forming and finishing processes and their reflections in tool marks. In a welcome departure from his recent testimony in federal court that "it's not very difficult" to distinguish between subclass and individual characteristics (3), Nichols now states that "[t]he difficulty of addressing subclass characteristics is not in debate."

References

1. Masson JJ. Confidence level variations in firearms identifications through computerized technology. AFTE J 1997;29(1):42–4.
2. Biasotti A, Murdock J, Moran BR. Scientific issues. In: Faigman DL, Kaye DH, Saks MJ, Sanders J, Cheng EK, editors. Modern scientific evidence: the law and science of expert testimony, Vol. 4. St. Paul: West, 2006–2007;544–79.
3. Transcript of hearing. United States vs. Diaz (ND Cal. January 24, 2007).
4. Nichols R. The scientific foundations of firearms and tool mark identification: a response to recent challenges. CAC News 2006;2nd Quarter: 8–27.
5. Grzybowski R, Miller J, Moran B, Murdock J, Nichols R, Thompson R. Firearm/toolmark identification: passing the reliability test under federal and state evidentiary standards. AFTE J 2003;35(2):209–41.
6. Peterson JL, Markham PN. Crime laboratory proficiency testing results, 1978–1991, I: identification and classification of physical evidence. J Forensic Sci 1995;40:994–1008.
7. Thompson E, DeKinder J. Range of exclusions. AFTE J 2006;38(1): 51–5.
8. Peterson JL, Markham PN. Crime laboratory proficiency testing results, 1978–1991, II: resolving questions of common origin. J Forensic Sci 1995;40:1009–29.
9. Gunther JD, Gunther CO. The identification of firearms. New York: John Wiley & Sons, Inc., 1935.
10. Biasotti AA, Murdock J. "Criteria for identification" or "state of the art" of firearms and toolmark identification. AFTE J 1984;16(4):17–25.
11. Biasotti A. A statistical study of the individual characteristics of fired bullets. J Forensic Sci 1959;4(1):34–50.
12. Nichols R. Firearm and toolmark identification criteria: a review of the literature. J Forensic Sci 1997;42(3):466–74.

Adina Schwartz, J.D., Ph.D.
John Jay College of Criminal Justice
City University of New York
New York, NY
E-mail: aschwartz@jjay.cuny.edu