COMMENTARY

Criminalistics

Using the likelihood ratio in bloodstain pattern analysis

Daniel Attinger PhD | Kris De Brabanter PhD | Christophe Champod PhD

1 Struo LLC, Ames, Iowa, USA
2 Department of Statistics, Iowa State University, Ames, Iowa, USA
3 Department of Industrial Manufacturing & Systems Engineering, Iowa State University, Ames, Iowa, USA
4 Ecole des Sciences Criminelles, Faculty of Law, Criminal Justice and Public Administration, Université de Lausanne, Lausanne, Switzerland

Correspondence
Daniel Attinger, PhD, Struo LLC, 1908 Northwestern ave, Ames, IA 50010, USA. Email: daniel.atinger@gmail.com
Kris De Brabanter, PhD, Department of Statistics, Iowa State University, 2419 Snedecor Hall, 2438 Osborn Dr, Ames, IA 50011-1210, USA. Email: kbrabant@iastate.edu

Funding information
KDB acknowledges funding by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreement 70NANB20H019 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, Duke University, University of California Irvine, University of Virginia, West Virginia University, University of Pennsylvania, Swarthmore College, and University of Nebraska, Lincoln.

Abstract

There is an apparent paradox that the likelihood ratio (LR) approach is an appropriate measure of the weight of evidence when forensic findings have to be evaluated in court, while it is typically not used by bloodstain pattern analysis (BPA) experts. This commentary evaluates how the scope and methods of BPA relate to several types of evaluative propositions and methods to which LRs are applicable. As a result of this evaluation, we show how specificities in scope (BPA being about activities rather than source identification), gaps in the underlying science base, and the reliance on a wide range of methods render the use of LRs in BPA more complex than in some other forensic disciplines. Three directions are identified for BPA research and training, which would facilitate and widen the use of LRs: research in the underlying physics; the development of a culture of data sharing; and the development of training material on the required statistical background. An example of how recent fluid dynamics research in BPA can lead to the use of LR is provided. We conclude that an LR framework is fully applicable to BPA, provided methodic efforts and significant developments occur along the three outlined directions.

KEYWORDS
blood, bloodstain pattern analysis, courtroom testimony, likelihood ratio, statistics

Highlights
• The likelihood ratio is rarely used in bloodstain pattern analysis (BPA), even though this is a useful measure of the weight of forensic evidence.
• This shortcoming is attributed to specificities in scope and methods and to gaps in the underlying science base.
• Three lines of efforts in research and training are recommended to promote the use of likelihood ratios in BPA.
• We describe how to estimate a likelihood ratio relevant to BPA, based on recent fluid dynamics research.

1 INTRODUCTION

In evaluative reporting, there is a current trend to evaluate findings based on the concept of likelihood ratio (LR). Recent guidelines recommending the use of LR have been issued by the UK Association of Forensic Science Providers (AFSP) [1], then adapted by the European Network of Forensic Science Institutes [2], the National Institute of Forensic Science, Australia and New Zealand (NIFS) [3], and recently advised by the UK Forensic Science Regulator (FSR), the UK Charted Society of Forensic Science, and the Royal...
The present contribution has been motivated by observations made in relation to the way BPA findings are reported in court. Current reporting practices in BPA rely on a classic approach [5] where BPA experts indicate that some of their observations are “consistent with” some stated allegations, generally without weighing the strength of their observations. As Evett et al. summed up [6] (p. 237): “At worst, the use of this phrase [consistent with] lays one open to the criticism of partiality. At best, it does nothing to convey an assessment of the weight of evidence in favor of one or other of the stated propositions.”

To the best of our knowledge, attempts to use LRs in BPA are very few and very recent: an oral presentation was given by Leon Meijsink at the 2019 European Conference of the International Association of the Bloodstain Pattern Analysts, with abstract in Ref. [7]; Zadora et al. discussed the construction of LRs based on the interpretation of spectroscopy measurement, toward evaluating the age of dried bloodstains [8–11]. There is also a current lack of guidance on how BPA evaluations should be reported. For instance, the Bloodstain Pattern Analysis Subcommittee of the Organization of Scientific Area Committees (OSAC) for Forensic Science issued in the 2020 ANSI/ASB Standard 031, Standards for Report Writing in Bloodstain Pattern Analysis [12]. The document gives guidance with regards to the formatting aspects of the reports, but none on how the findings should be assessed and reported in the discipline. Similarly, another standard on the validation of procedures, ANSI/ASB 072 “Standard for the validation of procedures in bloodstain pattern analysis” [13], has been reported to be vague and unfit for purpose [14]. In 2020, the UK FSR added to its code of practice and pattern analysis” [13], has been reported to be vague and unfit for purpose [14]. In 2020, the UK FSR added to its code of practice and conduct an appendix on BPA [15] that is also silent on how the findings are evaluated and simply refer to the ANSI/ASB Standard 031.

With this contribution, we would like to provide context and facts to nurture the debate on a possible migration away from sole ipse dixit opinion of the expert to a more structured way of forming and expressing opinions. Since LRs are not necessarily used by BPA experts, this commentary first describes in detail the LR as a scientific approach to evaluating observations used in different domains of forensic science and then explores what the use of LRs would entail for BPA.

When used to evaluate forensic findings, the LR, also known as the Bayes factor, is the ratio between the probabilities of the observations under two competing and mutually exclusive propositions [16]. The propositions typically represent the contrasting allegations made by the parties, whereas the observations are the forensic findings. The LR represents numerically the capability of the observations to discriminate two propositions; its value indicates which proposition is more supported by the observations. The logarithm in base 10 of the LR is also referred as the weight of evidence to be assigned to the observations.

There are significant advantages to use LRs in evaluative forensic opinions in court [17]. It forces forensic practitioners to consider their findings in the perspective of both the prosecution and the defense, rather than, for example, only looking through the prism of the thesis of one side. Doing so ensures a balanced approach in the sense that results are assessed in a way that considers the allegations of both sides. It allows bringing a fair assessment to the court whose main duty is to act as a referee between the views of each party.

Another advantage of the LR approach is the adherence to a logical framework that invites the forensic practitioner to assess the probability of the forensic observations under both propositions, and not the probabilities of the propositions themselves, that being the reserved duty of the court. One of the key logical requirements of forensic testimony is indeed to avoid an error known as “transposing the conditional” [18] and making sure that the forensic practitioners reserve their assessments to the observations and not to the propositions themselves. This requirement does not imply that all practitioners will reach the same LR assignments as they may invoke different knowledge, data, and expertise to make them, but at least their opinions will withstand scrutiny in a formal and logical sense. To achieve consistency in opinions, two mechanisms can be foreseen as identified in the FSR document [4]: a full disclosure of the data used by the practitioner and its limitation (which is important to transparency), and the regular calibration of expert’s judgement through, for example, proficiency testing and collaborative exercises.

Finally, the LR approach strives toward transparency in the sense that the forensic practitioners are invited to describe the data and explain the methods and knowledge that contributed to their assignments of probabilities.

There are certain difficulties to the use of LRs by forensic experts in court. For instance, the use of LRs is grounded on the specification of two mutually exclusive propositions representing the defense and prosecution allegations. This basic requirement is not always met, as, for example, when the defense exercises its right to remain silent. Also, the propositions put forth by either the defense or the prosecution might change during the course of the proceedings.

The availability of at least a pair of propositions (referring respectively to the position of the prosecution and defense) is a prerequisite for evaluative reports, meaning reports that assess the observations toward the use in a court of law. In the situation where there is no constituted defense proposition or where the defense remains silent, the forensic findings are best presented in a report that is more of an investigative nature, offering lines of investigation or possible reconstructions, and not through an evaluative report [19]. In BPA, we recognize that a lot of activities are investigative in nature and will proceed to evaluation only when clarity on the alleged propositions is made. This commentary is primarily concerned with the evaluative testimony. Finally, we note that the defense proposition does not need to be restricted to a single option but may cover multiple alleged activities that will be treated separately or jointly as shown in Ref. [16].

Nevertheless, the LR approach is recommended by the European Network of Forensic Science Institutes [2] for evaluations of forensic
evidence in court and is routinely used for DNA evidence and other types of forensic evidence.

2 | A LIKELIHOOD RATIO CAN BE MORE OR LESS COMPLEX TO ASSIGN

The complexity to establish LRs depends on two main factors:

1. The method and knowledge used to evaluate the forensic observations

As reviewed in Ref. [20], the LR has been used in evaluations by forensic experts of various fields to evaluate the relations between the evidence and a source, either an individual source such as a person of interest in a crime or a set of sources sharing common identities such as a brand and model, such as pens or cameras.

These evaluations are made either for analysis obtained from recovered material traces, where the observations are some physicochemical or biological part of the source left at the scene, or for pattern or signal evidence, which is an impression left either on a physical surface or in a numerical form by the source, where the observations will be some geometrical, structural, or signal characteristics.

For material traces, source-based evaluations have been made based on chemicals or biological analyses of material traces such as ink [21]; gasoline [22] or inflammable liquid [23]; drugs [24]; glass fragments [25], and, of course, DNA [26]. All these source-based evaluations relied on statistically assessing the importance of similar (or dissimilar) features between the observations and reference data [23].

For pattern evidence, source-based evaluations have been developed, for example, in the areas of fingerprints [27], facial images [28], and voice recordings [29] (not a pattern per se but an analogue or digital trace).

Most of these source-based evaluations involve a process where features of the recovered specimens are compared with features or outputs of a known source, such as handwriting compared with the writing of an individual [30], fake identity documents (compared with elements from fake document factory [31]), or digital images associated with a given camera sensor [32].

Some of the above evaluations involve a mathematical process, which quantifies the correspondence of features between evidence and source, and also rely on databases to determine how rare the corresponding features are.

The availability of databases is critical to the assignment of probabilities to the observations. Indeed, [23] mentions that the performance of an LR evaluation depends on several factors, including the scarcity of the databases used as populations; the mismatch in the conditions of the materials in the population databases and in the specimens; and the degraded quality or quantity of the materials. Also, approaches based on LRs need to be validated, via performance assessments, measurements of discriminative power, and calibration [16,23,33]. There are two main components that will impact the LR at source level: the within-source variability and the between-sources variability. DNA evidence is among the fields where assignments of LRs at source level are the most advanced. The within-source variability is almost null (the DNA profile is stable over the lifetime), and the availability of a biological model, coupled with known allele frequency data, makes the LR calculation well understood, standardized, and widely accepted [34]. For some areas (such as glass fragments, fibers, or footwear marks), the knowledge associated within- and between-sources variability is informed by structured data and documented knowledge.

At times, experts will thus rely more on informed judgement than structured datasets, for example, to assess how a feature may evolve over time due to the wear of a shoe. By “informed judgement,” we mean that the expert relies partly on their personal experience, that is, the expert’s training and professional experience in the forensic technique, up to mere opinion (on the extreme right of the axis). The more they base their assignments of probability on relevant, peer-reviewed, publicly available, and robust data, the greater is the trustworthiness of those assignments. The above leads us to suggest a hierarchy of methods and knowledge used by forensic scientists to assess their findings. This hierarchy is reported in Figure 1, which may help compare the methods and scope of BPA with those of other forensic disciplines.

The horizontal axis of Figure 1 is linked to the number of variables, nature of data, and expertise required. On the left side of the horizontal axis, we have knowledge derived from studies, ideally published, and peer-reviewed, where the relevant features have been systematically measured and studied statistically, together with knowledge from models based on sciences such as physics, chemistry, or biology. On the right side of the horizontal axis, specialists will rely on knowledge derived from personal experience, that is, the expert’s training and professional experience in the forensic technique, up to mere opinion (on the extreme right of the axis). The more specialists base their assignments of probability on relevant, peer-reviewed, publicly available, and robust data, the greater is the trustworthiness of those assignments. The more they base their assignments on their recalled experience and knowledge and on their intuition, the more these assignments will be open to justified challenge. Indeed, when more and more
2. The type of evaluative propositions

Applications of LRs in the context of forensic expertise [20] are not only limited to evaluate the relation of the evidence with either its source, but also to the associated activities, or the offense itself. These three types of propositions are what Cook et al. define as the hierarchy of propositions [35], explaining that propositions of lower hierarchy (about source) are easier to weigh than higher ones (about activity or offense). This hierarchy is reflected on the vertical axis of Figure 1. For the first type of evaluative statement, source-based evaluations, it is easier to apply LRs to material traces than patterns [36], because the material is of similar nature as the source (albeit in lower quantity), while patterns involve the printing of features from the source to the evidentiary findings, a transfer process that is not always reproducible. As mentioned in Ref. [37], “In case of crime scene markings created by one object leaving markings of itself on another object—such as a fingerprint onto a surface, a firearm barrel onto a bullet, or teeth onto skin—the faithfulness of the transfer from the original to the receiving surface, and the ability of the receiving surface to retain the impression unchanged, are essential to the probativeness of the comparison of the mark on the receiving surface to a suspected source.”

The second type of evaluative statements where LRs are used involves activity propositions, where it is asked to consider the meaning of the forensic observations in the context of the activities that led to them. Evaluations of activity add to source-level considerations, the necessity to consider parameters such as the transfer and transport (how much material would be exchanged under these alleged actions; how far from the source would material be found), persistency over time (how the material will potentially change, deteriorate, or disappear over time), the efficiency of the recovery methods used to collect the traces, and the presence by chance of the material in the background [35]. Evaluations of activity do not only aim at associating evidence to a source, but also, and often mainly, consider the spatial motion or temporal modification of the evidentiary material. Examples are the assessment of DNA findings considering either a direct or a secondary transfer [38] or the assessment of fibers [39] or glass [40] if a given set of actions occurred. Consideration of activity relies both on databases of experimental cases and on models of the transport, transfer, or persistence phenomena involved, be they physical, chemical, or biological in nature. Again, the models need to be validated against well-designed and available data.

The third type of evaluations relates to judicial matters such as the committed offense, which involve, in general, considerations outside the domain of expertise of the forensic expert [35]. Therefore, Figure 1 represents the increasing difficulties associated with the different methods and propositions evaluated with LRs. The difficulties increase along the diagonal direction, from the “less complex” to the “more complex” realm. Indeed, forensic evaluations become more complex with increasing level in the hierarchy of propositions, with the unavailability of methods, models, and data, and with increasing number of variables that may contribute to the assignments of probabilities. To be clear, the term “less complex” is not used to mean that any forensic evaluation is trivial, but it indicates that in some activities of forensic science, the methods of evaluation are established and well documented, while in others, practitioners have to resort on their personal experience through methods that are less documented in the peer-reviewed literature.

In this commentary, we limit the discussion to the first two types of evaluative statements, to which LRs have been applied: source-based and activity-based.

3 | WHAT ARE THE EVALUATIVE STATEMENTS AND METHODS RELEVANT TO BPA?

Let us consider now what type of evaluative statements are expected from BPA experts, and what methods are used to evaluate these statements. To do so, we review, in Table 1, the goals and methods of BPA and the forensic discipline that examines traces of blood for the purpose of crime scene reconstruction and court
### Table 1: List of topics of investigation and goals of BPA analysis, with methods used and type of evaluative statements

| General BPA investigation topic (and specific names) | Methods used | Typical goal | Desired goal | Type of evaluative statement |
|-----------------------------------------------------|--------------|--------------|--------------|-----------------------------|
| 1. Location of origin of a blood spatter (height of source of drip pattern; determination of area of origin; determination of area of convergence) | Measurement of features; comparison with features known from existing experiments; determination of impact angles; assumption of straight trajectories; inspection of intersecting trajectories | Determination of a location in a 3D space, with a general estimate of the spatial uncertainty | Evaluation of the same, with estimation of the uncertainty specific to the spatter pattern of interest | Activity (transport) |
| 2. Evaluation of motion based on appearance of multiple blood stains (effect of horizontal motion on bloodstain appearance; reconstruction of cast-off patterns) | Inspection of stain size and shapes, and spatial distributions of stains | Determination of the sequence of movements, and the speed, of a bleeding person or bleeding object | Evaluation of the same, with estimation of the uncertainty specific to the spatter pattern of interest | Activity (transport) |
| 3. Classification of bloodstain patterns with respect to their physical cause (identification of the following patterns: low/medium/high velocity spatter; impact patterns, gunshot patterns (forward/backward spatter); spatter patterns; arterial gushing; transfer patterns, cast-off pattern) | Qualitative or quantitative examination of stains (number, size, spatial distribution); comparisons with patterns generated during training of analyst; uses of classification trees [43,86], or artificial intelligence [76] | Opinion on the possible physical mechanism(s) causing the pattern; exclusion of some physical mechanism(s) | Evaluation of possible physical causes, with a score or other quantitative estimate of certainty as, for example, done in figure 3 of Ref. [87] | Activity (cause) |
| 4. Evaluation of the volume of a stained area | Inspection of the stain; comparison with experiments where relevant parameters are systematically explored | Determination of the volume of the blood that generated the stain | Evaluation of the same, with estimation of the uncertainty specific to the situation of interest | Activity (transport) |
| 5. Evaluation of the time when the blood trace was produced [54] | Inspection of the stain; measurement of the ambient conditions; comparison with experiments where relevant parameters are systematically explored; modelling of transport phenomena | Determination of the time when blood was spilled, with estimate of uncertainty specific to situation of interest | Evaluation of the same | Activity (transport) |
testimony. The table of contents of two successful textbooks on BPA helped to prepare Table 1: the book by MacDonell [41] (and the associated laboratory manual [42]), which—while dated—inspired the traditional basic BPA training; and that of a more recent book by Bevel and Gardner [43], a recommended reading for the BPA certification of the International Association for Identification. Table 1 lists five general BPA topics with their associated methods, goals, and type of evaluative statement. Note that the expressed goals are borrowed directly from the above textbooks. However, when adopting a LR-based approach, the purpose is not to “determine”, but to provide a relative assessment of the probabilities associated with the observations (the effects) given the alleged propositions (the causes).

3.1 | BPA evaluations of observations are not about source, but about activities

Considering that source evaluations are the most current and most straightforward applications of the LR in forensic evaluations, we can readily answer why LR is rarely used in BPA: LR is rarely associated with BPA evaluations, because these are essentially activity level reporting, as shown in the last column of Table 1. Source evaluations are not done in BPA: helping to establish the common origin of a bloodstain with its source is a fundamental function of forensic genetics but is carried out outside of the traditional scope of BPA. Thus, the scope of BPA makes it amenable to LR in a more complex way than source-level evaluations, as per Figure 1. This difficulty is also present in the interpretation of patterns considering activity level propositions. Patterns such as footwear or fingerprints have rarely been assessed in the light of activity level propositions, most of the considerations remained about the source. Assigning LRs for patterns considering activity level propositions is an emerging forensic topic [44–46].

Bloodstain pattern analysis indeed focuses on helping towards questions of activity, involving either the determination of the transport—that is, the temporal or spatial evolution—of blood during the activity (how far, in what direction, and how long ago), or the determination of the type of activity (the cause of a blood pattern).

For instance, BPA analysts are trained at associating patterns of millimeter-sized elliptical stains with broad causes such as spatter events. The “spatter” cause is an abstract category, subsuming a wide variety of causes such as spitting blood, stepping in a blood pool, hitting a person with a bullet, and snapping bloody fingers. These spattering events have also completely different characteristics (velocity, impulse, and shape of the impactor) than the resulting distribution of sizes and shapes of the stains in the pattern. Determination of activity based on inspection of stain patterns is thus a process where the observations are associated with an abstract cause, rather than the association of evidence with class characteristics. Class characteristics associate evidence of the kind of blood type, fibers, or paint with a given set of individual sources, as bullets with specific marks can be associated with a gun model. The recommendations for courtroom testimony in Ref. [43] mention that bloodstain pattern analysis is a “class characteristic” process. This definition may apply to a small subset of blood patterns called transfer patterns, which may share characteristics of an object involved in the creation of the blood trace, such as the weaves of a cloth or the ridges of a finger. However, for most blood patterns, the association of a blood pattern with its cause is not a class characteristics process, because there is typically little-to-no comparable features between the blood pattern and its cause.

3.2 | BPA evaluations rely on a wide range of methods and a complex science base that is still being built

Classification of findings with respect to their cause can certainly be done by considering similarities between features of the evidentiary observations and features of known data produced by a given cause. This task is more difficult than classification by similar features between questioned and known samples of the same kind, because establishing the database requires the design of experiments producing samples based on the set of causes of interest. Firing a model of gun (the cause) and collecting the bullets (the findings) might be a relatively simple example of matching an effect to a cause, in the sense that the striaions on the bullet might match striation patterns of the barrel of the weapon, or match the striaions on other bullets fired by the same model of gun. In BPA, the task of designing experiments corresponding to the specific abstract cause of spatter generation will involve generating events involving, for example, somebody coughing blood, firing bullets, simulating the motion of fists, rods hitting blood pools or body parts, at a variety of velocities, and along different orientations and directions [47,48]. To be representative of all the possible outcomes of these causes, the parameters of the experiments need to be chosen according to a scientific knowledge of the associated transport phenomena, so that no significant region of the design parameter space is omitted. This is difficult to do whether the relationship between cause and effect is not well understood, because of, for example, scientific knowledge gaps in the underlying fluid dynamics. In a related way, [36] describes the purpose of BPA in contrast to the identification and classification problems presented above. It describes BPA as an attempt to infer causes from observed evidence: “Forensic examiners in disciplines like crime scene investigation, arson, and blood spatter analysis attempt to reconstruct a crime based on evidence found at the crime scene, which can be viewed as attempts to infer the causes of observed effects. This can be a challenging task because it is difficult to carry out realistic controlled experiments that would allow one to reliably distinguish between competing explanations (e.g., fires that develop naturally versus those using an accelerant). Statistical collaboration with practitioners in relevant disciplines will be valuable in strengthening inferences in these settings.” Similarly, [35] also mentions the inherent difficulty in generating statistical data based on a cause such as “kicking bleeding bodies”: “However, this is not
a straightforward matter to deal with: it is similar to the issues that arise in reconstruction problems, a subject which will be left to a future paper.” Also, for the purpose of peer review, the databases are better made publicly available. Recent studies [49–52] have shown that classification in BPA typically has error rates of about 10%. These numbers are difficult to interpret, because the data used in the studies are not publicly available. Recently, a larger study [53] confirmed the above error rates, while providing more information on the data and methods. Let us stress that a reduction in the error rates by several orders of magnitude is needed to make the related analyses and LRs significant in court. Considering both of these dimensions (level on the hierarchy and nature of the knowledge used), it is clear that assignment of LR in BPA is in the complex realm. Thus, BPA specialists currently borrow more from their personal experience than from structured and published data.

4 | PROPOSITION OF THREE DIRECTIONS TO FACILITATE THE USE OF LRS IN BPA

Despite the inherent difficulty that the scope of BPA is about activities rather than source, it seems possible to develop BPA along the three directions below to facilitate and widen its use of LRs. Note that, a clear benefit of the proposed directions is to root the BPA assessment in a logical and transparent process, a form of evidence-based assessment more than opinion-based.

4.1 | Proposed direction 1: Better understand the physics associated with BPA

The branch of physics describing the complex events connecting the cause of a blood pattern and the pattern itself is called fluid dynamics. It describes how a bullet atomizes blood; how the impact of a fist on a moving face squeezes and splashes blood; how airborne droplets move along trajectories curved by gravity and slowed by drag; and how drops spread, wick, and dry onto materials as complex as fabrics. Only recently, peer-reviewed fluid dynamic studies have started to uncover the complex physics underlying BPA [54–72]. The governing equations of fluid dynamics are called the Navier–Stokes equations; they are the expression of Newton’s second law for fluids. These equations involve nonlinear and dissipative terms. As their name suggests, nonlinear terms such as inertial forces exhibit a nonlinear relationship between cause and effect. Linear extrapolations of uncertainties from experiments under specific conditions to casework involving different conditions may therefore lead to large, unknown errors. Nonlinearity thus makes the design of experiments and databases complicated. Effects of dissipative terms such as viscous dissipation or species diffusion are that information about the initial conditions of a bloody event may be lost during the evolution of the resulting blood pattern. Even the rheology of blood, the description of the fundamental mechanisms by which viscous dissipation occurs, is a topic of current studies [65,73,74]. It is currently not fully clear whether the blood substitutes used for BPA research and training behave like fresh blood [74] in several situations of interest. These difficulties currently limit the ability to reconstruct (going backward in time to identify the causes) a crime scene and to provide corresponding evaluative statements.

For instance, one could argue that information on the average size of the stains would tell whether a spatter is caused by a gunshot or the impact of, for example, a baseball bat. This argument has been made [41] in the past. It has been recently shown that this distinction is not straightforward [75], and even quite inaccurate in situations involving muzzle gases interacting with droplets [76]. Certainly, a deeper knowledge of the physics involved (as per horizontal axis of Figure 1) would improve the ability to associate patterns with their physical cause.

4.2 | Proposed direction 2: Create public databases of BPA patterns, and/or make existing databases public

There is tradition and increased trend in various scientific disciplines to make data publicly, freely, and widely available, but this is not yet the case in BPA. Only few BPA databases are publicly available [53,77–80]. Similarly, lots of peer-reviewed BPA studies produce conclusions on the basis of bloodstain pattern data that are not accessible. The open sharing of data is useful for peer review and necessary to support the determination of rarity and significance of a conclusion. Developing a culture of sharing and publishing BPA data, either scattered, specific, or organized in databases, will contribute to an open science attitude that is much needed in forensic science [81]. This effort may not only involve experimental results from forensic researchers, but also involve casework data from forensic practitioners.

4.3 | Proposed direction 3: Develop BPA training material that discusses the likelihood ratio and its related statistical foundations

The widely used BPA book by James et al. [5] does not mention the concept of likelihood ratio and recommends to perform experiments to demonstrate that traces are “consistent” with an activity, involving blood transfer, swinging a bloody object, timing of the event, a beating, blows, or stab wounds, a given position of a victim raising their arms, or high velocity spatter. Doing so gives BPA the role of supporting one of either the prosecution or the defense theses, without assigning weight to the findings in the context of both propositions. Indeed, the chapter on court testimony in Ref. [5] mentions that experts often influence a verdict and remind the reader that the weight of the evidence is decided by the jury or judge. The uncertainty is presented in a binary manner: “Remember not every question derived from
the bloodstain patterns will have a definitive answer*, rather than along the more realistic probabilistic spectrum. The textbook by Bevel and Gardner does not mention the likelihood ratio either [43]. In their chapter on court testimony, the word uncertainty is not mentioned. The closest is a discussion about the determination of angles of impact from stain inspection, where the authors mention error rates while possibly meaning uncertainty. The chapter also involves a statement mistrusting the legal system and overvaluing the work of the bloodstain pattern analysis: "It is clear and apparent that the legal system as a whole has forgotten that truth is truth."

Approaches based on LRs may take years to propagate into the methods used in crime scene reconstruction and court testimony: the culture of certainty needs to change; new methods need to be designed and validated based on research findings; the associated uncertainties need to be quantified; practitioners need to be trained to the new methods; and standards of accreditation need to account for these changes.

5 | EXAMPLE OF THE USE OF LR IN BPA

Since BPA is about activities involving reconstructed measurements (such as time and location), this creates opportunities to apply LRs in BPA. Remember that the LR applies to any situation where the observations are weighted against each other in the light of two propositions whatever these are. Careful inspection of the first column of Table 1 shows that many topics of BPA investigations are not about quantifying the likelihood of one proposition, but rather about providing a measurement, such as a location (rows #1, 2), speed (#2), volume (#4), or time (#5) and its meaning in the context of the alleged circumstances. The opportunity to use measurements to evaluate LRs can be illustrated by the recent work by Smith et al. [54]. In a laboratory setting, they estimated the time duration \( t \) between the generation of a blood pool and the observation of the drying blood pool via high-resolution photographs. They developed a physical model that estimates \( t \) based on the following evidential observations: time-stamped photographs, with scale, of the shape, size of the blood pool, nature of the surface on which the blood pool resides, and climate measurements. While the study acknowledges that "many questions remain" [54], the proposed estimate is based on a physical model established from transport equations that describe how a complex fluid like blood flows, and how the mass and temperature of the blood pool evolve. While the data and results are currently preliminary, a detailed data collection plan is presented to quantify the uncertainty associated with the estimation of the time. Let us fast forward a few years and assume that these additional developments of the method have been successfully done; that the assumption of normality has been verified; that uncertainty has been quantified; and that the model can be applied to crime scenes. Assume that such model applied to a given crime scene estimates the time \( t = 480 \text{ min} \pm a^*18 \text{ min} \), where we set \( a = 1 \) to obtain a 68.27% confidence interval (based on normality). If the proposition of the prosecution (\( H_p \)) is that \( t < 7 \text{ h} (420 \text{ min}) \), and that of the defense that \( t > 7 \text{ h} (420 \text{ min}) \), it would be easy to express the likelihood ratio of the evidence under either proposition (see also [82]). The LR is

\[
\text{LR} = \frac{P(\text{evidence}|H_p)}{P(\text{evidence}|H_p)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{420} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{1}{2330}.
\]

This value of the LR means that the findings are 2330 times more likely if the defense proposition is true than if the prosecution proposition is true.

Note that a frequentist approach is used in the above illustration about how LRs can be used in BPA. While the use of LRs in forensic science has often been associated with a Bayesian approach to statistics, there is no requirement to assign LRs using a full Bayesian methodology. For instance, the above illustrative example does not require the determination of a prior, the value of which is rarely universally accepted. Besides illustrating how LRs can be used in BPA, the above discussion on Ref. [54] illustrates how the scientific method (here fluid dynamics and experiment design) helps determine the data that need to be collected. Indeed, without knowledge of the intricacies of the complex, multistage, and multidimensional drying process of a complex fluid like blood, a brute force data collection approach might not have identified the relevant data to collect (here the ratio of wet vs. total area of the pool and its perimeter) and might not have been able to subsume those contributions in a single algebraic equation based on physical principles. Note also that, physics-based methods such as the use of dimensionless numbers can reduce the amount of data to collect by several orders of magnitude [66,83].

Similar considerations can be done regarding the determination of the height of a blood source, or its distance from a wall, as long as the reconstructed measurement is provided in a physically and statistically sound way [60,84]. These illustrations show that the LR applies to BPA, at least in principle. Finally, LR approaches may be developed on the sole basis of extensive and representative statistical analysis of findings associated with real casework, such as the survey by Briggs for extraneous blood on clothing [85].

6 | SUMMARY AND CONCLUSION

We conclude that the LR framework is applicable to BPA, but that it is a complex task to do. A structural reason for this complexity is the scope of BPA, which is about evaluations at the level of activities rather than source. A conjective reason is about the state of the art of BPA-related fundamental knowledge, data, and training material. To facilitate applications of LR in BPA evaluations, it is proposed to develop the discipline along three directions: better physical understanding, solid and available data, and appropriate training. As an example relevant to BPA, thoughts are presented on...
how to develop an estimate of the LR of forensic observations given two propositions related to the time of generation of a blood pool, on the basis of existing fluid dynamics modelling and preliminary experiments.

**ORCID**

Daniel Attinger: https://orcid.org/0000-0002-2382-1149

Christophe Champod: https://orcid.org/0000-0002-4035-2698

**REFERENCES**

1. Association of Forensic Science Providers. Standards for the formulation of evaluative forensic science expert opinion. Sci Justice. 2009;49(3):161–4.

2. Willis S, McDermott S, O’Donnell G, Barrett A, Rasmusson B, Hoglund T, et al. ENFSI guideline for evaluative reporting in forensic science. Contract No.: HOME/2010/ISEC/MO/4000001759. 2015. http://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf. Accessed 25 Aug 2021.

3. National Institute of Forensic Science Australia and New Zealand. An introductory guide to evaluative reporting. 2017. https://www.anzpsa.org.au/ArticleDocuments/220/An%20Introductory%20Guide%20to%20Evaluative%20Reporting.PDF.aspx. Accessed 15 Jul 2021.

4. The Forensic Science Regulator in Partnership with the Chartered Society of Forensic Sciences and the Royal Statistical Society. Development of evaluative opinions: codes of practice and conduct for forensic service providers. Contract No.: FSR-C-118. Birmingham, UK: Forensic Science Regulator; 2021. https://www.gov.uk/government/publications/development-of-evaluative-opinions. Accessed 25 Aug 2021.

5. James SH, Kish PE, Sutton TP. Principles of bloodstain pattern analysis: theory and practice. Boca Raton, FL: CRC Press; 2005. p. 431, 435, 443–78, 488, 492.

6. Evett IW, Jackson G, Lambert JA, McCrossan S. The impact of the principles of evidence interpretation on the structure and content of statements. Sci Justice. 2000;40(4):233–9. https://doi.org/10.1016/S1355-0306(00)71993-9

7. Leon Meijrink NFI. A logical framework approach for evaluating BPA evidence in casework. IABPA J Bloodstain Pattern Anal. 2019;34(3):8–9.

8. Menzýk A, Damin A, Martyna A, Alladio E, Vincenti M, Martin G, et al. Toward a novel framework for bloodstains dating by Raman spectroscopy: how to avoid sample photodamage and subsampling errors. Talanta. 2020;209:120565. https://doi.org/10.1016/j.talanta.2019.120565

9. Khandasamy SR, Fikiet MA, Mistek E, Ahmed Y, Halámková L, Bueno J, et al. Bloodstains, paintings, and drugs: Raman spectroscopy applications in forensic science. Forensic Chem. 2018;8:111–33. https://doi.org/10.1016/j.forc.2018.02.002

10. Zadora G, Menzýk A. In the pursuit of the holy grail of forensic science – spectroscopic studies on the estimation of time since deposition of bloodstains. Trends Analyt Chem. 2018;105:137–65. https://doi.org/10.1016/j.talanta.2018.04.009

11. Martyna A, Menzýk A, Damin A, Michalska A, Martin G, Alladio E, et al. Improving discrimination of Raman spectra by optimising preprocessing strategies on the basis of the ability to refine the relationship between variance components. Chemometr Intell Lab Syst. 2020;202:104029. https://doi.org/10.1016/j.chemolab.2020.104029

12. AAFS Standards Board. ANSI/ASB, standard 031 – standard for report writing in bloodstain pattern analysis. 2020. http://www.asbstandardsboard.org/published-documents/bloodstain-pattern-analysispublished-documents/. Accessed 15 Jul 2021.

13. AAFS Standards Board. ANSI/ASB, standard 072 – standard for the validation of procedures in bloodstain pattern analysis. 2019. http://www.asbstandardsboard.org/published-documents/bloodstain-pattern-analysis-published-documents/. Accessed 15 Jul 2021.

14. Morrison GS, Neumann C, Geoghegan PH. Vacuous standards – subversion of the OSAC standards-development process. Forensic Sci Int Synerg. 2020;2:206–9. https://doi.org/10.1016/j.fsisy.2020.06.005

15. Forensic Science Regulator. Codes of practice and conduct – bloodstain pattern analysis. Contract No.: FSR-C-102. Birmingham, UK: Home Office Forensic Science Regulator; 2020. https://www.gov.uk/government/publications/bloodstain-pattern-analysis-codes-of-practice. Accessed 25 Aug 2021.

16. Aitken CCG, Taroni F, Bozza S. Statistics and the evaluation of evidence for forensic scientists. 3rd ed. Chichester: John Wiley & Sons Ltd.; 2020. p. 101–278, 637–54.

17. Robertson B, Vignaux GA, Berger CEH. Interpreting evidence – evaluating forensic science in the courtroom. 2nd ed. Chichester: John Wiley & Sons, Ltd.; 2016. p. 9–67.

18. Evett IW. Avoiding the transposed conditional. Sci Justice. 1995;35(2):127–31.

19. Jackson G, Jones S, Booth G, Champod C, Evett IW. The nature of forensic science opinion – a possible framework to guide thinking and practice in investigations and in court proceedings. Sci Justice. 2006;46(1):33–44. https://doi.org/10.1016/S1355-0306(06)71565-9

20. Morrison GS, Enzinger E. Score based procedures for the calculation of forensic likelihood ratios – scores should take account of both similarity and typicality. Sci Justice. 2018;58(1):47–58. https://doi.org/10.1111/j.1751-5823.2017.00605

21. Neumann C, Margot P. New perspectives in the use of ink evidence in forensic science: Part III: operational applications and evaluation. Forensic Sci Int. 2009;192(1–3):29–42. https://doi.org/10.1016/j.forsciint.2009.07.013

22. Vergeer P, Bolck A, Peschier LJ, Berger CE, Hendrikse JN. Likelihood ratio methods for forensic comparison of evaporated gasoline residues. Sci Justice. 2014;54(6):401–11. https://doi.org/10.1016/j.scijus.2014.04.008

23. Zadora G, Martyna A, Ramos D, Aitken C, Zadora G, Martyna A, et al. Likelihood ratio models for classification problems. Chichester: John Wiley & Sons Ltd; 2014. p. 151–80.

24. Bolck A, Weyermann C, Duypury L, Esseiva P, van den Berg J. Different likelihood ratio approaches to evaluate the strength of evidence of MDMA tablet comparisons. Forensic Sci Int. 2009;191(1–3):42–51. https://doi.org/10.1016/j.forsciint.2009.06.006

25. Zadora G, Ramos D. Evaluation of glass samples for forensic purposes – an application of likelihood ratios and an information–theoretical approach. Chemometr Intell Lab Syst. 2010;102(2):63–83. https://doi.org/10.1016/j.chemolab.2010.03.007

26. Foreman LA, Champod C, Evett IW, Lambert JA, Pope S. Interpreting DNA evidence: a review. Int Stat Rev. 2003;71(3):473–95. https://doi.org/10.1111/j.1751-5823.2003.tb00207.x

27. Champod C, Lennard C, Margot P, Stolovic M. Fingerprints and other ridge skin impressions. 2nd ed. Boca Raton, FL: CRC Press, Taylor and Francis; 2016. p. 33–126.

28. Ali T, Spreeuwers L, Veldhuis R, Meuwly D. Biometric evidence evaluation: an empirical assessment of the effect of different training data. IET Biometrics. 2014;3(4):335–46. https://doi.org/10.1049/iet-bmt.2014.0009

29. Gonzalez-Rodriguez J, Rose P, Ramos D, Toledano DT, Ortega-Garcia J. Emulating DNA: rigorous quantification of evidential weight in transparent and testable forensic speaker recognition. IEEE Trans Audio Speech Lang Process. 2007;15(7):2104–15. https://doi.org/10.1109/taslp.2007.902747
66. Attinger D, Moore C, Donaldson A, Jafari A, Stone HA. Fluid dynamics topics in bloodstain pattern analysis: comparative review and research opportunities. Forensic Sci Int. 2013;231(1–3):375–96. https://doi.org/10.1016/j.forsciint.2013.04.018

67. Kim S, Ma Y, Agrawal P, Attinger D. How important is it to consider target properties and hematocrit in bloodstain pattern analysis? Forensic Sci Int. 2016;266:178–84. https://doi.org/10.1016/j.forsciint.2016.05.015

68. Comiskey PM, Yarin AL, Kim S, Attinger D. Prediction of blood back spatter from a gunshot in bloodstain pattern analysis. Phys Rev Fluids. 2016;1(4):043201. https://doi.org/10.1103/PhysRevFluids.1.043201

69. Comiskey PM, Yarin AL, Attinger D. Hydrodynamics of back spatter by blunt bullet gunshot with a link to bloodstain pattern analysis. Phys Rev Fluids. 2017;2(7):073906. https://doi.org/10.1103/PhysRevFluids.2.073906

70. Comiskey PM, Yarin AL, Attinger D. High-speed video analysis of forward and backward spattered blood droplets. Forensic Sci Int. 2017;276:134–41. https://doi.org/10.1016/j.forsciint.2017.04.016

71. Taylor MC, Laber TL, Epstein BP, Zamzow DS, Baldwin DP. The effect of firearm muzzle gases on the backspatter of blood. Int J Legal Med. 2011;125(5):617–28. https://doi.org/10.1007/s00414-010-0462-4

72. Williams EMP, Graham ES, Jermy MC, Kieser DC, Taylor MC. The dynamics of blood drop release from swinging objects in the creation of cast-off bloodstain patterns. J Forensic Sci. 2019;64(2):413–21. https://doi.org/10.1111/1556-4029.13855

73. Brust M, Schaefer C, Doerr R, Pan L, Garcia M, Arratia PE, et al. Rheology of human blood plasma: viscoelastic versus Newtonian behavior. Phys Rev Lett. 2013;110(7):078305. https://doi.org/10.1103/PhysRevLett.110.078305

74. Lee S-Y, Seo Y-I, Moon B-S, Kim J-P, Goh J-M, Park N-K, et al. Study on development of forensic blood substitute: focusing on bloodstain pattern analysis. Forensic Sci Int. 2020;316:110461. https://doi.org/10.1016/j.forsciint.2020.110461

75. Siu S, Pender J, Springer F, Tulleners F, Ristenpart W. Quantitative differentiation of bloodstain patterns resulting from gunshot and blunt force impacts. J Forensic Sci. 2017;62(5):1166–79. https://doi.org/10.1111/1556-4029.13418

76. Liu Y, Attinger D, De Brabanter K. Automatic classification of bloodstain patterns caused by gunshot and blunt impact at various distances. J Forensic Sci. 2020;65(3):729–43. https://doi.org/10.1111/1556-4029.14262

77. Attinger D, Liu YU, Faflak R, Rao Y, Struttman BA, De Brabanter K, et al. A data set of bloodstain patterns for teaching and research in bloodstain pattern analysis: gunshot backspatters. Data Brief. 2019;22:269–78. https://doi.org/10.1016/j.dib.2018.11.075

78. Attinger D, Liu Y, Bybee T, De Brabanter K. A data set of bloodstain patterns for teaching and research in bloodstain pattern analysis: impact beating spatters. Data Brief. 2018;18:648–54. https://doi.org/10.1016/j.dib.2018.02.070

79. Laber TL, Epstein BP, Taylor MC. High speed digital video analysis of bloodstain pattern formation from common bloodletting mechanisms. IABPA News. 2008;24(2):4–12.

80. Basu N, Bandyopadhyay SK. Initial data release of regular blood drip stain created by varying fall height, angle of impact and source dimension. Data Brief. 2016;8:1194–205. https://doi.org/10.1016/j.dib.2016.07.003

81. Chin J, Grows B, Mellor D. Improving expert evidence: the role of open science and transparency. Ottawa Law Rev. 2019;50(2):365–410.

82. Weyermann C, Ribaux O. Situating forensic traces in time. Sci Justice. 2012;52(2):68–75. https://doi.org/10.1016/j.scijus.2011.09.003

83. Yarin LP. The Pi-Theorem: applications to fluid mechanics and heat and mass transfer. Berlin/Heidelberg: Springer; 2012. p. 1–102.

84. Camana F. Determining the area of convergence in blood-stain pattern analysis: a probabilistic approach. Forensic Sci Int. 2013;231(1–3):131–6. https://doi.org/10.1016/j.forsciint.2013.04.019

85. Briggs TJ. The probative value of bloodstains on clothing. Med Sci Law. 1978;18(2):79–83. https://doi.org/10.1177/002580247801800202

86. Esperança P. Morphoanalyse des traces the sang: une approche methodique [Bloodstain pattern analysis: a methodical approach]. Lausanne: Presses Polytechniques et Universitaires Romandes; 2019.

87. De Brabanter K, Karsmakers P, De Brabanter J, Suykens JAK, De Moor B. Confidence bands for least squares support vector machine classifiers: a regression approach. Pattern Recognit. 2012;45(6):2280–7. https://doi.org/10.1016/j.patcog.2011.11.021

How to cite this article: Attinger D, De Brabanter K, Champod C. Using the likelihood ratio in bloodstain pattern analysis. J Forensic Sci. 2021;00:1–11. https://doi.org/10.1111/1556-4029.14899