Northumbria Research Link

Post-refereeing version
Final version published in: Journal of Forensic Sciences, Vol. 60(2), 539-541.
doi:10.1111/1556-4029.12691

Commentary on: Gauriot R, Gunaratnam L, Moroni R, Reinikainen T, Corander R. Statistical challenges in the quantification of gunshot residue evidence. J Forensic Sci 2013;58(5);1149-1155.

Matteo Gallidabino*, Alex Biedermann, Franco Taroni

University of Lausanne, École des Sciences Criminelles, Institut de Police Scientifique, Bâtiment Batochime, 1015 Lausanne-Dorigny, Switzerland

---

Abstract

Commentary on: Gauriot R, Gunaratnam L, Moroni R, Reinikainen T, Corander R. Statistical challenges in the quantification of gunshot residue evidence. J Forensic Sci 2013;58(5);1149-1155.

---

* Corresponding author: matteo.gallidabino@unil.ch

Sir,

we would like to comment on the paper by Gauriot et al. [1], "Statistical challenges in the quantification of gunshot residue evidence". In this paper, the authors deal with the assessment of the evidential value of gunshot residue (GSR) particles using a probabilistic approach. The discussion is partially based on graphical probability networks (i.e., Bayesian networks) inspired by previous literature [2], and invokes various scenarios.

As starting point, we wish to emphasize that we are glad to see that discussions on the forensic interpretation of GSR particles stimulate strong interest among researchers worldwide. This topic is intricate and several concerns are regularly raised during meetings, while rather few works have been published so far. The paper by Gauriot et al. contributes to move discussions ahead by raising questions regarding the general understanding of aspects that affect the coherent evaluation of results of GSR analyses. However, we do not fully agree with all the points addressed by the authors. We would like to revisit some of their conclusions and discuss two points in particular: first, the choice of propositions of interest in a GSR-related case, and second, the difference between subjectivity and arbitrariness. The viewpoint that we seek to justify is that none of these two points represents a drawback for probabilistic approaches to GSR interpretation, contrary to what might be understood from Gauriot et al.'s discourse.

Our first point relates to the definition of propositions of interest and is crucial for the application of probabilistic approaches in any field of forensic science, not only GSR-analysis. It is a subtle step and should be faced with utmost care. It is widely accepted that at least two competing propositions must be considered: one for the prosecution and one for the defence [3]. In previous works on the evaluation of GSR particles [2, 4], propositions of the following kind were used as an example:

H_p: the suspect has discharged a firearm;

H_d: the suspect has not discharged a firearm.

Although it is clear that the simple negation of the first proposition is rarely a helpful alternative (because one may ask what else 'happened' if Hp is *not* the case), the generality of the discussion supposed that, in an actual case at hand, the proposition Hd *will* be further specified in order to enable the scientist to assess probabilities for outcomes given Hd. Note that, without an explicit alternative scenario under Hd it may be difficult to assign probabilities for outcomes in a defensible way. We will come back to this issue in due course.

Gauriot et al. initially adopt the same pair of propositions as mentioned above. However, they consider that these propositions do not allow the scientist to take into account the wide range of possible scenarios and, thus, they argue that other hypotheses should be considered. For example, the defence may consider that the suspect has not discharged a firearm, but he was standing next to the shooter, or touched a GSR-contaminated surface (e.g., the victim or a spent cartridge). To bridge this gap, the authors propose propositions such as "the suspect was or was not contaminated by standing near the shooter" and "the suspect was or was not contaminated by touching a GSR-contaminated surface" (p. 1152). On first sight, such propositions undoubtedly appear intuitively attractive, essentially because they provide a full account of the findings. That is, as their name says, they *explain* the scientist's findings. On close inspection, however, we must realise that they provide no guidance to a jury. The reason for this is that, generally, the probability of the findings given an explanation is the same, or about the same, as that given the first proposition, so that the resulting likelihood ratio is one. This is why such propositions have been recognised as 'explanations', that is allegations that explain the findings but that do not enable the strength of findings to be assessed [5].

The problem with the explanations invoked above is not that they suggest an alternative activity such as standing nearby or touching a surface. The complication arises from the fact of factoring observations into the explanations, in particular the expression 'was contaminated'. To illustrate that this observation-driven feature should not be part of the propositions, consider the case of the smash of a window where the suspect argues that he is not the breaker, but was standing nearby. In such a case, the denominator of the likelihood ratio represents the probability of the findings (which may be *no* glass) assuming that the suspect is not the breaker, but a bystander. This will require the scientist to assess the probability of the findings given additional *uncertain events* of transfer and no transfer as a non-smashing bystander. This is the same logic of reasoning as used for the numerator where the findings are assessed under the proposition that the suspect is the breaker, with an extension to additional conditioning events of transfer and no transfer [6]. Stated otherwise, invoking the alternative proposition that the suspect did *not* commit the activity alleged by the first proposition, but was standing close while it was committed by someone else (the true offender) does not mean that *necessarily* transfer occurred. Transfer is an uncertain event both under the first and the second proposition and its probability needs to be assigned depending on the framework of circumstantial information (e.g., [6]).

This is not to doubt that, in real cases, the primary use of a firearm is often not the only possible source of GSR-like particles on a suspect's hands. Indeed, contamination effects due to proximity to a discharging firearm or handling GSR-contaminated surfaces are well known and documented. Several studies regarding this problem are reported in the literature [7-9]. It should be noted, however, that during an evaluative stage (as opposed to an investigative stage) of a firearm-related crime, the principal interest of competing parties at trial is the role (i.e., activity) of an accused in a well-defined event. Thus, when evaluating results of GSR analyses, propositions should reflect this exigency appropriately. That is, propositions should clearly state what the suspect allegedly did and did not do. Additional special circumstances are part of the conditioning information. Most importantly, it should be reminded that *meaningful* and *operationally helpful* propositions are defined within the framework of circumstances as given by the non-scientific evidence in a case as it has been presented to the scientist [5].

Explanation-based reporting is predominant in current practice. Indeed, most forensic laboratories report the detection of GSR particles on hands of a suspect using statements ascertaining the extent of compatibility of those particles with the direct use of a weapon or contamination due to proximity to a discharging firearm *or* handling GSR-contaminated surfaces (see Schwoeble & Exline [7] for some examples of suggested conclusions). Usually, this approach offers *no* discrimination between such explanations, which is in agreement with the well-accepted view among practitioners that the sole count of GSR particles collected on hands does rarely allow one to arrive at conclusions that are in some way discriminative. This, however, is due to an inherent limitation of the nature of the evidence itself, and not the result of a weakness of the interpretative models. In fact, previous literature about the application of probabilistic approaches to GSR evidence never claimed to overcome this situation. On the contrary, probabilistic evaluative procedures actually provide a rationale for the limited discriminative capacity. In particular, they help to logically structure the reasoning procedure, and enhance the crucial understanding of, for example, the differences between propositions and explications.

This is a first step towards an interpretation of evidence that may be more helpful for the judicial system. Indeed, conclusions using explanations offers *no* guidance to the Court as to how it ought to revise its view regarding the various competing scenarios: the presence of any particles is ‘explained’ away by alternative activities. For example, when finding particles the suggestion is that this does not necessarily mean that the suspect discharged the firearm, while any absence of particles is said *not to be incompatible* with the discharge of a firearm (i.e., no particles may be found even though the suspect discharged a firearm). Inherent in such statements is the well-known, and widely banned, confusing expression ‘consistent with’ [3].

The above impasse is inherent to the investigative perspective that starts by looking at the findings to produce potential explanations. This is different for the evaluative perspective where the propositions are formed prior to looking at the findings, based on the agreed background of non-scientific information, and where any results are subsequently assessed with respect to the existing propositions. The difference between these two settings is fundamental and worthy to be reemphasized: the former (investigation) is about asking 'what are the potential explanations for my findings?' whereas the latter (evaluation) is about asking 'if these are my views about the case prior to considering the evidence, based on the non-scientific evidence, how ought these views be modified in the light of the findings?'. Only the latter leads to an expression of the strength of the evidence, positive, negative or neutral, as the case may be.

More generally, it should also be noted that consideration of further analytical items of information, other that the simple count of GSR-like particles, could help reduce the number of situations where findings offer limited discriminative capacity. Intrinsic particle features such as particle composition and population distribution are still rarely taken into account. For example, compatibility with reference material could lead to more substantial weights of evidence under activity-level propositions, in either direction. Future works on this should thus be encouraged.

Another matter of concern pointed out by Gauriot et al. is the claim that the choices (referred to as 'subjective') made by the forensic expert in building a statistical model and assigning probabilities lead to an inevitable degree of arbitrariness in the conclusions. Regarding this observation, it is useful to emphasise the difference between the terms "arbitrary" and "subjective". As a matter of general understanding, one could term "arbitrary" a judgement that is based solely on random choice or personal whim. In contrast to this, a "subjective" judgement is typically understood as one that is based on (or influenced by) personal feelings, tastes or opinions. The distinction between the two thus relies on how the particular judgement is formed. When it is based on no knowledge of the problem under study (or in intentional ignorance of relevant knowledge), then it is arbitrary. In turn, if it is based on previous experience or learning, then it is subjective in the sense of 'personal'. Stated otherwise, a subjective judgement is an opinion conditioned on and informed by relevant acquired information on the considered topic, while an arbitrary judgement is merely a random unjustified choice. These definitions are mutually exclusive, in the sense that a genuinely subjective judgement is not arbitrary and vice-versa. Thus, if it is said that the conclusions of scientists depend on various component assignments, this does not mean that the overall conclusion is arbitrary. It solely means that the conclusion translates the knowledge of the individual who issues the statement of interest. Different people can have different opinions about the same issue, which translates different states of knowledge.

Generally in science, including forensic science, models are necessarily subjective because, as noted for example by Lindley [10], every inferential model cannot be anything other than an approximation to the complexities of the real world, and judgement to decide how to model a phenomenon is necessarily personal. As anticipated above, different persons have different (personal) backgrounds, which may lead them to give more or less importance to target variables that compose a problem. They may provide different appreciations of dependencies between variables. This should not, however, be a cause for concern [11]. In particular, it does not indicate a 'problem' with the theory (of probability). It actually shows that the framework is capable to capture intersubjective differences. This perception also applies to assignments for the unknown parameters that are employed in probabilistic models. As each evaluator has access to different datasets and records, it is natural to see that the values chosen by scientists vary inter-individually. Notwithstanding, it may be that scientist can agree on at least a range of inter-subjective agreement.

In forensic science at large, an additional difficulty stems from the fact that each case is distinct in its own right, and perpetrated in conditions which may be hard to reproduce or approximate (because of the nature of the crime itself, or because some of this conditions are actually unknown) [12]. Thus, scientists are often called to exert their personal judgement, based on their personal scientific knowledge and documented competence in the particular area of specialisation. This situation was also accepted by the ENFSI best practice manual [13], where Section 9.1.1. states: "[…] there will rarely be any situation where all the information requirements are met, and the quality of what information is available may be very varied. So, there will always be an element of subjectivity in how this is used and what weight should be attached to the different aspects."

The fact that some extent of divergence in individual assignments may exist does not mean that conclusions arrived at through a particular model are necessarily 'invalid'. The best that scientists can do is to ensure that a model structure and adopted probability assignments are justified by their background knowledge. Subsequent probability calculus will only claim to ensure coherence in the manipulation of probabilities already assigned within the specified model. On this point, see also Lindley [10]: "There are therefore two aspects to our study: the construction of the model and the analysis of that model. The latter is essentially automatic; in principle it can be done on a machine. The former requires close contact with reality. To paraphrase and exaggerate de Finetti, think when constructing the model; with it, do not think but leave it to the computer." (p. 303)

For the purpose of illustration, let us reconsider the effect of considerations about the background presence of particles on the likelihood ratio as discussed in the paper of Gauriot et al. As mentioned above, belief in the occurrence of any kind of contamination in a particular case is a subjective judgement that should be evaluated with respect to the available circumstantial information. Arguably, it will be uncontroversial that the scientists' knowledge about the context, in particular characteristics of the suspect (e.g., his habits, professional activity, etc.), will shape their probability assignments for events of "high" or "low" levels of contamination. For example, if the suspect is known to be a regular sport shooter, a scientist may be entitled to retain higher beliefs in the event of a high level of background, than in a case where the suspect is a sedentary person. Such an assignment is neither arbitrary nor ambiguous. If stated sincerely, these beliefs reflect the scientist's current knowledge of the framework of circumstances *as it was presented* and it is weighted by the scientist's experience in the domain. The latter condition constitutes a fundamental point. In fact, GSR-experimented scientists would take into account contextual information on the basis of their documented learning about the factors which could influence background levels. Thus, the final judgement is not a matter of random choice (i.e., arbitrary), but an informed scientific opinion.

From our point of view, Gauriot at al.'s paper leaves the impression that there is a problem with probabilistic models, and hence with probability in general, for the assessment of GSR analyses. We concede that the application of probabilistic analyses to real world problems may indeed be challenging, but is this the 'fault' of probability? We do not think so because probability theory itself says nothing about how it ought to be applied. Instead we believe that the real difficulty lies in the world that surrounds us. Probability only claims to ensure a coherent processing of beliefs according to a structure that is defined by the scientist. Thus, the matter of sound use of probability rests essentially upon the efforts deployed by the scientist.

Today we have innovative approaches – such as Bayesian networks – that allow us to construct probabilistic models at unprecedented levels of complexity. Rather than running these aids down we should emphasise that never before we have been in a better position to bring probability to fruitful applications, both in forensic science and beyond. The very fact that such models may point out that conclusions may be sensible to input assignments shows where further research in probability elicitation is imperative.

References

[1] R. Gauriot, L. Gunaratnam, R. Moroni, T. Reinikainen, J. Corander, Statistical challenges in the quantification of gunshot residue evidence, Journal of forensic science, 58 (2013) 1149.

[2] A. Biedermann, S. Bozza, F. Taroni, Probabilistic evidential assessment of gunshot residue particle evidence, part 1: likelihood ratio calculation and case pre-assessment using Bayesian networks, Forensic science international, 191 (2009) 24.

[3] I.W. Evett, G. Jackson, J.A. Lambert, S. McCrossan, The impact of the principles of evidence interpretation on the structure and content of statements, Science & justice, 40 (2000) 233.

[4] B. Cardinetti, C. Ciampini, S. Abate, C. Marchetti, F. Ferrari, D. Di Tullio, C. D'Onofrio, G. Orlando, L. Gravina, L. Torresi, G. Saporita, A proposal for statistical evaluation of the detection of gunshot residues on a suspect, Scanning, 28 (2006) 142.

[5] I.W. Evett, G. Jackson, J.A. Lambert, More on the hierarchy of propositions: exploring the distinction between explanations and propositions, Science & justice, 40 (2000) 3.

[6] J.M. Curran, T.N. Hicks, J.S. Buckleton, Forensic interpretation of glass evidence, CRC Press, Boca Raton, FL, 2000.

[7] A.J. Schwoeble, D.L. Exline, Current methods in forensic gunshot residue analysis, CRC Press, Boca Raton, FL, 2000.

[8] D.M. Gialamas, E.F. Rhodes, L.A. Sugarman, Officers, their weapons and their hands: an empirical study of GSR on the hands of non-shooting police officers, Journal of forensic sciences, 40 (1995) 1086.

[9] J. Lebiedzik, D.L. Johnson, Rapid search and quantitative analysis of gunshot residue particles in the SEM, Journal of forensic science, 45 (2000) 83.

[10] D.V. Lindley, The philosophy of statistics, The statistician, 49 (2000) 293.

[11] P. Garbolino, Explaining relevance, Cardozo law review, 22 (2001) 1503.

[12] F. Taroni, C.G.G. Aitken, P. Garbolino, De Finetti's subjectivism, the assessment of probabilities and the evaluation of evidence: a commentary for forensic scientists, Science & justice, 41 (2001) 145.

[13] R. Beijer, J. Biegstraaten, J. Andrasko, L. Gunaratnam, L. Niewoehner, Best practice manual in forensic examination of gunshot residue, ENFSI Working Group FIREARMS, 2003.