JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, March 2017, Vol. 62, No. 2
doi: 10.1111/1556-4029.13274
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

CRIMINALISTICS

Fabiano Riva,[1,2] Ph.D.; Rob Hermsen,[3] B.Sc.; Erwin Mattijssen,[3] M.Sc.; Pascal Pieper,[3] B.Sc.; and Christophe Champod,[1] Ph.D.

Objective Evaluation of Subclass Characteristics on Breech Face Marks

ABSTRACT: Subclass characteristics can be found on the breech face marks left on spent cartridge cases. Even if they are assumed to be rare and their reported number is small, they can potentially lead to false associations. Subclass characteristics have been studied empirically allowing examiners to recognize them and to understand in which conditions they are produced. Until now, however, their influence on the identification process has not been studied from a probabilistic point of view. In this study, we aim at measuring the effect of these features on the strength of association derived from examinations involving subclass characteristics. The study takes advantage of a 3D automatic comparison system allowing the calculation of likelihood ratios (LRs). The similarities between cartridge case specimens fired by thirteen S&W .40S&W Sigma pistols are quantified, and their respective LRs are computed. The results show that the influence of subclass characteristics on the LRs is limited, even when these features are prevalent among the potential sources considered in a case. We show that the proportion of firearms sharing subclass characteristics should be larger than 40% of the pool of potential firearms for the effect to be significant.

KEYWORDS: forensic science, firearm identification, subclass characteristics, likelihood ratio, cartridge cases, 3D topographies, evaluation

When evaluating the strength associated with a comparison between marks left on cartridge cases, the firearm examiner has to evaluate the amount and the quality of the similarities (or the differences) between marks. During this phase, which is subjective in nature (1), firearm examiners will weigh the similarities in function of their quality and number based on their knowledge and experience. Examiners will traditionally distinguish the so-called individual characteristics from the characteristics that can be shared by more than one firearm, such as in the case of "subclass characteristics." When subclass characteristics are recognized, they will typically not be further considered in the assessment. General guidelines have been provided to examiners to help them identify these characteristics (2). Case reports involving subclass characteristics have been published in the literature alongside with research investigating the origin of them during the manufacturing processes (3–10). The number of actual cases reported in the literature is limited, probably due to the rare occurrence of these features in casework (2). However, recognizing subclass characteristics is not an easy task, and some have rightly indicated that the ability of examiners to detect them is not well established (11). Even if we assume that examiners possess the means to recognize and isolate subclass features, questions can be asked about how these characteristics could or should be used in the further steps of the identification process, namely the evaluation of the results. Until now, the value of the subclass characteristics has not been systematically studied from a probabilistic point of view, because they are simply ignored when identified.

This research investigates the impact of the presence of potential sources sharing subclass characteristics on the interpretation of the results of the comparison process. For this purpose, an automatic comparison system based on topographical measurements (3D) has been used to compare marks left on the primer cup of spent cartridge cases. This system has been coupled with a bi-dimensional statistical model allowing the calculation of likelihood ratios (LRs) using the scores generated by the system. The main strength of this system is the limited intervention of the human operator that ensures an objective procedure to evaluate the comparison results between two impressed marks, such as the breech face mark (BF). The system used for this study is fully described elsewhere (12). We recall that the system is based on the LR paradigm and its performance in terms of accuracy has been measured by the rates of misleading evidence and reported in (12). It is a score-based likelihood ratio system, a framework similar to the one used in other areas such as handwriting (13). The underpinning idea of this research is to avoid the need for the examiner to identify and then exclude these features from the comparison process, and to explore systematically the statistical impact of the presence of a given proportion of specimens sharing subclass characteristics on the overall evidential strength to be assigned to the forensic findings. By forensic findings, we mean the level of agreement between striated marks left on the breech face of fired cartridge cases as measured by our system. To assign LRs, two datasets of scores have to be established: the *within* distribution (scores resulting from comparisons between impressions left by the same firearm) and the *between* distribution (scores resulting from comparisons between impressions left by different firearms). The aim is then to

[1]School of Criminal Sciences, University of Lausanne, 1025 Lausanne-Dorigny, Switzerland.
[2]Institute of Legal Medicine, University of Bern, Bühlstrasse 20, Bern 3012, Switzerland.
[3]Netherlands Forensic Institute, PO Box 24044, 2490 AA, The Hague, The Netherlands.

Received 2 July 2015; and in revised form 9 May 2016; accepted 28 May 2016.

explore how the assignment of the LRs changes when specimens fired by firearms sharing subclass characteristics are present at a varying proportion in the *between* dataset.

Material and Methods

Specimens Used

Specimens fired from Sigma pistols were made available to us by Rivera (9) and Lightstone (10). The manufacturing process involved in the fabrication of the Sigma pistol slides is fully described in the above papers. Rivera (9) reported findings on two different S&W .40S&W Sigma pistols showing close serials numbers (serial #PVB7152 and #PVB7164). The impressions left on the primer cups by the pistols are prominent and continuous along the whole surface (Fig. 1).

Lightstone (10) published results based on 10 consecutively manufactured slides of S&W .40S&W Sigma pistols. The specimens, fired by the consecutively manufactured slides, show, except for the shears, less prominent striae compared to the specimens from Rivera (Fig. 2).

The comparison between known non match (KNM) specimens from Rivera shows a high level of similarity among the impressed breech face mark (Fig. 3).

For the specimens from Rivera, the presence and the magnitude of the shear marks are highly dependent on the ammunition type. For example, the Remington–Peters cartridge cases exhibited the most primer shear marks (Fig. 1 – middle), whereas the CCI specimens did not show any shear marks (Fig. 3).

The specimens' cartridge cases obtained have been organized in three different groups (Table 1):

A Three specimens fired by two firearms having close serial numbers (#PVB7152 and #PVB7164) and sharing prominent subclass characteristics on the breech faces (Fig. 3)—from firearms used by Rivera (9). This means two cartridge cases fired by the firearm #PVB7152 and one by the firearm #PVB7164;

B Ten specimens not showing obvious subclass characteristics (Fig. 2)—these specimens have been fired by the ten consecutively manufactured slides described by Lightstone (10). This means one cartridge case fired by each of the slides;

C Seven specimens fired by one S&W .40S&W Sigma pistol without any obvious relationship with the firearms used to fire the specimens in groups A and B.

For the rest of this study, we will posit that group A and group B represent specimens that share potential subclass characteristics within each group, even though they are more prominent for group A than for group B.

All specimens selected for this study have been fired using Federal ammunition .40S&W (FMJ—180 gr.) with a brass primer cup. This ammunition led to shear marks of a limited area (Fig. 1—right). The shear marks are decisive to help with the discrimination between cartridge cases sharing subclass characteristics. The ammunition used in this study tends to reduce the availability of these features for discrimination.

The Likelihood Ratio

The likelihood ratio (LR) is a numerical value that expresses the strength of the forensic evidentiary findings (E, commonly

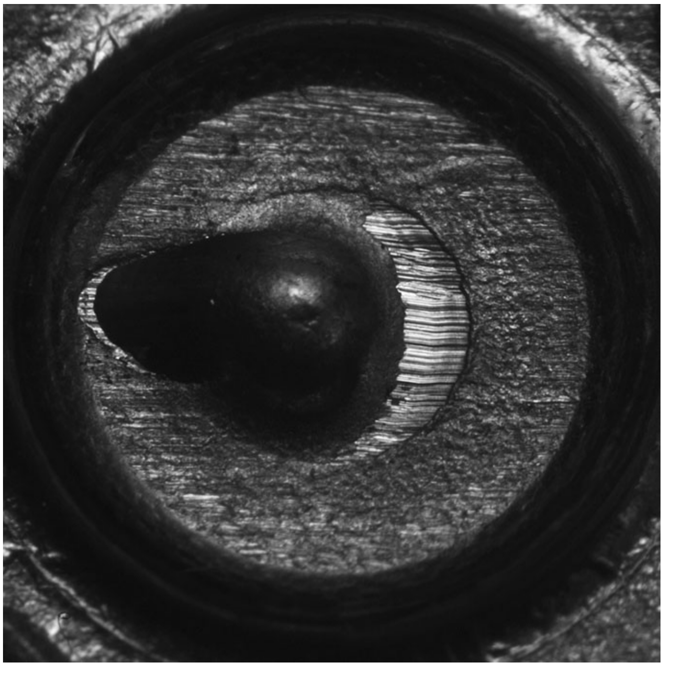

FIG. 2—*Image showing a breech face mark left by one of the ten consecutively manufactured slides used by Lightstone (10).*

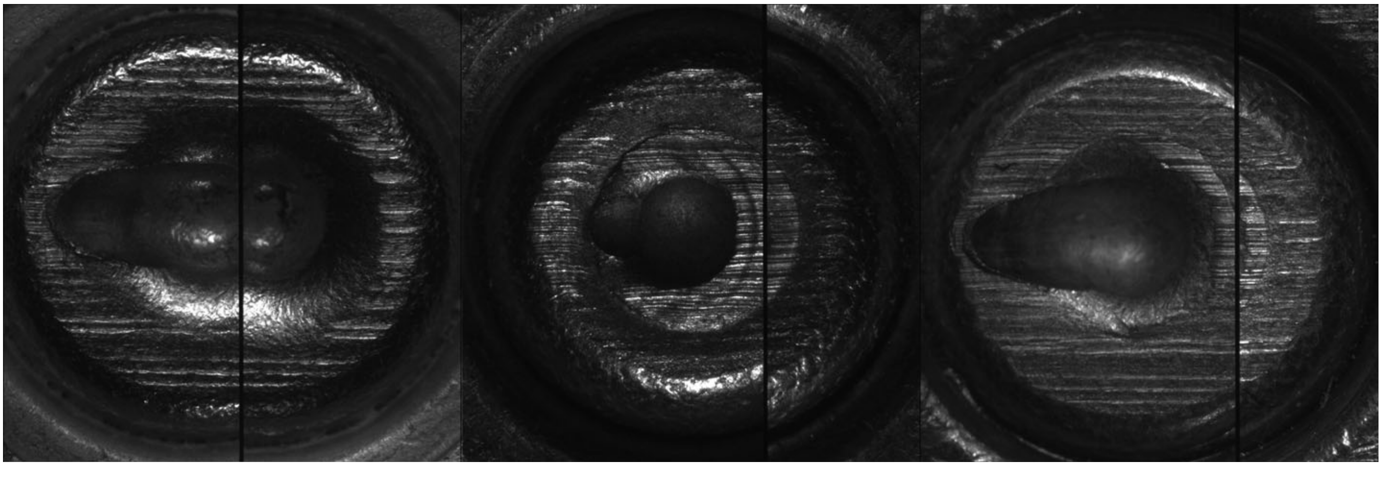

FIG. 1—*Images showing the comparison between specimens fired by the same S&W .40S&W Sigma pistol with three different brands of ammunitions. From left to right: respectively, CCI, Remington–Peters, and Federal.*

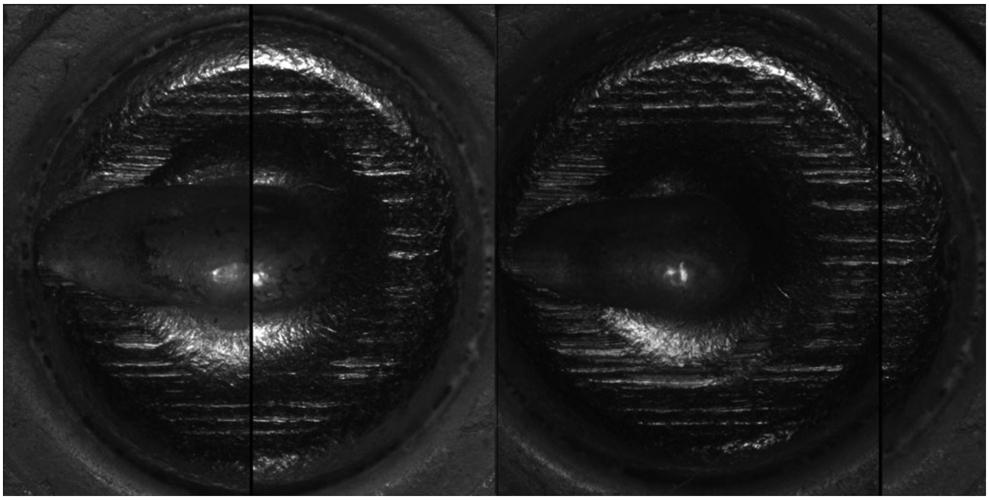

FIG. 3—*Breech face mark comparison between cartridge cases fired by the two pistols involved in the cases reported by Rivera (9). In this example, the CCI ammunition was used to produce the test fired.*

TABLE 1—*Description of the specimens used in this research.*

Group A		Group B	Group C
1 firearm (#PVB7152)	1 firearm (#PVB7164)	10 consecutively manufactured slides	1 firearm
1 cartridge case	2 cartridge cases	10 cartridge cases	7 cartridge cases

used for "evidence"). It is obtained by the ratio of the probabilities of the findings considered under two propositions that we relate, respectively, to the prosecution (H_p) and the defense (H_d) proposition (14). In the context of this study, we would like to assign the strength of the forensic findings conditional on the class characteristics. Hence, the respective propositions are formulated in the following way:

- H_p: The questioned cartridge case has been fired by the suspect firearm.
- H_d: The questioned cartridge case has been fired by another firearm with the same caliber and the same class characteristics as the suspect firearm.

To assign a likelihood ratio to a given comparison between two breech face impressions, the similarity scores from the automatic comparison system will be weighed against, respectively, the *within-source* density distribution and the *between-source* density distribution. It is important to stress that this leads to a score-based likelihood ratio. It is based on the scores derived from a comparison and the LR so computed refers to these metrics alone.

The *within source* relates to the distribution of the scores when it is known that the compared impressions originate from the same source (H_p). By "same source" in this study, we mean the given firearm used to generate the seven cartridge cases under group C in Table 1. All data points used to compute the *within-source* distribution are based on this firearm. The *between sources* relate to the distribution of the scores obtained when cartridge cases fired from different firearms are compared (H_d). Each score provided by the comparison system is based on different features (a full account can be found in (12)): the first on the Euclidean distance, the second is a correlation index, and the last is based on the properties of the normal vectors to the surface. To reduce the dimensionality of the problem, PCA (principal component analysis) is applied to allow the selection of the two principal components (PC 1 and PC 2) that offer the highest contribution to the variability. In this way, it is possible to describe the comparison between marks using just these two values that will act as similarity scores. A data point is a two-dimensional score (consisting of two values from the PCA) resulting from the comparison between two breech face impressions. The probability densities are estimated by parametric modeling using a bivariate normal distribution.

Ways to Construct the Between-source Distribution

Under H_d, we will consider two scenarios. The first represents a situation where only comparisons arising from cartridge cases not sharing subclass characteristics will be used (hence between groups A, B, and C). For the second, we will add to the first a given percentage of cases involving comparisons between cartridge cases sharing subclass characteristics (hence within groups A or B).

The aim is to investigate the behavior of the likelihood ratio when we move from the first scenario to the second. In doing so, we will understand the impact of the presence of firearms sharing subclass characteristics in the population of firearms used to construct the *between*-source distribution on the resulting LRs. We expect to observe decreasing LRs as we move from the first to the second scenario.

Experimental Design

The data points that allow to construct the *within* distribution come from the 21 pairwise comparisons between the seven cartridge cases of group C. They are depicted by the circles in Fig. 4. The data points toward the first *between* distribution, namely *without subclass,* are obtained from the pairwise comparisons between cartridge cases from different sources but not sharing any subclass characteristics (represented by triangles in Fig. 4). The data points of the *between with subclass* distribution are obtained from the comparisons between specimens from different sources showing subclass characteristics (represented by crosses in Fig. 4).

To explore the impact of subclass characteristics on the LRs, data from the *between with subclass* distribution are progressively added, to the data from the *between without subclass* distribution. We have 47 data points arising from comparisons *with subclass*. They are iteratively added (i = 1 to 47) one by one to the 43 data points of the set *without subclass*.

Hence, for i = 1 to 47, we proceed as follows (Fig. 5):

A Add to the *between without subclass* data points i randomly selected points from the *with subclass*.

B Fitting *between* distributions resulting from point A by means of a bi-dimensional normal distribution.

C Calculate a LR for each point of the *within* distribution (21 LRs). It represents LRs obtained when H_p is true, denoted $LR_i^{\boldsymbol{P}}$.

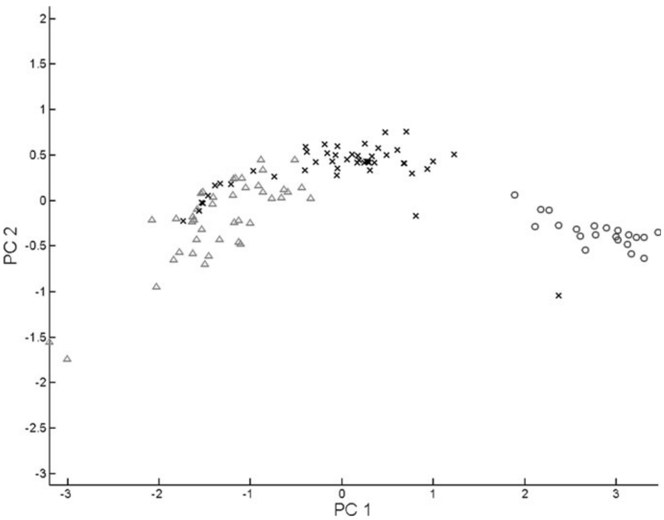

FIG. 4—*Bi-dimensional within (circles), between with subclass (crosses), and between without subclass (triangles) data points.*

D Calculate an LR for each point of the *between* distribution. The mean of these 43 + i LRs obtained when H_d is true, forming the set $LR_i^{\boldsymbol{D}}$.

E The steps A to D are repeated 500 times by resampling i with replacement. This allows calculating 500 sets of $LR_i^{\boldsymbol{P}}$ and $LR_i^{\boldsymbol{D}}$ from which we extract the mean.

At the end of the process, the results are represented by 47 mean values of $LR_i^{\boldsymbol{P}}$ and 47 mean values of $LR_i^{\boldsymbol{D}}$.

Results

LRs under H_p

Figure 6 shows the transition of the mean LRs—$LR_i^{\boldsymbol{P}}$ calculated with a total of 21 values for each value of i—under the prosecution proposition for comparisons between specimens fired by the same firearm as a function of the percentage of *with subclass* data points added to the *without subclass* data points. The means are shown with ± 1 standard deviation.

The results show a steep decrease in the mean LR in function of the increasing percentage of comparisons affected by potential subclass characteristics. This phenomenon is caused by a gradual overlapping (increasing with i) between the *within* and the *between* distributions.

LRs under H_d

Figure 7 shows the transition of the mean LRs under the defense hypothesis ($LR_i^{\boldsymbol{D}}$) for comparisons between specimens fired by different firearms as a function of the percentage of *with subclass* data points added to the *without subclass* data points. The means are shown with ± 1 standard deviation.

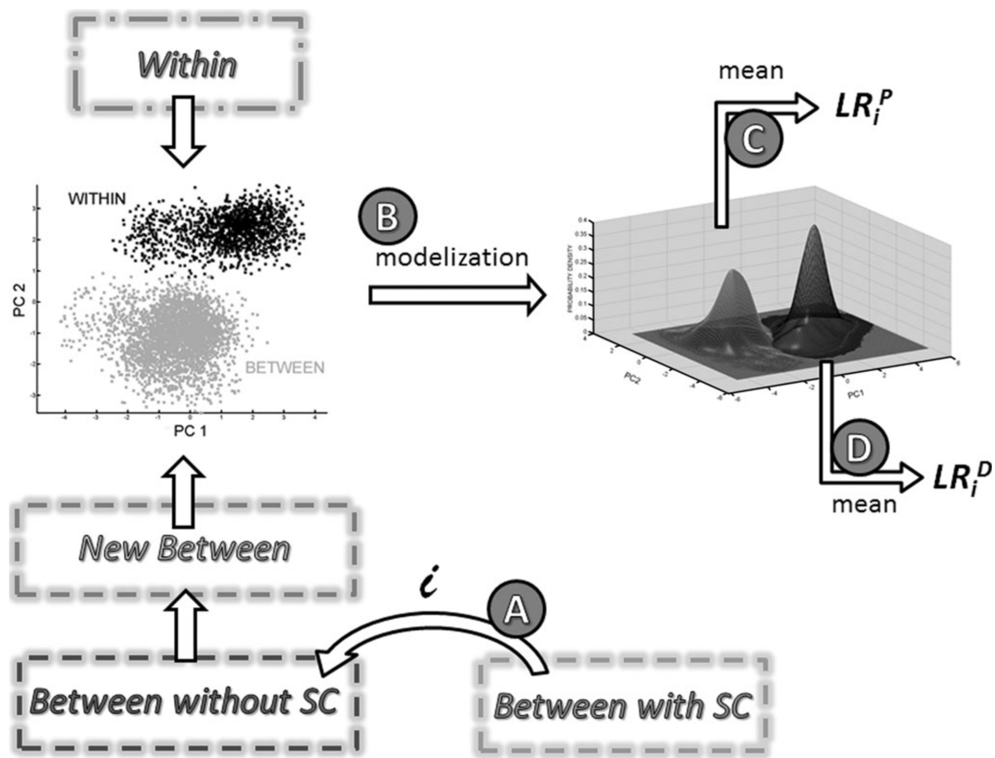

FIG. 5—*Resampling scheme adopted to increase the number of with subclass characteristics cases into the between distribution from* i = *1 to 47.*

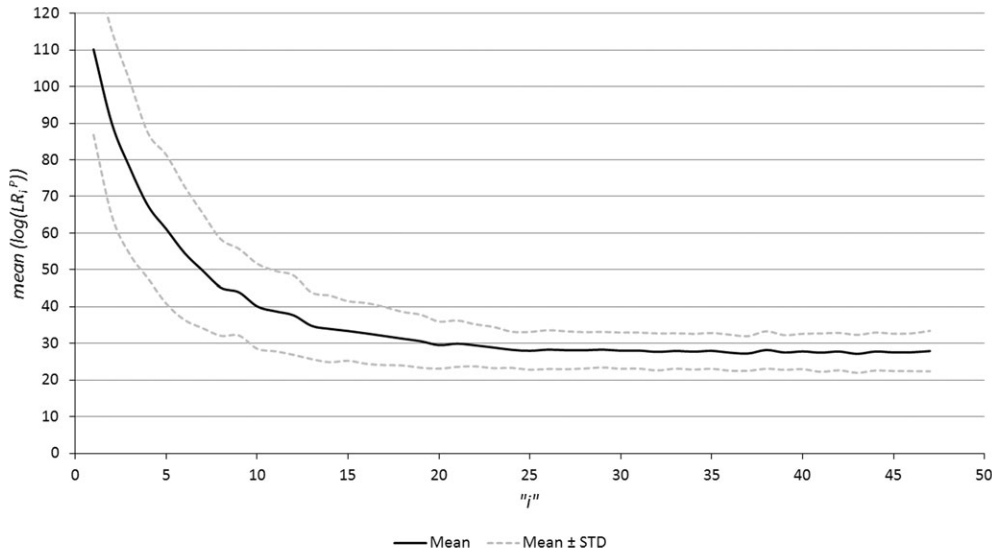

FIG. 6—*Evolution of the mean LRiP as a function of the number of with subclass data points added to the without subclass data points (43 points in total).*

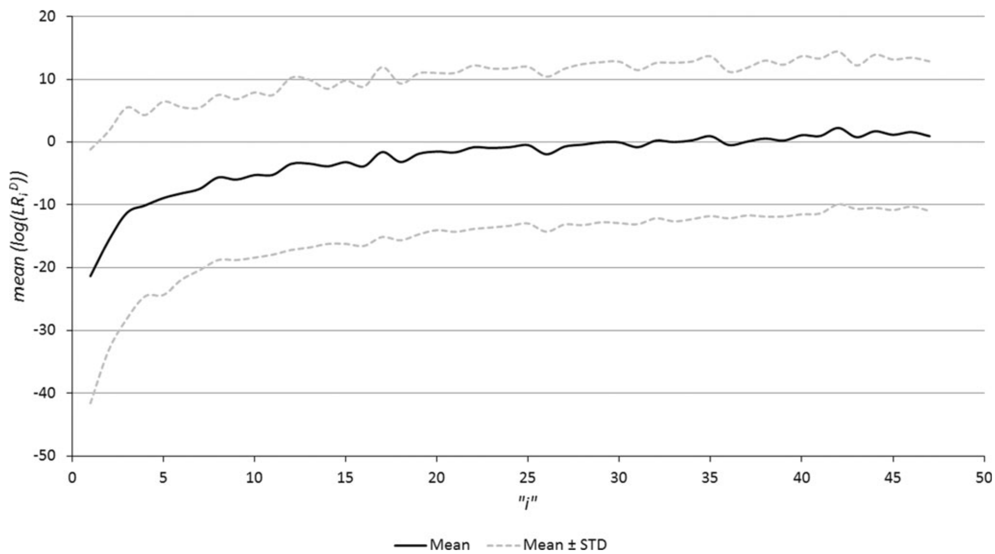

FIG. 7—*Evolution of the mean LRiD as a function of the number of with subclass data points added to the without subclass data points (43 points in total).*

The results show a reduction in the average level of support toward H_d when the percentage of "subclass specimens" increases, to the point that it can even support H_p. However, to obtain mean LRs greater than one, the percentage of comparisons between specimens sharing potential subclass characteristics added to the *between* distribution is large, namely greater than 40% of the global population of interest. This means that for a pool of 100 firearms, there are 40 cartridge cases fired from different firearms sharing subclass characteristics.

Discussion and Conclusion

This study takes advantage of an automatic comparison system to quantify the similarities between breech face marks on primer cups of spent cartridge cases. This system allows generating similarity scores that serve the probabilistic approach underpinning these results. The aim of the study was to assess the impact of the presence of subclass characteristics on the evidential strength that can be assigned to a given comparison between two breech face marks.

This objective would have been difficult to fulfill using a traditional comparison technique, namely relying on the judgment of a firearm examiner taking advantage of a comparison macroscope. Indeed, methods based on human experts suffer from operator influences and potential bias that put heavy constrains on the experimental design (15). In addition, such empirical methods would offer no mechanism to measure and quantify the similarities and difference between two marks.

Contrary to the traditional approach, where subclass characteristics have to be recognized and ruled out, the approach proposed in this study allows accounting for the possibility of subclass characteristics without the need for them to be recognized or excluded from the population at the outset of the evaluation phase.

The results showed that the LRs under H_p decrease when specimens sharing potential subclass characteristics are added into the *between* distribution. This is the logical effect of adding specimens that are expected to be close to each other in the considered *between* population. Under H_d, the effect is to reduce the level of support in favor of that proposition (hence an increase in the LR) up to a point where the findings do not provide support for any proposition.

This effect has been observed by contrasting a situation where subclass is ignored with a situation with an increasing number of cases involving subclass characteristics, up to more than 50%. It is obvious that in casework, such a mixing proportion of up to 50% is not realistic. The reported number of firearms showing subclass characteristics is notoriously low, and the number of reported items per subclass is equally low (approximately 2—from Rivera—to 10—from Lightstone). The specimens from Rivera (9) have very well-defined subclass characteristics and represent probably one of the best known non matches discovered in the community. So far, to our knowledge, no additional firearm has been reported with the type of breech face characteristics as the two reported by Rivera.

The critical question is what should be the proportion of specimens sharing subclass characteristics into the relevant population considered for a specific casework? We are confident to say that 50% is too high for the reasons exposed above. In our opinion, a proposition of 10% (this value is not meant to be a threshold) is already an extreme scenario—even though it is known that subclass features may be observed in S&W Sigma breech faces—and that would mean an addition of about five cases in our experiments. Under such condition, the LR under H_p would reduce but not to the point to make the findings inefficient to support that view. Likewise, the LR under H_d will increase but still remain largely below 1 (or a $\log_{10}$(LR) below 0), hence providing support for that proposition. Overall, the information capability of the system reduces but not to the point to make it completely uninformative with the regard to the propositions at hand. It means that once we operate in a probabilistic framework, the possibility of the presence of subclass characteristics can be accounted for and, within reasonable circumstances, they will have no drastic impact on the evidential strength of the evidence.

However, when a LR-based system such as the above is used, the reference population of specimens should also contain specimens sharing subclass characteristics, if they exist. These cases tend to significantly decrease the LR under H_p.

Finally, even if the results obtained in this study illustrate the impact of subclass characteristics for a given make and model of firearm, they cannot be easily transposed to all firearms at this stage. We remain conscious of the limitation of the sample used here. It is known that the quality and the quantity of these features will vary as a function of the type of firearms and the manufacturing process. However, we believe that the specific case of S&W Sigma considered here can be seen as a worst-case scenario.

Acknowledgments

We would like to thank Gene Rivera and Laura Lightstone for their great help and for providing us the main part of the specimens used in this research.

References

1. AFTE Criteria for Identification Committee. Theory of identification, range of striae comparison reports and modified glossary definitions – an AFTE criteria for identification committee report. AFTE J 1992;22 (2):336–40.
2. Murdock J, Biasotti A, Moran B. Firearm and toolmark identification – scientific issues. In: Faigman D, Kaye D, Saks M, Sanders J, Cheng E, editors. Modern scientific evidence: the law and science of expert testimony. St. Paul, MN: Thompson-West, 2006-2007;§36:6–§36:13.
3. Matty W. Raven 25 automatic pistol breech face tool marks. AFTE J 1984;16(3):57–60.
4. Uchiyama T. Similarity among breech face marks fired from guns with close serial numbers. AFTE J 1986;18(3):15–52.
5. Thompson E. Phoenix arms breech face tool marks. AFTE J 1994;26 (2):134–5.
6. Lardizabal PP. Cartridge case study of the Heckler & Koch USP. AFTE J 1995;27(1):49–51.
7. Lopez L, Grew S. Consecutively machined Ruger bolt faces. AFTE J 2000;32(1):19–24.
8. Coffman B. Computer Numerical Control (CNC) production tooling and repeatable characteristics on ten Remington model 870 production run breech bolts. AFTE J 2003;35(1):75–8.
9. Rivera GC. Subclass characteristics in Smith & Wesson SW40VE sigma pistols. AFTE J 2007;39(3):247–56.
10. Lightstone L. The potential for and persistence of subclass characteristics on the breech faces of SW40VE S&W Sigma pistols. AFTE J 2010;42 (4):308–22.
11. Schwartz A. A systemic challenge to the reliability and admissibility of firearms and toolmark identification. Columbia Sci Technol Law Review 2005;6:1–42.
12. Riva F, Champod C. Automatic comparison and evaluation of impressions left by a firearm on fired cartridge cases. J Forensic Sci 2014;59 (3):637–47.
13. Davis LJ, Saunders CP, Hepler A, Buscaglia J. Using subsampling to estimate the strength of handwriting evidence via score-based likelihood ratios. Forensic Sci Int 2012;216(1–3):146–57.
14. Kerkhoff W, Stoel RD, Mattijssen EJAT, Hermsen R. The likelihood ratio approach in cartridge case and bullet comparison. AFTE J 2013;45 (3):284–9.
15. Buckleton J, Triggs C, Taroni F, Champod C, Wevers G. Experimental design for acquiring relevant data to address the issue of comparing consecutively manufactured tools and firearms. Sci Justice 2008;48(4):178–81.

Additional information and reprint requests:
Fabiano Riva, Ph.D.
Institute of Legal Medicine
University of Bern
Bühlstrasse 20
3012 Bern
Switzerland
E-mail: fabiano.riva@irm.unibe.ch