Proficiency Test Results from Peterson and Markham Article - Firearms

Source: “Crime Laboratory Proficiency Test Results, 1978-1991, II: Resolving Questions of Common Origin,” *Journal of Forensic Sciences*, Vol. 40, No. 6, November 1995, pp. 1009 - 1029. (12 separate tests involving between 42 and 173 laboratories.)

From Table 8, page 1019:

Total comparisons = 2106

False identifications = 12
False eliminations = 17

True identification conclusions = 905
True elimination conclusions = 954

True identifications judged inconclusive = 43
True eliminations judged inconclusive = 175

Total true identifications = 905 + 43 + 17 = 965
Total true eliminations = 954 + 175 + 12 = 1141

Total identification conclusions offered = 905 + 12 = 917
Total elimination conclusions offered = 954 + 17 = 971

Total inconclusives = 43 + 175 = 218

Data Analysis – Firearms

Test Sensitivity = true IDs offered/true IDs = 905/965 = 93.78%

Test Specificity = true eliminations offered/true eliminations = 954/1141 = 83.61%

False positive error rate (false or mis-identifications) = false positive responses/total true eliminations = 12/1141 = **1.05%**

False negative error rate (false or mis-eliminations) = false negative responses/total true identifications = 17/965 = **1.76%**

Inconclusive rate = 218/2106 = 10.35%

Proficiency Test Results from Peterson and Markham article - Toolmarks

Source: “Crime Laboratory Proficiency Test Results, 1978-1991, II: Resolving Questions of Common Origin,” *Journal of Forensic Sciences*, Vol. 40, No. 6, November 1995, pp. 1009 - 1029. (12 separate tests involving between 72 and 163 laboratories.)

From Table 13, page 1024:

Total comparisons = 1961

False identifications = 30
False eliminations = 44

True identification conclusions = 646
True elimination conclusions = 755 + 53 + 44 = 852

True identifications judged inconclusive = 83 + 48 = 131
True eliminations judged inconclusive = 258

Total true identifications = 646 + 44 + 48 = 821
Total true eliminations = 852 + 30 + 258 = 1140

Total identification conclusions offered = 646 + 30 = 676
Total elimination conclusions offered = 852 + 44 = 896

Total inconclusives = 83 +258 + 48 = 389

Under toolmarks, the authors include a category of “unjustified exclusions.” An example: two wires cut by different areas on the cutting edge of a single pair of wire cutters was marked by a participant as an elimination. While this mistake would be understandable if one merely considers microscopic correspondence and ignores the larger picture, it was properly categorized as an unjustified exclusion, and counted here as a false negative. In other cases, however, the responses were correct from a scientific perspective (only false positives and false negatives matter), but incorrect from a training and quality assurance perspective. For my purposes, the scientific propositions trump quality assurance considerations, and thus the remaining “unjustified exclusions” were counted as correct responses.

Data Analysis – Toolmarks

Test Sensitivity = 646/821 = 78.68%

Test Specificity = 852/1140 = 74.74%

False positive error rate = 30/1140 = **2.63%**

False negative error rate = 44/821 = **5.36%**

Inconclusive rate = 389/1961 = 19.84%

Points of Explanation, Clarification, and Discussion

In the text of the article, the authors put forth the following figures for the firearms tests:

- 88% of responses agreed with the manufacturer of the test.
- 1.4% of responses disagreed with the manufacturer.

The 88% figure approximately equals 100 – (1.4 + 10.35), which = 88.25. The 1.4% figure derives from the total false positive and negative responses, divided by the total number of comparisons.

The authors put forth the following figures for the toolmarks tests:

- 74% of responses agreed with the manufacturer of the test.
- 4% of responses disagreed with the manufacturer.

The 74% figure derives from the following: 1 – [30 + 141 + (389 – 48)]/1961 = .7389. This does not exactly parallel the 88% figure above, owing to the complicating factor of the 48 true inconclusives. The 4% figure derives from (30 + 41)/1961. This is analogous to the 1.4% figure for firearms, except for toolmarks the authors did not count unjustified exclusions as responses that disagreed with the manufacturer of the test.

Test Sensitivity and Test Specificity are, along with false positives and negatives, indicators of test validity and overall quality (here a test = a microscopic examination). These are commonly used indicators for laboratory tests such as the one for HIV. For our purposes these indicators are somewhat less useful because not all conclusions are positively "yes" or "no." "Inconclusive" is often a perfectly legitimate and necessary result. Test Specificity is more problematic than Test Sensitivity because eliminations are, properly, seldom effected on the basis of microscopic marks but rather on the basis of significantly different class characteristics. If, for example, a validity or proficiency test packet contains mostly bullets with the same rifling characteristics, yet are from different barrels, a preponderance of inconclusive conclusions will properly result. This in turn would yield what seems to me to be an artificially low Test Specificity figure. Much of this stems from the fact that, unlike blood or fingerprints, barrels change over time and from firing to firing.

A note of caution: These CTS data are very roughly indicative of the validity of microscopic examinations. But one must avoid falling victim to using the above figures by themselves to assess overall probabilities, either in general or for a particular case. Leaving aside the issues of individual examiner abilities, case difficulty, and the validity of these data, etc., in order to calculate a probability that a bullet was fired from a particular barrel, we must invoke a base rate, or prior odds. That discussion lies beyond the scope of this document, but suffice it to say that an examiner should <u>not</u> assert that the probability he will mis-identify in a bullet comparison is 1.05%.

Finally, any errors or relevant omissions contained in this document are solely the responsibility of the author, Stephen Bunch, Firearms-Toolmarks Unit, FBI Laboratory. Likewise, any opinions expressed are not necessarily those of the FBI Laboratory or the SWGGUN membership.