Journal Pre-proof

A review of Likelihood Ratios in Forensic Science based on a critique of Stiffelman "No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials"

John Buckleton, Bernard Robertson, James Curran, Charles Berger, Duncan Taylor, Jo-Anne Bright, Tacha Hicks, Simone Gittelson, Ian Evett, Simone Pugh, Graham Jackson, Hannah Kelly, Tim Kalafut, Frederick R. Bieber

PII: S0379-0738(20)30113-4

DOI: https://doi.org/10.1016/j.forsciint.2020.110251

Reference: FSI 110251

To appear in: *Forensic Science International*

Received Date: 31 January 2020

Accepted Date: 11 March 2020

Please cite this article as: Buckleton J, Robertson B, Curran J, Berger C, Taylor D, Bright J-Anne, Hicks T, Gittelson S, Evett I, Pugh S, Jackson G, Kelly H, Kalafut T, Bieber FR, A review of Likelihood Ratios in Forensic Science based on a critique of Stiffelman "No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials", *Forensic Science International* (2020), doi: https://doi.org/10.1016/j.forsciint.2020.110251

A review of Likelihood Ratios in Forensic Science based on a critique of Stiffelman "*No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials*"

John Buckleton[1,2], Bernard Robertson[3], James Curran[1], Charles Berger[4], Duncan Taylor[5,6], Jo-Anne Bright[2*], Tacha Hicks[7,8], Simone Gittelson[9], Ian Evett[10], Simone Pugh[11], Graham Jackson[12], Hannah Kelly[2], Tim Kalafut[13], Frederick R. Bieber[14]

1. Department of Statistics, University of Auckland, New Zealand
2. Institute of Environmental Science and Research, Auckland, New Zealand
3. Barrister, Wellington; Visiting Scholar, Department of Law, Auckland University of Technology, Auckland, New Zealand
4. Leiden University, Institute for Criminal Law and Criminology, PO Box 9520, 2300 RA Leiden, The Netherlands
5. Forensic Science SA, PO Box 2790, Adelaide, SA 5000, Australia
6. School of Biological Sciences, Flinders University, GPO Box 2100 Adelaide SA, Australia 5001
7. Unité de Génétique Forensique, Centre Universitaire Romand de Médecine Légale, Centre Hospitalier Universitaire Vaudois et Université de Lausanne, Lausanne, Switzerland
8. Fondation pour la formation continue Universitaire Lausannoise (UNIL-EPFL), 1015 Dorigny, Switzerland
9. District of Columbia Department of Forensic Sciences, Washington DC, USA
10. Principal Forensic Services Ltd, Bromley, Kent, UK
11. California Department of Justice, Redding, CA, USA
12. Advance Forensic Science, St. Andrews, UK; Abertay University, Bell Street, Dundee, UK
13. Defense Forensic Science Center, Forest Park, GA USA
14. Harvard Medical School, Boston, MA USA

* Corresponding author jo.bright@esr.cri.nz

Highlights

- Recent criticism reprises some objections to the use of *LR*s in forensic science
- We find that *LR*s do not infringe the ultimate issue
- *LR*s do not change the beyond reasonable doubt standard

- *LR*s do not infringe the presumption of innocence
- Propositions should be exhaustive in the context of the case.

Abstract

Stiffelman [1] gives a broad critique of the application of likelihood ratios (*LR*s) in forensic science, in particular their use in probabilistic genotyping (PG) software. These are discussed in this review.

*LR*s do not infringe on the ultimate issue. The Bayesian paradigm clearly separates the role of the scientist from that of the decision makers and distances the scientist from comment on the ultimate and subsidiary issues.

*LR*s do not affect the reasonable doubt standard. Fact finders must still make decisions based on all the evidence and they must do this considering that all evidence, not just that given probabilistically.

*LR*s do not infringe on the presumption of innocence. The presumption of innocence does not equate with a prior probability of zero but simply that the POI is no more likely than anyone else to be the donor.

Propositions need to be exhaustive within the context of the case. That is, propositions deemed relevant by either defense or prosecution which are not fanciful must not be omitted from consideration.

Keywords: Likelihood ratio, reasonable doubt, presumption of innocence, exhaustiveness

Stiffelman [1] gives a broad critique of the application of likelihood ratios (*LR*s) in forensic science, in particular their use in probabilistic genotyping (PG) software. In response, we clarify the true role of likelihood ratios in forensic science, focusing on Stiffelman's main arguments, which are:

1. *LR*s speak to the ultimate issue;
2. *LRs* persuade a jury that they should convict without proof beyond a reasonable doubt;
3. *LR*s infringe on the presumption of innocence;
4. The requirement for propositions to be exhaustive;
5. Traditional DNA evidence is not subject to the same criticisms.

Stiffelman's central argument is that the use of *LRs* is incompatible with the law. Although not stated explicitly, it could be inferred that she favors the prohibition of their use, specifically in relation to DNA mixture evidence. However, the use of *LR*s perfectly matches the Federal Rules of Evidence's definition of relevance (see FRE, 401):

> *""""Relevant* evidence' *means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence."*

*LR*s are a measure of how much more probable or less probable the evidence makes the propositions that are being considered.

*LR*s are based directly on the laws of probability and are integral to a logical framework of probabilistic inference that has extensive application throughout the scientific world. They have a strong claim to being based directly on rules of formal logic. They are used in many areas outside PG, including medical diagnosis testing, speaker identification, parentage analysis, and identification of missing persons and disaster victims. In many cases *LR*s are the only approach that can be logically supported.

Ignoring the strengths of this logical framework would be detrimental to the judicial process, as modern PG software, all of which use *LR*s, offer improvements over previous methods in distinguishing donors from non-donors to DNA mixtures. Arguments that suggest removal of *LR*s from the judicial process are, in our opinion, arguments against a basic precept of law – namely, that evidence-based rational beliefs (e.g., about whether the defendant is a donor to the DNA evidence) should be updated with the presentation of evidence. This is the essence of the framework; the *LR* is simply the updating mechanism.

1. *LR*s speak to the ultimate issue

Perhaps the most succinct statement of Stiffelman's concerns about *LR*s occurs at pg. 146:

> *"These statistics, posing as the likelihood of guilt or innocence, create a trial by mathematics — fundamentally inconsistent with the constitutional norms of the criminal justice system."*

If this statement correctly summarizes her concern, then the author must have some fundamental misunderstanding of *LR*s. This is because "posing as the likelihood of guilt or innocence" is exactly what *LR*s avoid doing.

In the forensic DNA context, *LR*s speak only to the probability of observing the DNA results given certain propositions or hypotheses that are typically subsidiary to the ultimate issue. *LRs* explicitly avoid mention or inference of guilt or innocence, and they do not speak directly about the probability of the propositions themselves. We emphasize that the *LR* is no more than a measure of evidential strength, albeit the best measure of which we are aware. It cannot determine the truth of any proposition, their probability, nor can it determine guilt or innocence; those are matters for the fact finder. Bayes' Theorem[1], which provides the justification for the *LR*, unequivocally makes this distinction.

Evett and Weir [2, 3] outlined principles for evidence interpretation in 1997. Modified slightly, these are:

1. Interpretation of scientific evidence is carried out within a framework of circumstances. The interpretation depends on the structure and content of the framework.
2. Interpretation is only meaningful when two or more competing propositions are considered.
3. The role of the forensic scientist is to assign the probabilities of the evidence given the propositions that are considered.

In the specific context of evidential items containing DNA from a single source, the questions might be: "What is the probability of observing the evidence if the person of interest is the

[1] As all theorems, Bayes' Theorem can be proven from basic laws, in this case the basic laws of probability.

donor?" and "What is the probability of observing the evidence if some unknown person is the donor?" (See also [4] pg. 181-182). Note that these questions ask only about the probability of observing the evidence, and neither ask nor address the probability that a person is or is not the donor.

Bayes' Theorem provides a model that clearly distinguishes the role of the scientist and that of the fact finders [2, 4]. The role of the scientist is to advise the fact finders on the strength of the evidence by computing the *LR*. Any consideration of the prior or posterior odds (or the probability) of the propositions is left to the fact finders.

To confuse the posterior odds with the *LR* is an error termed "the error of the transposed conditional" or "the prosecutor's fallacy" [5]. It was reported as early as 1899 in the Dreyfus case [6] and is still common today. Stiffelman misstates the *LR* in the form of the prosecutor's fallacy numerous times, to wit:

> Pg. 118: *"A likelihood ratio (LR) compares the probabilities of two different hypotheses that seek to explain a given piece of evidence."*

Rephrasing this example to correctly phrase the *LR* and avoid committing the fallacy of the transposed conditional, we could state:

> *"A likelihood ratio compares the probabilities of observing the evidence under two different hypotheses."*

> Pg. 121: *"In this hypothetical, let us assume it is reported that, given the DNA mixture found in the evidence sample, it is 1,000,000 times more probable that the sample originated from Defendant 1 and two unknown unrelated individuals, than three unknown unrelated individuals."*

Properly rephrased:

> *"The DNA mixture observed in the evidence sample is 1,000,000 times more probable if the sample originated from Defendant 1 and two unknown unrelated individuals than if it originated from three unknown unrelated individuals."*

> Pg. 141: *"Thus, the technician proposes that it is the defendant's DNA, and weighs that possibility against the possibility that it is not the defendant's DNA in the mix".*

Properly rephrased:

> *"The scientist weighs the probability of observing the evidence if the defendant's DNA is in the mixture against the probability of observing the evidence if the defendant's DNA is not in the mixture."*

Note that the probability of the propositions (or hypotheses) belong to the court—it is never the place of the scientist to propose *"that it is the defendant's DNA"*, or even to propose the probability that it is the defendant's DNA. This is a profound misstatement of the *LR*.

> Pg. 144: *"LRs are designed to specifically present the relative probabilities that the evidence came from the defendant versus the probability that it came from an unknown unrelated individual."*

Properly rephrased:

> *"LRs are designed to specifically present the relative probabilities of the evidence if it came from the defendant versus if it came from an unknown, unrelated individual."*

There is another serious error embedded in Stiffelman's statement at pg. 146 quoted above in this section. The output from a PG system is an *LR* concerning the probabilities of observing the DNA typing results given different propositions about the origin of the DNA. The results of the DNA comparison are therefore one ingredient to help the court assess whether or not a given person is the source of the DNA. But a DNA comparison – whether made by *LR*s, Random Match Probability (RMP), or any other method – cannot address or answer the question of what activities might have led to the DNA being deposited. DNA evidence cannot be directly connected to guilt or innocence unless the multiple inferential steps needed to make such a connection are accepted by the jury or the court. This aspect is outlined by many authors and is known as the hierarchy of propositions (see for example [7-11]).

Both the prosecutor's fallacy and a misunderstanding of the hierarchy of propositions are serious but common mistakes. We welcome the opportunity presented by Stiffelman's article to discuss these with our legal colleagues.

2. Likelihood ratios persuade a jury that they should convict without proof beyond a reasonable doubt

Stiffelman writes:

> Pg. 139: "*…the introduction of probability evidence in criminal trials invites the jury to render a guilty verdict with proof less than beyond a reasonable doubt.*"
>
> Pg. 140: *"Any time a jury is urged to convict based on an argument that there is a likelihood of guilt to some percentage short of 100%, they are accepting the conviction of a certain percentage of innocent men and women."*

It is certainly correct that, when reasoning with uncertainty, a fact finder in a criminal trial can never be 100% certain of guilt or innocence. This is true when interpreting all evidence, not just evidence that is assessed numerically. Whatever the evidence, the jury members are deciding in the face of uncertainty. Whether or not probability theory is invoked, it must be accepted that the only certain way to avoid convicting even one innocent person is to declare "not guilty" in every case, regardless of how much evidence is presented. If that is not the policy, then society must accept that innocent people will sometimes be convicted. This does not mean that jurors have not reached a state of "beyond reasonable doubt."

The law does not equate "beyond reasonable doubt" with 100% certainty. These points are made by Stiffelman on pg. 140:

> *"We also know that beyond a reasonable doubt does not mean beyond all doubt."*

This seems to logically conflict with her statement on pg. 141:

> *"We don't want a system in which we articulate that it is acceptable if a certain percentage of those convicted are innocent, no matter how small that percentage might be. Of course, we all know that innocent men and women are convicted of crimes. But that cannot be an explicit aim of our criminal justice system."*

We note that recognizing the unavoidable non-zero probability of making a wrong decision, and not striving to reduce the probability of wrong decisions, are two very different philosophical

stances to take. Forensic scientists take the former stance and not the latter since the whole purpose of forensic science is to reduce the probability of wrong decisions by making observations and reporting on their evidential strength. We offer two definitions from law dictionaries of the legal meaning of "beyond reasonable doubt":

From The People's Law Dictionary:

> *Adj. Part of jury instructions in all criminal trials, in which the jurors are told that they can only find the defendant guilty if they are convinced "beyond a reasonable doubt" of his or her guilt. Sometimes referred to as "to a moral certainty," the phrase is fraught with uncertainty as to meaning, but try: "you better be damned sure." By comparison it is meant to be a tougher standard than "preponderance of the evidence" used as a test to give judgment to a plaintiff in a civil (non-criminal) case. (See: reasonable doubt, moral certainty, conviction).*[2]

From Collins Dictionary of Law:

> *The standard of proof in criminal cases in the UK, higher than the civil standard of the BALANCE OF PROBABILITIES. Contrasted with the balance of probabilities, it is not a matter of weighing up both sides and deciding who has won. Thus, if matters are evenly balanced, the accused must be acquitted. Juries when charged are often reminded that they are allowed to have doubts. The doubt must be a real doubt before they acquit -it must not be a fanciful doubt.*[3]

Neither of these definitions imply that "beyond reasonable doubt" is absolute certainty.

We reprise here, importantly, that an *LR* addresses the probability of the evidence, not issues of guilt or innocence. The *LR* in this context applies to a single item of evidence, which might be outweighed by other items of evidence.

3. *LR*s infringe on the presumption of innocence

There seem to be two possible explanations for the belief that *LR*s infringe on the presumption of innocence. One is that the numerator of the *LR* requires the probability of the evidence if the POI is a donor. It appears that Stiffelman interprets this as an assumption that the POI is a donor. A simple example shows that this is not so:

About 3.3% of Californians are surfers[4]. Hence, based on no other information, the probability that you are a surfer if you come from California is about 3.3%. This makes absolutely no statement at all about whether you are from California. In the same way, calculating the probability of the evidence if the POI is or is not a donor makes absolutely no statement at all about the probability that the POI is a donor.

The second possibility is that Stiffelman is not really referring to the *LR* but to the operation of Bayes' Theorem which requires the assessment of prior odds. The term "prior odds" can be used in two ways. If we are considering an item of DNA evidence, the prior odds of the DNA having

---

[2] From https://dictionary.law.com/Default.aspx?selected=59 The People's Law Dictionary

[3] From https://legal-dictionary.thefreedictionary.com/beyond+a+reasonable+doubt Collins Dictionary of Law.

[4] These data are taken from https://www.surfrider.org/coastal-blog/entry/how-many-surfers. We might have different numbers if we were to make a new survey of all the people of California or of only people between 15 and 35 years old. This illustrates another important point: probabilities depend on our knowledge and assumptions.

come from the defendant are the odds assessed based on any evidence presented other than the DNA evidence. "Prior odds" can also refer to the odds assessed before *any* evidence has been considered. Stiffelman seems to be referring to the latter, for example on pg. 138:

> *"The court posited that the presumption of innocence should be represented as zero, which renders the application of Bayes' theorem useless — all posterior probabilities would be zero. Setting the odds to one doesn't work because it presumes an equal chance of guilt and innocence, and zero renders the numbers useless. But these, of course, are not the only options."*

An assignment of a prior probability of zero that a person is a donor to the DNA sample is equivalent to stating that no evidence could ever change your view. Yet, as Stiffelman states, prior odds of either zero or one are certainly not the only options. Collins[5] gives: "If you presume that something is the case, you think that it is the case, although you are not certain." This indeed implies that a presumption of innocence does not necessitate prior odds of either zero or one.

When assigning prior odds, one asks questions of the form "What is the probability that Mr. Smith is a donor to the DNA?" without any consideration of the DNA evidence. These cannot be assessed by the scientist. When considered by the fact finders, it is unlikely that prior odds will be formulated numerically – and there is no requirement that they be numerical.

There is considerable published scholarship on this subject. Some of the best comes from "Interpreting Evidence" [4]:

> "*It is sometimes objected that assessing prior odds (or probability) of guilt in criminal cases 'appears to fly in the face of the presumption of innocence'. Those espousing this view cannot mean that prior odds should be set at zero ... Assigning a probability of zero is to express an unshakeable belief that something cannot be true. The consensus in the common law legal literature is that the presumption of innocence does not entail zero prior odds, but that it is simply a restatement of principles that:*
>
> - *the prosecution must prove the case 'beyond reasonable doubt';*
> - *the prosecution must present a case to be answered before the accused is required to consider offering evidence; and*
> - *the court reaches its decision only on the evidence presented and does not count the fact of arrest and so on against the accused. Before the court had heard any evidence, it should presume nothing simply from the fact that the person has been accused. In that situation, the accused is to be treated like anyone else, because anyone could be accused. In an identity case, for example, the presumption of innocence could sensibly be interpreted to mean that the accused is just as (or no more) likely to be the perpetrator as anyone else."*[6]

---

[5] From https://www.collinsdictionary.com/us/dictionary/english/presume

[6] The statement that "*the accused is to be treated like anyone else"* suggests a prior probability of guilt of n/N, where n is the number of non-assumed donors to the evidence and N is the size of the population of potential perpetrators. In an identity case, one can even start with a prior probability of 1 in 7 billion, which will rapidly be cut down by constraints as to the time and place of the offense, the requirement for the perpetrator to be an able-bodied male, etc. We refer the reader to "Interpreting Evidence[6]"[6] for a valuable section on prior odds and other informative references.

The prior odds needed for the application of Bayes' Theorem represent the fact finders' assessment of all the non-DNA evidence, which could be anything from negligible to substantial, and may support either the case of the prosecutor or of the defense. This would not be assessed at the start of the trial but rather after the presentation of all the non-DNA evidence. Clearly, these prior odds could be anything between zero and very large, depending on the strength of the other evidence presented. This process is not limited to DNA evidence or *LR*s.

In conclusion, contrary to Stiffelman's concerns, there is absolutely no conflict between the use of *LR*s and the presumption of innocence.

4. The requirement for propositions to be exhaustive

> Pg. 119: *"Suppose you returned home to find your dog sitting next to a torn-up pillow, and feathers are everywhere. You could compare the hypothesis that your home was ransacked by burglars with the hypothesis that the pillows you bought at Ikea were designed to explode after six months. After gathering all the relevant data, like that your door was locked when you came home, and nothing else was broken, and the strength of the pillow fibers, you could come up with a likelihood ratio. Let's assume you determine the likelihood ratio that, given the evidence,* ***it was 10,000 times more likely that the pillow just fell apart than it was likely your home was ransacked by burglars****."*

The bold emphasis (added) is yet another instance of Stiffelman's committing the so-called prosecutor's fallacy, as well as an example of a misleading *LR*. For Bayes' Theorem to apply, it is necessary that the two hypotheses being compared are mutually exclusive, but they do not have to be exhaustive. However, if the forensic scientist is given two propositions to consider, and neither are true, it may be that there is missing information that, if known, would lead to the third proposition that is actually true. In Stiffelman's example, this third proposition would be that the dog tore up the pillow. The remedy is straightforward: ensure that all relevant case information is provided to the scientist[7]. When prosecution and defense have given their views of the events, all relevant propositions can be considered, and those propositions can be seen as exhaustive in the context of the case.

More simply stated, the relevant information in a case needs to be considered for any investigation or evidence – including an *LR* – to have meaning. This is the cornerstone of the first principle of evidence evaluation given in the introduction from Evett and Weir [2, 3]. In the example given by Stiffelman, relevant background information would include the knowledge that the homeowner owned a dog that was found sitting beside the torn-up pillow.

As Stiffelman notes on pages 141-142, the defense is not required to provide their theory of events nor to present any case at all:

> *"…the initial problem is that the crime lab proposes the defense hypothesis. Not only does the defendant not need to propose such a hypothesis, but also, there's more than one potential defense hypothesis."*

Stiffelman is correct that, in an adversarial system, the defense is not required to provide their theory of events. The defense should be entitled to all alternatives consistent with exoneration.

[7] As the value of the results depends on the information and propositions considered, scientists are to indicate in their reports what propositions they have considered. A note of caution should be included to alert the reader that if the information changes, the scientist should be contacted, as this will impact the value of the results.

This may encompass very many possibilities. The application of Bayes' Theorem does not require that there are only two propositions; in fact, there can be many. In such cases, the simple odds form of Bayes' Theorem does not apply, but the more general form is straightforward to use. The result is a Bayes' factor which is the multi-proposition analogue of the *LR*.

Taking up Stiffelman's pillow example, we might consider the evidence that there is a torn-up pillow and feathers are everywhere. The background information is that the door was locked when the occupants came home, nothing else was broken, knowledge of the strength of the pillow fibers, and that a dog is present in the house. The prosecution alleges that the house was ransacked by burglars. For some reason, the alternative proposition considered is that the pillows purchased at IKEA were designed to explode after six months. There are no data to indicate that pillows from IKEA explode.

In principle, the defense is entitled to all propositions other than burglary. It is impossible to enumerate these, but a start might be that the pillow was manufactured containing explosives, extraterrestrial aliens entered the house, ants entered the house and are responsible for the damage, or the dog is responsible for the damage. Propositions that are fanciful may be eliminated. The involvement of aliens and the exploding pillow qualify as fanciful, at least in the opinion of the authors. The same goes for the ants because it is highly unlikely that ants would tear apart a pillow even if they entered the house. This leaves the most plausible remaining alternative proposition as the involvement of the dog. Focusing on this most relevant proposition as the alternative to burglary will, in this case, produce a lower *LR* (i.e., providing greater support for the defense) than the Bayes factor produced from a consideration of all propositions. As such, it is conservative and common to approximate the probability of the findings considering multiple propositions as the probability of the findings given the single most significantly contributing proposition.

Many laboratories include a comment in their report inviting the defense (and prosecution) to suggest an alternative hypothesis if they desire. This cannot be done on the stand; rather it must be done ahead of trial to allow reanalysis, re-reporting, and peer review. It must be clear that re-evaluation of the evidence is required whenever a new proposition becomes relevant. This, too, is not specific to an approach which applies Bayes' Theorem.

5. Traditional DNA evidence is not subject to the same criticisms

> Pg. 127: *"The weight of a single-source, gold-standard DNA match is most frequently reported as a Random Match Probability (RMP). ... an RMP expresses the rarity of a certain profile, which makes it in essence a frequency statement of certain genes. It's simply based on the frequency of certain genetic markers.... You'll note, neither of these pieces of evidence, nor the conclusions drawn, require making any assumptions, or postulating any hypotheticals about the suspect's presence in the sample. In other words, the rarity of the profile, is independent of any particular accused."*

We quickly dispense with Stiffelman's error that "*neither of these pieces of evidence, nor the conclusions drawn, require making any assumptions*." In the US, the RMP may be used for single-source DNA profiles or mixtures. It is calculated as the sum of the estimated genotype probabilities for all genotypes that could be the (or one of the) donors to the mixture. What makes the RMP relevant evidence is an implicit alternative hypothesis that the donor is an unknown member of the general population. The RMP then contains the same proposition as the

denominator of the *LR*. Indeed an *LR* for a high template single-source profile is simply the reciprocal value of the RMP.

In conclusion, previous methods such as RMP made assumptions that were in no meaningful way different from those used in assigning an *LR*.

6. Conclusions

*LR*s do not infringe on the ultimate issue of guilt or on subsidiary issues of the source of the DNA or the activity that occurred. In fact, the very purpose of the Bayesian paradigm is to clearly separate the role of the scientist from that of the decision makers and distance the scientist from comment on the ultimate and subsidiary issues.

*LR*s do not affect the reasonable doubt standard. Fact finders must still make decisions based on all the evidence and they must do this considering that all evidence, not just that given probabilistically.

*LR*s do not infringe on the presumption of innocence. *LR*s from PG evidence consider propositions about the source of the DNA. Even then, they do not speak directly to the probabilities of the propositions but rather to the probabilities of the evidence given the propositions.

Propositions do need to be exhaustive within the context of the case. That is, propositions deemed relevant by either defense or prosecution which are not fanciful must not be omitted from consideration.

DNA evidence reported by RMP makes some of the same assumptions (either implicitly or explicitly) as evidence reported by *LR*.

Stiffelman commented: *"When dealing with such small amounts of DNA, there is much greater ambiguity as to how the DNA ended up on the object*", and "*Although the possibility of transfer clearly affects the evidentiary weight of these small DNA samples, this is not accounted for when LRs are reported to a jury."* We agree with the first statement, but certainly not with the second. *LR*s cannot be seen as separate from the propositions they address. If an *LR* only addresses whether an accused or someone else was the donor of the DNA, then indeed this does not account for any ambiguity about the transfer. But when e.g. the defense contests when or through which activity the DNA was transferred, this becomes part of the relevant propositions and will be accounted for by an *LR* addressing those propositions. To tackle such questions at activity level, we use." It is true that minute traces of DNA can sometimes provide weaker evidence than, for instance, a large bloodstain, but this is because of the greater uncertainty of how the DNA ended up on the object, and not because of probabilistic methods. In fact, to tackle such questions at activity level, we need probabilistic models, either numerical or not, for transfer, persistence, and background levels of DNA at least.

It is crucial to distinguish the value of DNA profiling comparisons from statements regarding what activities led to the presence of the DNA, or the ultimate issue from decisions of guilt or innocence. The activity leading to the presence of the DNA can be addressed by propositions and *LR*s, but probabilistic genotyping provides robust information for making inferences about the source of the DNA, not on how the DNA was transferred, nor on the relevance of the evidence for the ultimate issue.

Stiffelman's article presents an opportunity for discourse on issues concerning the presentation of scientific evidence in the form of an *LR* in court. All of these issues have been the subject of scholarly discussion in the past. Nonetheless it is valuable to constructively review these concerns with our legal colleagues.

Acknowledgements

This work was supported in part by grant 2017-DN-BX-K541 from the US National Institute of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of their organizations.

Disclaimer

The opinions or assertions contained in this document are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Department of Justice, or the Department of the Army, or the Department of Defense.

Competing interests

Buckleton, Bright, and Taylor are the developers of STRmix.

References

[1] Stiffelman B. No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials. Berkeley Journal of Criminal Law. 2019;24:110-46.

[2] Evett IW, Weir BS. Interpreting DNA Evidence – Statistical Genetics for Forensic Scientists. Sunderland: Sinauer Associates, Inc.; 1998.

[3] Evett IW, Jackson G, Lambert JA, McCrossan S. The Impact of the Principles of Evidence Interpretation on the Structure and Content of Statements. Science and Justice. 2000;40:233-9.

[4] Robertson B, Vignaux GA, Berger CEH. Interpreting Evidence: Evaluating Forensic Science in the Courtroom, Second Edition: John Wiley & Sons, Ltd; 2016.

[5] Thompson WC, Schumann EL. Interpretation of statistical evidence in criminal trials - The prosecutors fallacy and the defence attorneys fallacy. Law and Human Behavior. 1987;11:167-87.

[6] Taroni F, Champod C, Margot P. Forerunners of Bayesianism in early forensic science. Jurimetrics. 1998;38:183-200.

[7] Gill P, Hicks T, Butler JM, Connolly E, Gusmão L, Kokshoorn B, et al. DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence - Guidelines highlighting the importance of propositions: Part I: evaluation of DNA profiling comparisons given (sub-) source propositions. Forensic Science International: Genetics. 2018;36:189-202.

[8] Cook R, Evett IW, Jackson G, Jones PJ, Lambert JA. A hierarchy of propositions: Deciding which level to address in casework. Science and Justice. 1998;38:231-40.

[9] Cook R, Evett IW, Jackson G, Jones PP, Lambert JA. A model for case assessment and interpretation. Science and Justice. 1998;38:151-6.

[10] Gittelson S, Kalafut T, Myers S, Taylor D, Hicks T, Taroni F, et al. A Practical Guide for the Formulation of Propositions in the Bayesian Approach to DNA Evidence Interpretation in an Adversarial Environment. Journal of Forensic Sciences. 2016;61:186-95.

[11] Evett IW, Jackson G, Lambert JA. More on the hierarchy of propositions: exploring the distinction between explanations and propositions. Science and Justice. 2000;40:3-10.