Science and Justice 56 (2016) 374–379

Contents lists available at ScienceDirect

Science and Justice

journal homepage: www.elsevier.com/locate/scijus

Position paper

What should a forensic practitioner's likelihood ratio be?☆

Geoffrey Stewart Morrison [a,b,*,1], Ewald Enzinger [a]

[a] Morrison & Enzinger, Independent Forensic Consultants, Vancouver, British Columbia, Canada & Corvallis, Oregon, USA
[b] Department of Linguistics, University of Alberta, Edmonton, Alberta, Canada

ARTICLE INFO

Article history:
Received 11 May 2016
Accepted 12 May 2016

Keywords:
Likelihood ratio
Precision
Reliability
Verbal scale

ABSTRACT

We argue that forensic practitioners should empirically assess and report the precision of their likelihood ratios. Once the practitioner has specified the prosecution and defence hypotheses they have adopted, including the relevant population they have adopted, and has specified the type of measurements they will make, their task is to empirically calculate an estimate of a likelihood ratio which has a true but unknown value. We explicitly reject the competing philosophical position that the forensic practitioner's likelihood ratio should be based on subjective personal probabilities. Estimates of true but unknown values are based on samples and are subject to sampling uncertainty, and it is standard practice to report the degree of precision of such estimates. We discuss the dangers of not reporting precision to the courts, and the problems with an alternative approach which instead reports a verbal expression corresponding to a pre-specified range of likelihood ratio values. Reporting precision as an interval requires an arbitrary choice of coverage, e.g., a 95% or a 99% credible interval. We outline a normative framework which a trier of fact could employ to make non-arbitrary use of the results of forensic practitioners' empirical calculations of likelihood ratios and their precision.

1. Introduction

We assume the reader is familiar with the introduction to the current special issue [1]. The present position paper is a contribution to the debate as to whether forensic practitioners should calculate and report the precision of the likelihood ratios which they present to the courts. We discuss differences in philosophical understanding as to the nature of the forensic practitioner's likelihood ratio. We argue that the forensic practitioner's likelihood ratio should be an empirically calculated estimate of a true but unknown value, and that the precision of the estimated value should be empirically calculated and reported. We argue against the use of likelihood ratios which are directly the result of subjective judgement. We also argue against the use of multilevel ordinal scales consisting of verbal expressions and associated ranges of likelihood ratio values. Finally, we outline a normative framework which could allow the trier of fact to make principled non-arbitrary use of the reported imprecision of forensic practitioners' likelihood ratios.

2. Different philosophical understandings as to what constitutes the forensic practitioner's likelihood ratio

As outlined in the introduction to the current special issue [1], a number of researchers and practitioners believe that forensic practitioners should empirically calculate and report the precision of the likelihood ratios which they present in court. Papers espousing this position and proposing concrete procedures for calculating precision include Chakraborty et al. [2], Balding [3], Weir [4], Curran et al. [5], Curran [6], Morrison et al. [7], Beecham & Weir [8], Curran & Buckleton [9], Morrison et al. [10]; Morrison [11], Nordgaard & Höglund [12], Hancock et al. [13], Alberink et al. [14], Zhang et al. [15], Taylor et al. [16], Kaplan Damary et al. [17]. Others, for example Taroni et al. [18,19], argue that it is not appropriate to calculate the precision of a likelihood ratio since a likelihood ratio is a personal belief which has no true value to be estimated. We find Sjerps et al. [20] counterarguments to the latter position convincing (see also Kaplan Damary et al. [17]), and we believe that the likelihood ratio which a forensic practitioner presents to the court should not be a personal belief. Once the forensic practitioner has stated what they understand to be the relevant circumstances of the case and the prosecution and defence hypotheses they have adopted, including the definition of the relevant population, and they have stated the type of measurements they will make on the known-origin sample, the questioned-origin specimen, and a sample of the relevant population, their task is to calculate an estimate of a likelihood ratio which has a true (but unknown) value. Without these specifications there is no true likelihood ratio value to be estimated. If one were to change

☆ This paper is part of the Virtual Special Issue entitled: Measuring and Reporting the Precision of Forensic Likelihood Ratios, [http://www.sciencedirect.com/science/journal/13550306/vsi], Guest Edited by G. S. Morrison.

* Corresponding author.
E-mail address: geoff-morrison@forensic-evaluation.net (G.S. Morrison).

1 This author is also the Guest Editor for the special issue. The present paper was written before the Guest Editor handled any of the other submissions, and the Editor in Chief had editorial responsibility for the present paper.

http://dx.doi.org/10.1016/j.scijus.2016.05.007

the specifications then the question being asked and answered would change. A likelihood ratio is the answer to a specific question, and changing the question will lead to a different likelihood ratio with a different true value. For example, a likelihood ratio based on the widths of objects will not only have a different value compared to a likelihood ratio based on the lengths of objects, it will be the answer to a different question.[2]

If a forensic practitioner's task were to assign personal probabilities to arrive at a likelihood ratio which does not depend on specifications including the types of measurements to be made, such a likelihood ratio could not be said to be an estimate of a true but unknown value. In contrast, we believe that the task of the practitioner is to calculate an estimate of the true but unknown value of the likelihood ratio given a specified relevant population and one or more specified types of measurement. This disagreement in belief as to what constitutes the forensic practitioner's likelihood ratio is a philosophical one. It cannot be resolved empirically. We believe that this philosophical disagreement is the fundamental source of dispute between those who believe one should calculate the precision of likelihood ratios and those who believe that one should not. Although both groups may use the term "likelihood ratio", they are referring to different things.

The strict terminological issue could perhaps be resolved by maintaining the distinction between the terms "likelihood ratio" and "Bayes factor": A likelihood ratio is calculated using sample data and maximum likelihood estimates of parameters, whereas a Bayes factor is calculated using posterior predictive probability distributions derived by combining sample data with prior probability distributions and integrating over nuisance parameters. The debate could therefore be rephrased as: Should forensic practitioners calculate estimates of likelihood ratios and their precision versus should they calculate Bayes factors? (The wider debate would also include a third option: Should forensic practitioners assign likelihood ratio values based directly on subjective judgement rather than on explicit calculations using quantitative data?) The term "likelihood ratio" is, however, often used as a cover for both the more restrictive definition of likelihood ratio and for Bayes factor (we will generally use it as a cover term in the present paper). Also, note that the debate is not actually about whether to use frequentist calculations based only on sample data, or Bayesian calculations that take account of both sample data and prior probability distributions. Both sides can advocate Bayesian approaches, albeit different Bayesian approaches. Sjerps et al. [19] argue in favour of calculating and reporting Bayesian posterior probability distributions rather than point-value Bayes factors, whereas Brümmer & Swart [21] argue in favour of the latter.

3. Why it is important to calculate and report the precision of the forensic practitioner's likelihood ratio

If one believes that the forensic practitioner's task is to calculate an estimate of a likelihood ratio which has a true but unknown value, since those calculations will be based on sample data, not oracular knowledge of population distributions, the calculated likelihood ratio value will be influenced by sampling uncertainty. Since the calculated value has uncertainty it should be accompanied by a measure of that uncertainty, i.e., a measure of its precision. This is standard practice (e.g., International Organization for Standardization [22], United Kingdom Accreditation Service [23]) and has been explicitly recommended for forensic science (e.g., National Research Council [24]).

Not to report the precision of a likelihood ratio value could be highly misleading to the court. For example, if the forensic practitioner's best estimate of the likelihood ratio is 10, but this estimate is produced by a system (an implementation of a method) which when tested under conditions reflecting those of the case is found to have a 95% credible interval of plus or minus two orders of magnitude, then the 95% credible interval in this case would be 1/10 to 1000. Knowing this, a judge at an admissibility hearing might reasonably decide that the evidence is not sufficiently reliable to warrant admission, or the trier of fact might reasonably decide that the likelihood ratio value is not substantially different from 1 and that the evidence should therefore not change their belief as to the relative probabilities of the prosecution and defence hypotheses.[3] Even if the forensic practitioner's best estimate of the likelihood ratio were relatively extreme compared to the bound of its credible interval closest to the neutral value of 1, e.g., a likelihood ratio of 1000 with a 95% credible interval of plus or minus one order of magnitude and a lower bound of this interval at 100, the trier of fact might choose to be conservative, and, based on the reported precision, they might use a likelihood ratio value closer to 1 than the forensic practitioner's best estimate, e.g., they may choose to use 100 rather than 1000. If the forensic practitioner did not report precision, then the court would be denied the information necessary to make a reasonable decision on admissibility or on what might constitute a reasonable degree of conservatism. A trier of fact might then be misled into assigning a more extreme strength of evidence to the forensic practitioner's likelihood ratio conclusion than they would have done had they known about the precision of the system used to calculate that likelihood ratio. Alternatively, a trier of fact may be incredulous as to the apparent degree of precision of a likelihood ratio reported as a point value. This might lead them to use a more conservative strength of evidence than if they had actually been presented with the results of an empirical assessment of the precision of the system, e.g., they could choose to use a likelihood ratio of 10 when the forensic practitioner's best estimate was 1000 and the unreported lower limit of the 95% credible interval would have been 100.

The discussion in the preceding paragraph raises a distinction to be made between the likelihood ratio reported by the forensic practitioner and the likelihood ratio actually used by the trier of fact. These will not necessarily have the same value. The effective likelihood ratio that the trier of fact employs, i.e., the extent by which they update their beliefs with respect to the relative probabilities of the competing prosecution and defence hypotheses, will likely depend on the trier of fact's assessment of how much they trust what the forensic practitioner reports. For example, if a practitioner reports a likelihood ratio of 1000 and their appearance and manner instil confidence, the trier of fact might use an effective likelihood ratio of 1000, but if the practitioner's appearance and manner do not instil confidence the trier of fact might be less trustful of what the practitioner reports and use an effective likelihood ratio of 100 instead. Supplying the trier of fact with empirical information about the precision of the system used to calculate the forensic practitioner's likelihood ratio would hopefully lead to the trier of fact choosing their effective likelihood ratio on a more relevant basis than the forensic practitioner's appearance and manner.

[2] For this reason, it is important that the forensic practitioner clearly explain to the judge at an admissibility hearing and to the trier of fact at trial what the question is that the forensic practitioner has set out to address. Only with a clear understanding of the question can the judge or trier of fact decide whether the forensic practitioner has set out to answer a question that is of interest to the court, and only with a clear understanding of the question can the trier of fact understand the forensic practitioner's answer to that question.

[3] The *Daubert* ruling states that "The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate" [William Daubert et al. v Merrell Dow Pharmaceuticals Inc., 1993, 509 US 579, Section II C]. We think, however, that the value of the strength of evidence could play a role in deciding upon admissibility. When the likelihood ratio is 10 and the 95% credible interval is plus or minus two orders of magnitude, the testimony could be ruled inadmissible because if the likelihood ratio is not substantially different from 1 it may be that it will not "help the trier of fact… to determine a fact in issue" [Federal Rule of Evidence 702(a) as amended Apr. 17, 2000, eff. Dec. 1, 2000; Apr. 26, 2011, eff. Dec. 1, 2011]. In contrast, this would not be of concern if the value of the likelihood ratio were extreme (far from 1) compared to the width of its credible interval. Note that our discussion relates to the absolute value of the likelihood ratio compared to its precision, not to the absolute value alone. A likelihood ratio of 2 with a high degree of precision could still help the trier of fact to determine a fact in issue.