

## Accepted Manuscript

Title: A response to “Likelihood ratio as weight of evidence: A closer look” by Lund and Iyer

Authors: Simone Gittelson, Charles E.H. Berger, Graham Jackson, Ian W. Evett, Christophe Champod, Bernard Robertson, James M. Curran, Duncan Taylor, Bruce S. Weir, Michael D. Coble, John S. Buckleton



PII: S0379-0738(18)30275-5  
DOI: <https://doi.org/10.1016/j.forsciint.2018.05.025>  
Reference: FSI 9319

To appear in: *FSI*

Received date: 9-4-2018  
Accepted date: 15-5-2018

Please cite this article as: Simone Gittelson, Charles E.H.Berger, Graham Jackson, Ian W.Evett, Christophe Champod, Bernard Robertson, James M.Curran, Duncan Taylor, Bruce S.Weir, Michael D.Coble, John S.Buckleton, A response to “Likelihood ratio as weight of evidence: A closer look” by Lund and Iyer, Forensic Science International <https://doi.org/10.1016/j.forsciint.2018.05.025>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## A response to “Likelihood ratio as weight of evidence: A closer look” by Lund and Iyer

Simone Gittelson<sup>1</sup>, Charles E. H. Berger<sup>2</sup>, Graham Jackson<sup>3,4</sup>, Ian W. Evett<sup>5</sup>, Christophe Champod<sup>6</sup>, Bernard Robertson<sup>7</sup>, James M. Curran<sup>8</sup>, Duncan Taylor<sup>9,10</sup>, Bruce S. Weir<sup>11</sup>, Michael D. Coble<sup>12,13</sup>, John S. Buckleton<sup>11,14</sup>

1. Centre for Forensic Science, University of Technology Sydney, P.O. Box 123, Broadway, NSW 2007, Australia
2. Leiden University, Institute for Criminal Law and Criminology, PO Box 9520, 2300 RA Leiden, The Netherlands
3. Consultant Forensic Scientist, Advance Forensic Science, St. Andrews, Scotland, UK
4. School of Science, Engineering and Technology, Abertay University, Dundee, Scotland, UK
5. Principal Forensic Services Ltd., 34 Southborough Road, Bickley, Bromley, Kent, BR1 2EB, UK
6. Ecole des Sciences Criminelles, Faculty of Law, Criminal Justice and Public Administration, Université de Lausanne, Batochime – quartier Sorge, CH-1015 Lausanne-Dorigny, Switzerland
7. Barrister, Wellington, New Zealand
8. Department of Statistics, University of Auckland, PB 92019 Auckland, New Zealand,
9. Forensic Science South Australia, 21 Divett Place, Adelaide, SA 5000, Australia
10. School of Biological Sciences, Flinders University, GPO Box 2100 Adelaide SA, Australia 5001
11. University of Washington, Department of Biostatistics, Seattle, WA 98195, United States
12. National Institute of Standards and Technology, Applied Genetics Group, 100 Bureau Drive, Gaithersburg, MD 20899, United States
13. Center for Human Identification, University of North Texas Health Science Center, 3500 Camp Bowie Blvd, Fort Worth, TX 76107, United States
14. Institute of Environmental Science and Research Limited, Private Bag 92021, Auckland, 1142 New Zealand

### Highlights

- All agree that *LRs* shall not be imposed on others; this is not current practice.
- Presenting both an *LR*, and the basis for it, is current best practice.
- *LRs* should not only be assigned where adequate empirical information is available.
- Even when an opinion is purely subjective, it should be in the form of an *LR*.
- The *LR* is the single most informative summary of evidential weight.

### Abstract

Recently, Lund and Iyer (L&I) raised an argument regarding the use of likelihood ratios in court. In our view, their argument is based on a lack of understanding of the paradigm. L&I argue that the decision maker should not accept the expert’s likelihood ratio without further consideration. This is agreed by all parties. In normal practice, there is often considerable and proper exploration in court of the basis for any probabilistic statement. We conclude that L&I argue against a practice that does not exist and which no one advocates. Further we conclude that the most informative summary of evidential weight is the likelihood ratio. We state that this is the summary that should be presented to a court in every scientific assessment of evidential weight with supporting information about how it was constructed and on what it was based.

### Keywords

Forensic evidence interpretation, evidential weight, LR, Bayesian approach, Bayes’ theorem

### Introduction

Peer-reviewed publications in statistics (e.g., [1-10]), law (e.g., [11-20]) and forensic science (e.g., [21-30]) have extensively discussed the evaluation and presentation of evidence to a court for over 40 years.

Although we may not yet be the majority, the consensus in the forensic statistics, forensic science, and academic legal communities has moved towards a Bayesian approach, and in particular, the assignment of the likelihood ratio ( $LR$ ). Authoritative publications (e.g., [31-47]) and a lengthy list of key quotes and references in Supplementary Material 1 demonstrate this.

Lund & Iyer [48] (hereafter “L&I”) raised the following argument:

Consider the equation: Posterior odds<sub>DM</sub> =  $LR_{DM}$  x prior odds<sub>DM</sub> (1)

where DM stands for the decision maker (this is L&I Equation 1). They suggest that substitution of the expert’s likelihood ratio,  $LR_{Expert}$ , for  $LR_{DM}$  (this is L&I Equation 2) has “no basis in Bayesian decision theory...” (L&I page 3). They further suggest, “it is necessary to conduct an uncertainty evaluation regarding the potential difference between  $LR_{DM}$  and  $LR_{Expert}$ , requiring consideration of the range of results attainable under a wide-ranging and explicitly defined class of models” (L&I page 3). They then go on and perform a sensitivity analysis only on  $LR_{Expert}$ , and not on any difference between  $LR_{DM}$  and  $LR_{Expert}$ . This argument, if valid, would seem to apply to any information passed from an expert witness to the fact finders. This would include any probabilistic statement whether  $LR$ -based or otherwise.

We are concerned about potential retrogression of interpretation science, with powerful negative impacts on the effective delivery of forensic science – as a result of an incorrect interpretation of theory. The risk is real. The argument posed by L&I was forwarded in an Australian case on the 23<sup>rd</sup> of October 2017, 11 days after its appearance: stimulated, no doubt, by somewhat sensationalist press coverage [49].

G.S. Morrison’s response [50] correctly identifies L&I’s main argument against the use of an  $LR$  as a “straw man argument”. In the following sections we discuss this in detail.

### Logical fallacy: L&I’s straw man argument

L&I state on page 3:

“Our paper explicitly identifies the swap from Eq. (1) to Eq. (2) as having no basis in Bayesian decision theory”

and begin their “List of Concerns” on page 6 with:

“The recommendation that an individual substitute someone else’s  $LR$  for his or her own, as represented in Eq. (2), is indefensible, rather than normative, under the subjective Bayesian paradigm.”

These statements are significantly misleading. No one advocating the Bayesian approach to evidence evaluation has ever argued against the self-evident truth that a trier of fact updates their prior odds with their own personal  $LR$  to obtain their own personal posterior odds according to Eq. (1). L&I’s suggestion of the contrary with the words “swap” and “recommendation” in their paper is a wrong interpretation of theory: this is their straw man argument.<sup>1</sup> Contrary to their assertions, we completely agree with L&I that their Eq. 2 is indefensible. However, to our knowledge, no one has ever suggested this swap, let alone recommended it. L&I do not give any reference for such a recommendation.

If our previous writings did not explicitly address this implausible view, we explicitly state here:

To update their prior odds to their posterior odds, a trier of fact must assign their own  $LR$ .

However, it is very unlikely that triers of fact think explicitly in terms of prior and posterior odds. To that extent, the argument and concern over fact finders’ use of an  $LR$  is more academic and theoretical than practical.

---

<sup>1</sup> A straw man argument is a logical fallacy where a person attacks a claim or argument that was never made by the opposing party.

## Misconception about how the trier of fact assigns their likelihood ratio

On page 6, L&I state:

“However, there are many reasons why an  $LR$  value offered by the expert may differ from that of the DM.”

Yes, of course, the  $LR$  of the DM may differ from the  $LR$  of the expert! This statement does not recognise that this is a standard and legally enshrined aspect of the fact-finding process. As a consequence of their straw man argument (see previous section), L&I focus on the question “whose likelihood ratio?”, whereas the question of interest is “how does the trier of fact assign their own likelihood ratio?”

The trier of fact cannot assign their likelihood ratio  $LR_{DM}$  for the evidence without the expert witness’s testimony on this evidence. In the fact-finding process, a forensic scientist is an expert witness that possesses the required knowledge for evaluating the scientific results in forensic science. It is the forensic scientist’s legal obligation to present to the trier of fact an assessment of the scientific results with regard to the issue of interest to the court. This value of the evidence takes the form of an  $LR$ .<sup>2,3</sup> It is broadly accepted that there is no other statistical value that represents the value of the evidence (or weight of evidence) with regard to two competing propositions either at all, or as well. The forensic scientist assigns her  $LR_{Expert}$  using the most appropriate methods and data available given her expert knowledge, results of validation studies, and the relevant contextual information available to her (this includes the collection of scenarios that are being considered by the trier of fact). The fact-finding process unfolds as follows:

1. The forensic scientist informs the trier of fact of the scientific results and what these results mean with regard to the issue of interest to the court. The latter information is summarized as the  $LR_{Expert}$ . It is the forensic scientist’s job to explain her  $LR_{Expert}$  to the trier of fact. This explanation includes, but is not limited to, the propositions considered, her assignment of probabilities of the observations given the truth of these propositions, the methods used, the robustness of these methods, the validation of these methods, the scientific principles underlying these methods, the peer-reviewed literature on these methods, the scientist’s qualifications, her performance in organised assessments of competence, and the laboratory’s accreditation.
2. The forensic scientist is cross-examined. The forensic scientist responds to counter views by other scientists, and may be required to consider the probability of the evidence given further propositions at the direction of the court.
3. The judge or jury form their personal view which might be visualised as a likelihood ratio,  $LR_{DM}$ , although it is certainly not this formal. The jury accepts the scientist’s likelihood ratio  $LR_{Expert}$ , rejects it, or modifies it as their own. Hence, the jury has the choice of believing, not believing, or partially believing the expert witness.

The jury members are not obliged to use Bayes’ theorem. What is important is that the jury members will have had the benefit of hearing an explanation of the pertinent expert considerations in arriving at a balanced assessment of the probative value of the evidence.

At the end of their Discussion section, L&I conclude (page 22):

“More broadly, objective descriptions of procedures followed and outcomes obtained throughout investigation of the case and broader experience may present a promising path to ensuring transferability of information from a forensic scientist to DMs.”

<sup>2</sup> Note that in statistics this value is called a Bayes factor. In forensic science, the Bayes factor is commonly referred to as the likelihood ratio ( $LR$ ). To facilitate understanding in the forensic community, we use the term “likelihood ratio” in this paper.

<sup>3</sup> The term *weight of evidence* dates back to Alan Turing, who defined the *weight of evidence* as the logarithm (base 10) of the likelihood ratio ([51] p. 63).

Far from being a novel vision of the fact-finding process, this no more than a description of the current state of affairs.

### Misconception about the formulation of propositions

In a section entitled “Whose Scenarios?” L&I state (page 6):

“...the *LR* still depends upon the collection of scenarios that are considered as well as the corresponding weights given to them by the DM, neither of which is known to the expert.”

and

“The question remains: How sensitive is the *LR* value to any particular definition of a relevant population?”

Formulating a relevant pair of propositions is an essential step in the process of evaluating the forensic scientist’s results. A relevant pair of propositions is one that represents the collection of scenarios and the population considered by the DM at an appropriate level in the hierarchy of propositions [52, 53] that the forensic scientist can help with (e.g., [54-56]). For the past 20 years, the theory that is known as *Case Assessment and Interpretation (CAI)* [53, 57-62] has emphasized the importance of communicating the information required by the expert to formulate a relevant pair of propositions. Ideally the collection of scenarios and the population considered by the DM are determined in discussion with the prosecution and defence.<sup>4</sup>

An expert’s *LR* value for a pair of source level propositions depends on the definition of the relevant population, which in its turn depends on the alternative proposition. This is not a question that remains. This is the reason why our publications explain how to formulate propositions, and emphasize the importance of formulating a relevant pair of propositions (e.g., [55, 56, 63-68]).

### Misconception about the Bayesian definition of a probability function

On page 6, L&I state:

“...probability functions  $\Pr[y|x, H_j]$  ( $j=0,1,\dots,N$ ) are rarely known in any authoritative sense”

and on page 20:

“...any provided *LR* value would require an accompanying uncertainty statement (...) characterizing the analyst’s belief regarding its deviation from the “true value”, which the Bayesian paradigm defines as the *LR* value a given DM would arrive at following careful review of the complete body of evidence considered by the expert.”

These statements are incorrect from a Bayesian perspective. In the Bayesian approach, probability functions are not “known”. A probability function is a description of a state of knowledge, and therefore an *LR* is a description of a state of knowledge (e.g., [69, 70]). There is no such thing as a “true value” of the *LR*. There is  $LR_{\text{expert}}$  that is based on the expert’s state of knowledge and  $LR_{\text{DM}}$  that is based on the DM’s state of knowledge.

The truth lies in the propositions: either the prosecution proposition is true or the defence proposition is true. But the truth or otherwise of each of these is not something for the expert to consider. It is not the expert’s role to make statements about the propositions (e.g., [71] p. 27). It is important to emphasize that the Bayesian definition of a probability function does not make the expert’s probability functions arbitrary.

---

<sup>4</sup> Ideally, the communication of this information takes place before the expert’s testimony, yet unfortunately this communication sometimes only occurs in court.

On the contrary, calibration of *LRs* (e.g., [72, 73]) and extensive validation studies (e.g., [46, 74, 75]) demonstrate the appropriateness of the numerical values of *LRs* reported by experts.

### **Confusion between the weight of evidence and the communication of the weight of evidence**

L&I state (page 2):

“Practitioners adhering to Bayesian principles appear to consider likelihood ratio to be the only logical approach for expert communication...”

on page 20:

“...our concerns apply to any framework motivating the use of an *LR* as a means for experts to communicate their findings.”

and on page 23:

“Additionally, we hope the forensic science community comes to view the *LR* as one *possible*, not normative or necessarily optimum, tool for communicating to DMs.”

We carefully separate the concept of value of the evidence (or weight of evidence) from how this value should be communicated to the trier of fact. The former is logically an *LR*. The matter of communication is open for discussion and any suggestions for improvement would be welcomed by the community. Communication of the *LR* to the trier of fact does not require the normative theory of Bayesian theory or Bayesian reasoning.

L&I make the reasonable statement that the trier of fact may need additional information from the expert witness in order to inform her “accept, reject, or partially accept” decision. Forensic scientists are prepared for, and readily are presenting information about the principles and assumptions on which any opinion is based. We agree with Morrison [50]:

“Transparent implementation of the likelihood ratio framework (...) is actually the solution to the problem.”

### **Sensitivity Analysis**

Central to the discussion in L&I is an assessment of sensitivity. In their example, they examine the range of models that may be fitted to one set of data. While doing a sensitivity analysis is good practice, this particular choice often represents a small fraction of the variation in *LRs* that can be obtained.

It is little recognized that there are inevitable judgements in any assignment of probability. If we consider even the simplest situations such as the roll of a die, there are many assumptions. Any realistic problem tends to involve many of these. The logical sequel of this is that any summary statistic, whether an *LR* or not, is based on a number of judgements which have occurred prior to the numerical aspects. There is an understandable, but unhelpful, tendency by metrologists to concentrate on those numerical aspects of uncertainty that they believe that they can assess. The best we can offer the court - and we are only obliged to offer the best evaluation of the evidence available, not some hypothetical and unattainable perfect evaluation of the evidence - is our best assessment of the factors underlying any summary. This value of the evidence is open to examination, often vigorous. This is common ground among all forensic scientists.

### **The L&I proposal is the status quo**

L&I suggest the use of the *LR* combined with a sensitivity analysis. We broaden this from the modelling choices to include all choices and to any type of evidence, whether offered by an *LR* or any other method.

This is normal practice. In testimony, choices and assumptions in the evaluation of evidence are often explored, sometimes in depth.

## Conclusion

L&I try to contribute to the way we analyse the sensitivity of our *LR* to various modelling assumptions and parameters. Almost certainly unwittingly they make a number of false assumptions about the current practice in using *LRs*. They proceed to reject practices that do not exist and suggest an approach that is already in place. This does not only result in a straw man argument, but also risks the rejection of progress made in forensic science.

Their argument supposes that forensic scientists would impose their *LR* on the decision maker (DM). In reality however, the DM will only use the expert's *LR* if they agree or trust the expert to do better than themselves. They might defer to someone more knowledgeable but they are not obliged to do so. The DM can accept, reject, or adapt the expert's *LR*. Free appreciation of evidence by the DM is enshrined in most codes of law.

L&I perpetuate the age-old misconception that *LRs* should be used only where "adequate empirical information is available" and precise calculations can be performed ([48] p. 3 "computing a likelihood ratio"). This risks giving experts that cannot assign an *LR* through calculation the freedom to abandon logic and the basic laws of probability theory. Our position is that even when an opinion is purely qualitative it should logically be based on a careful consideration of the two questions which inform the numerator and denominator of the *LR*. Whether qualitative or quantitative, the *LR* is the key.

It is good to do a sensitivity analysis. But it should be kept in mind that evidence evaluation is generally about order of magnitude and logical assessment, not about computation with many decimal digits. L&I seem keen to demand a level of perfection of anyone using *LRs* that is in stark contrast with the quiet acceptance of out-dated practices that abandon logic, such as categorical conclusions. This is where L&I hit a nerve because their (implied) message risks harm to progress made in recent years. Especially when L&I note that it "is currently being evaluated as a candidate framework for adoption in the United States" ([48] p. 3). *LRs* have been presented in US courts for over 20 years, and with the implementation of probabilistic genotyping software, they are becoming the most common framework for presenting the value of DNA evidence in the United States today. In addition, their framework is increasingly being adopted for presenting the value of other evidence in the United States [76].

## Acknowledgements

J.S. Buckleton, J.M. Curran and B.S. Weir were supported in part by grant number 2017-DN-BX-0136 and J.M. Curran by grant number 2015-DN-BX-K049 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 the U.S. Department of Justice. The authors would also like to thank Johanna Veth and Maarten Kruijver for their valuable comments on a draft of this paper.

## References

- [1] D.V. Lindley, Probability and the law, *Journal of the Royal Statistical Society, Series D (The Statistician)* 26(3) (1977) 203-220.
- [2] D.V. Lindley, A problem in forensic science, *Biometrika* 64(2) (1977) 207-213.
- [3] I.W. Evett, A Quantitative Theory for Interpreting Transfer Evidence in Criminal Cases, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 33(1) (1984) 25-32.
- [4] I.W. Evett, P.E. Cage, C.G.G. Aitken, Evaluation of the Likelihood Ratio for Fibre Transfer Evidence in Criminal Cases, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 36 (1987) 174-180.
- [5] C. Neumann, I.W. Evett, J. Skerrett, Quantifying the weight of evidence from a forensic fingerprint comparison: a new paradigm, *Journal of the Royal Statistical Society, Series A (Statistics in Society)* 175(2) (2012) 371-415.
- [6] S. Bozza, F. Taroni, R. Marquis, M. Schmittbuhl, Probabilistic evaluation of handwriting evidence: likelihood ratio for authorship, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 57(3) (2008) 329-341.
- [7] S.E. Fienberg, J.B. Kadane, The Presentation of Bayesian Statistical Analyses in Legal Proceedings, *Journal of the Royal Statistical Society, Series D (The Statistician)* 32(1/2) (1983) 88-98.
- [8] D.A. Berry, I.W. Evett, R. Pinchin, Statistical Inference in Crime Investigations Using Deoxyribonucleic Acid Profiling, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 41(3) (1992) 499-531.
- [9] C.G.G. Aitken, D. Lucy, Evaluation of trace evidence in the form of multivariate data, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 53 (2004) 109-122.
- [10] T. Tvedebrink, P.S. Eriksen, H.S. Mogensen, N. Morling, Evaluating the weight of evidence by using quantitative short tandem repeat data in DNA mixtures, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 59(5) (2010) 855-874.
- [11] M.O. Finkelstein, W.B. Fairley, A Bayesian Approach to Identification Evidence, *Harvard Law Review* 83(3) (1970) 489-517.
- [12] R.O. Lempert, Modeling Relevance, *Michigan Law Review* 75(5/6) (1977) 1021-1057.
- [13] D. Kaye, The Laws of Probability and the Law of the Land, *The University of Chicago Law Review* 47(1) (1979) 34-56.
- [14] B. Robertson, G.A. Vignaux, Probability - The Logic of the Law, *Oxford Journal of Legal Studies* 13(4) (1993) 457-478.
- [15] F. Taroni, C. Champod, P. Margot, Forerunners of Bayesianism in early forensic science, *Jurimetrics Journal* 38 (1998) 183-200.
- [16] C.G.G. Aitken, F. Taroni, Fundamentals of statistical evidence - a primer for legal professionals, *International Journal of Evidence and Proof* 12(3) (2008) 181-207.
- [17] S.E. Fienberg, M.J. Schervish, The Relevance of Bayesian Inference for the Presentation of Statistical Evidence and for Legal Decisionmaking, *Boston University Law Review* 66(4) (1986) 771-798.
- [18] M. Redmayne, P. Roberts, C.G.G. Aitken, G. Jackson, Forensic science evidence in question, *Criminal Law Review* 5 (2011) 347-356.
- [19] D.H. Kaye, Likelihoodism, Bayesianism, and a pair of shoes, *Jurimetrics Journal* 53 (2012) 1-9.
- [20] B. Robertson, G.A. Vignaux, C.E.H. Berger, Extending the confusion about Bayes, *Modern Law Review* 74 (2011) 444-455.
- [21] I.W. Evett, What is the Probability that This Blood Came from That Person? A Meaningful Question?, *Journal of the Forensic Science Society* 23 (1983) 35-39.
- [22] I.W. Evett, A Bayesian Approach to the Problem of Interpreting Glass Evidence in Forensic Science Casework, *Journal of the Forensic Science Society* 26 (1986) 3-18.
- [23] I.W. Evett, J.A. Lambert, J.S. Buckleton, A Bayesian approach to interpreting footwear marks in forensic casework, *Science & Justice* 38(4) (1998) 241-247.
- [24] F. Taroni, J.A. Lambert, L.Fereday, D.J. Werrett, Evaluation and presentation of forensic DNA evidence in European laboratories, *Science & Justice* 42(1) (2002) 21-28.
- [25] I.W. Evett, Establishing the evidential value of a small quantity of material found at a crime scene, *Journal of the Forensic Science Society* 33(2) (1993) 83-86.
- [26] D.A. Stoney, Relaxation of the assumption of relevance and an application to one-trace and two-trace problems, *Journal of the Forensic Science Society* 34 (1994) 17-21.



- [27] D.J. Balding, P. Donnelly, Evaluating DNA Profile Evidence When the Suspect Is Identified Through a Database Search, *Journal of Forensic Sciences* 41(4) (1996) 603-607.
- [28] C. Champod, F. Taroni, Bayesian framework for the evaluation of fibre transfer evidence, *Science & Justice* 37 (1997) 75-83.
- [29] B.S. Weir, C.M. Triggs, L. Starling, L.I. Stowell, K.A.J. Walsh, J. Buckleton, Interpreting DNA mixtures, *Journal of Forensic Sciences* 42(2) (1997) 213-222.
- [30] J. Gonzalez-Rodriguez, J. Fierrez-Aguilar, D. Ramos-Castro, J. Ortega-Garcia, Bayesian analysis of fingerprint, face and signature evidences with automatic biometric systems, *Forensic Science International* 155 (2005) 126-140.
- [31] National Research Council (US) Committee on DNA Forensic Science: An Update, *The Evaluation of Forensic DNA Evidence*, National Academies Press, Washington D.C., 1996.
- [32] A. Carracedo, W. Bär, P. Lincoln, W. Mayr, N. Morling, B. Olaisen, P. Schneider, B. Budowle, B. Brinkmann, P. Gill, M. Holland, G. Tully, M. Wilson, DNA Commission of the International Society for Forensic Genetics (ISFG): Guidelines for mitochondrial DNA typing, *Forensic Science International* 110 (2000) 79-85.
- [33] G. Tully, W. Bär, B. Brinkmann, A. Carracedo, P. Gill, N. Morling, W. Parson, P. Schneider, Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles, *Forensic Science International* 124 (2001) 83-91.
- [34] N. Morling, R.W. Allen, A. Carracedo, H. Geada, F. Guidet, C. Hallenberg, W. Martin, W.R. Mayr, B. Olaisen, V.L. Pascali, P.M. Schneider, Paternity Testing Commission of the International Society of Forensic Genetics: Recommendations on genetic investigations in paternity cases, *Forensic Science International* 129 (2002) 148-157.
- [35] P. Gill, C.H. Brenner, J.S. Buckleton, A. Carracedo, M. Krawczak, W.R. Mayr, N. Morling, M. Prinz, P.M. Schneider, B.S. Weir, DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures, *Forensic Science International* 160 (2006) 90-101.
- [36] M. Prinz, A. Carracedo, W.R. Mayr, N. Morling, T.J. Parsons, A. Sajantila, R. Scheithauer, H. Schmitter, P.M. Schneider, DNA Commission of the International Society for Forensic Genetics (ISFG): Recommendations regarding the role of forensic genetics for disaster victim identification (DVI), *Forensic Science International: Genetics* 1 (2007) 3-12.
- [37] D.W. Gjerston, C.H. Brenner, M.P. Baur, A. Carracedo, F. Guidet, J.A. Luque, R. Lessig, W.R. Mayr, V.L. Pascali, M. Prinz, P.M. Schneider, N. Morling, ISFG: Recommendations on biostatistics in paternity testing, *Forensic Science International: Genetics* 1 (2007) 223-231.
- [38] Standards for the formulation of evaluative forensic science expert opinion, *Science & Justice* 49 (2009) 161-164.
- [39] C.G.G. Aitken, P. Roberts, G. Jackson, *Fundamentals of Probability and Statistical Evidence in Criminal Proceedings: Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses*, Royal Statistical Society, United Kingdom, 2010.
- [40] I.W. Evett, C.G.G. Aitken, C.E.H. Berger, J.S. Buckleton, C. Champod, J. Curran, A.P. Dawid, P. Gill, J. Gonzalez-Rodriguez, G. Jackson, A. Kloosterman, T. Lovelock, D. Lucy, P. Margot, L. McKenna, D. Meuwly, C. Neumann, N.N. Daeid, A. Nordgaard, R. Puch-Solis, B. Rasmusson, M. Redmayne, P. Roberts, B. Robertson, C. Roux, M.J. Sjerps, F. Taroni, T. Tjin-A-Tsoi, G.A. Vignaux, S.M. Willis, G. Zadora, Expressing evaluative opinions: A position statement, *Science & Justice* 51 (2011) 1-2.
- [41] A. Linacre, L. Gusmão, W. Hecht, A.P. Hellmann, W.R. Mayr, W. Parson, M. Prinz, P.M. Schneider, N. Morling, ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations, *Forensic Science International: Genetics* 5 (2011) 501-505.
- [42] P. Gill, L. Gusmao, H. Haned, W.R. Mayr, N. Morling, W. Parson, L. Prieto, M. Prinz, H. Schneider, P.M. Schneider, B.S. Weir, DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods, *Forensic Science International: Genetics* 6(6) (2012) 679-688.
- [43] W. Parson, L. Gusmão, D.R. Hares, J.A. Irwin, W.R. Mayr, N. Morling, E. Pokorak, M. Prinz, A. Salas, P.M. Schneider, T.J. Parsons, DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing, *Forensic Science International: Genetics* 13 (2014) 134-142.
- [44] S.M. Willis, C.G.G. Aitken, A. Barrett, C.E.H. Berger, A. Biedermann, C. Champod, T.N. Hicks, J. Lucena-Molina, L. Lunt, S. McDermott, L. McKenna, A. Nordgaard, G. O'Donnell, B. Rasmusson, M.J. Sjerps, F. Taroni, G. Zadora, *ENFSI Guideline for Evaluative Reporting in Forensic Science*, European Network of Forensic Science Institutes (ENFSI), Dublin, 2015.
- [45] G. Jackson, C.G.G. Aitken, P. Roberts, *Case Assessment and Interpretation of Expert Evidence: Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses*, Royal Statistical Society, United Kingdom, 2015.
- [46] D. Meuwly, D. Ramos, R. Haraksim, A guideline for the validation of likelihood ratio methods used for forensic evidence evaluation, *Forensic Science International* 276 (2017) 142-153.

- [47] K. Ballantyne, J. Bunford, B. Found, D. Neville, D. Taylor, G. Wevers, D. Catoggio, An Introductory Guide to Evaluative Reporting, National Institute of Forensic Science Australia New Zealand, 2017.
- [48] S.P. Lund, H. Iyer, Likelihood ratio as weight of evidence: A closer look, *Journal of Research of National Institute of Standards and Technology* 122(Article Number 27) (2017).
- [49] The Queen v. Clinton James Tuite, in: C.J. Supreme Court of Victoria (Ed.) S CR 2014 0007, Melbourne, Australia, 2017.
- [50] G.S. Morrison, A response to: "NIST experts urge caution in use of courtroom evidence presentation method", 2017.
- [51] I.J. Good, *Probability and the Weighing of Evidence*, Charles Griffin & Co. Ltd., London, 1950.
- [52] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, *Science & Justice* 38(4) (1998) 231-239.
- [53] I.W. Evett, P.D. Gill, G. Jackson, J. Whitaker, C. Champod, Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks, *Journal of Forensic Sciences* 47(3) (2002) 520-530.
- [54] I.W. Evett, G. Jackson, J.A. Lambert, More on the hierarchy of propositions: exploring the distinction between explanations and propositions, *Science & Justice* 40(1) (2000) 3-10.
- [55] B. Robertson, G.A. Vignaux, C.E.H. Berger, *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*, John Wiley & Sons, Ltd, Chichester, 2016.
- [56] S. Gittelsohn, T. Kalafut, S. Myers, D. Taylor, T. Hicks, F. Taroni, I.W. Evett, J.-A. Bright, J. Buckleton, A Practical Guide for the Formulation of Propositions in the Bayesian Approach to DNA Evidence Interpretation in an Adversarial Environment, *Journal of Forensic Sciences* 61(1) (2016) 186-195.
- [57] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A model for case assessment and interpretation, *Science & Justice* 38(3) (1998) 151-156.
- [58] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: Deciding which level to address in casework, *Science & Justice* 38(4) (1998) 231-240.
- [59] I.W. Evett, G. Jackson, J.A. Lambert, More on the hierarchy of propositions: exploring the distinction between explanations and propositions, *Science & Justice* 40(1) (2000) 3 - 10.
- [60] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, Case pre-assessment and review in a two-way transfer case, *Science & Justice* 39 (1999) 103-111.
- [61] G. Jackson, The scientist and the scales of justice, *Science & Justice* 40(2) (2000) 81-85.
- [62] 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(4) (2000) 233-239.
- [63] J. Buckleton, J.-A. Bright, D. Taylor, I.W. Evett, T. Hicks, G. Jackson, J.M. Curran, Helping formulate propositions in forensic DNA analysis, *Science & Justice* 54(4) (2014) 258-261.
- [64] A. Biedermann, T. Hicks, The Importance of Critically Examining the Level of Propositions When Evaluating Forensic DNA Results, *Frontiers in Genetics* 7 (2016) 8.
- [65] A. Biedermann, C. Champod, G. Jackson, P. Gill, D. Taylor, J. Butler, N. Morling, T. Hicks, J. Vuille, F. Taroni, Evaluation of Forensic DNA Traces When Propositions of Interest Relate to Activities: Analysis and Discussion of Recurrent Concerns, *Frontiers in Genetics* 7 (2016) 215.
- [66] J. Buckleton, J.-A. Bright, D. Taylor, *Forensic DNA Evidence Interpretation*, Second ed., CRC Press, Boca Raton, 2016.
- [67] C.G.G. Aitken, F. Taroni, *Statistics and the Evaluation of Evidence for Forensic Scientists*, 2nd ed., John Wiley & Sons Ltd, Chichester, 2004.
- [68] C. Champod, I.W. Evett, G. Jackson, Establishing the most appropriate databases for addressing source level propositions, *Science & Justice* 44 (2004) 153-164.
- [69] C.E.H. Berger, K. Slooten, The LR does not exist, *Science & Justice* 56 (2016) 388-391.
- [70] A. Biedermann, S. Bozza, F. Taroni, C.G.G. Aitken, Reframing the debate: A question of probability, not of likelihood ratio, *Science & Justice* 56(5) (2016) 392-396.
- [71] I.W. Evett, B.S. Weir, *Interpreting DNA Evidence*, Sinauer Associates, Sunderland MA, 1998.
- [72] N. Brummer, A. Swart, Bayesian calibration for forensic evidence reporting, INTERSPEECH 2014, 15th Annual Conference of the International Speech Communication Association, Singapore, 2014.
- [73] D. Ramos, J. Gonzalez-Rodriguez, G. Zadora, C.G.G. Aitken, Information-Theoretical Assessment of the Performance of Likelihood Ratio Computation Methods, *Journal of Forensic Sciences* 58(6) (2013) 1503-1518.
- [74] J.-A. Bright, D. Taylor, C. McGovern, S. Cooper, L. Russell, D. Abarino, J. Buckleton, Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles, *Forensic Science International: Genetics* 23 (2016) 226-239.

[75] T.R. Moretti, R.S. Just, S.C. Kehl, L.E. Willis, J. Buckleton, J.-A. Bright, D. Taylor, A.J. Onorato, Internal validation of STRmix™ for the interpretation of single source and mixed DNA profiles, *Forensic Science International: Genetics* 29 (2017) 126-144.

[76] M.J. Saks, J.J. Koehler, The Coming Paradigm Shift in Forensic Identification Science, *Science* 309 (2005) 892-895.

ACCEPTED MANUSCRIPT