



## **American Statistical Association Position on Statistical Statements for Forensic Evidence**

Presented under the guidance of the ASA Forensic Science Advisory Committee \*

January 2, 2019

### **Overview**

The American Statistical Association (ASA) has supported efforts to strengthen the inferential foundations that enable forensic scientists to report scientifically valid conclusions. This document focuses on how statements and opinions should be reported to accurately convey the strengths and limitations of forensic findings to ensure their scientific value is used appropriately in a legal context.

The ASA encourages the continued research necessary to provide the requisite scientific data to support rigorous inferential processes. This document does not advocate any particular method of statistical inference be adopted for all evidence.

The ASA believes that a statistical foundation is essential for identifying acceptable statements and reporting practices in forensic science because it provides a useful framework for assessing and expressing uncertainty. A statistical foundation supplies a set of principles for drawing conclusions from data and for expressing the risks of certain types of errors in measurements and conclusions. This framework applies throughout forensic science, but the discussion that follows is of special relevance to pattern, impression, and trace evidence. As the National Research Council Committee on Identifying the Needs of the Forensic Sciences Community emphasized, it is necessary to ascertain and describe uncertainties in measurements and inference.

This document aims to provide advice to forensic practitioners, but the intent is that the recommendations herein apply more broadly; i.e., to any persons providing (statistical) statements or advice regarding forensic evidence.

This document presents background information and views on the following question: When forensic science practitioners present the results of forensic examinations, tests, or measurements in reports or

---

\* This document is adapted from a document developed by a subcommittee of the National Commission for Forensic Science. The contributing authors of the subcommittee document were Matt Redle (subcommittee co-chair), Judge Jed S. Rakoff (subcommittee co-chair), Stephen E. Fienberg (task group chair), Alicia Carriquiry, Karen Kafadar, David H. Kaye, Peter Neufeld, Charlotte J. Word, and Paula H. Wulff. The members on the 2018 ASA Advisory Committee on Statistics and Forensic Science are: Karen Kafadar (committee chair), Hal Stern (vice-chair), Maria Cuellar, James Curran, Mark Lancaster, Cedric Neumann, Christopher Saunders, Bruce Weir, and Sandy Zabell.

testimony, what types of quantitative or qualitative statements should they provide, to ensure that their conclusions or opinions account for the accuracy and the uncertainty in the measurements or observations, and that they appropriately convey the strength of their findings within the context of the questions of interest for the judicial system? This document refers to such statements as “statistical statements.” Measurement precision, weight of evidence (the extent to which measurements or observations support specific hypotheses), and risks of incorrect conclusions are examples of such statistical statements.

For many types of evidence, forensic science practitioners are asked to determine the source of an item recovered during an investigation (often called a “questioned” sample). A source may be defined as an object, person, process, or location that generated the questioned sample. To achieve this task, practitioners focus primarily on ascertaining the presence/absence of corresponding features between the questioned sample and control samples from one or more known sources. They traditionally decide whether an association between the questioned sample and the control sample(s) is positive (often referred to as a “match”, an “inclusion” or “being consistent with”, or “having indistinguishable features”) or negative (an “exclusion” or “having sufficient differences”).

The traditional approach to the decision process, that supports the forming of opinions by forensic scientists, suffers from important shortcomings. Firstly, the terminology used to report conclusions is ambiguous: different people may understand it differently. For example, for some, “match” might imply that the features of two samples are indistinguishable, while, for others, it might imply that the questioned sample definitely originated from the same source as the control sample. Providing an opinion on the origin of a questioned sample from the observation that it is indistinguishable from a control sample requires knowledge of how common or rare the association is, based on empirical data linked to the case at hand. For example, glass fragments that are deemed to have “similar” trace element compositions provide some evidence that the fragments could have a common origin, but data on the specificity of the trace element compositions in a relevant population are needed to address whether the fragments came from a specific source.

To evaluate the weight of any set of observations made on questioned and control samples, it is necessary to relate the probability of making these observations if the samples came from the same source to the probability of making these observations if the questioned sample came from another source in a relevant population of potential sources.

Secondly, forensic conclusions often rely on the personal impressions of forensic science practitioners, and these are not sufficient for determining a level of uncertainty in measurements. Practitioners support a preference for a given hypothesis on the source of a questioned sample by using their subjective judgment. This judgment is developed through individual training and experience, or by reference to limited empirical studies of the reliability of the judgments of their peers. For instance, if examiners say they are 90 percent certain that a piece of evidence comes from a particular source and are asked to justify this level of certainty, they may cite years of experience. As such, this subjective certainty level reflects their impression of evidence encountered throughout their careers. However, although training and experience are important in applying valid techniques, practitioners’ subjective opinions are not sufficient for establishing the uncertainty in measurements or inferences.

Finally, forensic science practitioners do not currently make statistical assessments explicitly, but they may nevertheless present their findings in a manner that connotes a statistical assessment. For example, unless some lesser degree of confidence is provided, the statement that “the latent print comes from the defendant’s thumb” suggests certainty regarding the source of the latent print. Whether presented as an absolute certainty or more tentatively, such conclusions necessarily rest on an understanding of the variability of fingerprint features between multiple impressions of a given finger and on the frequencies of these features in a population.

Statistical statements should rely on: (1) a defined relevant database describing characteristics, images, observed data, or experimental results; (2) a statistical model that describes the process that gives rise to the data; and (3) information on variability and errors in measurements or in statistics or inferences derived from measurements. This information permits a valid statistical statement regarding the probative value of comparisons or computations (e.g., how rare is an observed positive association when two items arise from the same source and when they arise from different sources?).

The ASA recommends that trace, impression, or pattern evidence practitioners follow a valid and reliable process to determine the extent to which evidence supports the hypothesis of an association between a questioned sample and a sample whose source is known (such as a control sample from a person of interest). Reliability and validity should be established via scientific studies that have been subjected to independent scientific scrutiny. See the [Views Document on Technical Merit Evaluation of Forensic Science Methods and Practices](#) (*Adopted at NCFS Meeting #10, June 21, 2016*). Only when the reliability and validity of the process have been studied quantitatively can statements of the uncertainty in measurements and inferences be trustworthy.

Statistical models are most convincing when a scientific understanding of the physical process that generates the features exists. Sufficient knowledge of the process increases the likelihood of the development of a valid mathematical model. This approach has been successful for determining the probability that associations in pre-defined DNA features will exist among different individuals. However, the processes that give rise to variability in other types of trace and pattern evidence (e.g., friction ridge variability on finger tips, or variability in the chemical composition of fibers) may be too complex to describe and model from first principles. Consequently, efforts to provide statistical statements about the degree of support that observations provide in favor or against hypotheses of legal interest should rest on rigorous, large scale statistical studies of the variability of these evidence types. Depending on the assumptions of the researchers, such studies are likely to result in the development of several different models for a given evidence type, and by extension, in different inferences about the same hypotheses.

Statistical calculations used in judicial proceedings should be replicable, given the data and statistical model. Such replication is crucial when observations are largely subjective or when different statistical models are used, as the quantitative summary of the significance of the findings may vary across forensic science practitioners and laboratories. To assist other experts in replicating the statistical quantities that are reported, it is essential to state the measurements and the models or software programs that were used.

At the core of all statistical calculations, there must be data from a relevant population. To be applicable to casework, rigorous empirical studies of the reliability and accuracy of forensic science practitioners' judgments must involve materials and comparisons that are representative of evidence from casework. As noted below, the strength of evidence will depend in part on how common or rare the observations made on the questioned and control samples are in the relevant population. Consequently, it is important that forensic science practitioners clearly specify the relevant population behind the statistical statement. Communicating this information assists the judge in ruling on the admissibility of the evidence and the trier of fact at trial in making proper assessments about the statistical statement.

Any recommendation on presenting explicit probabilities needs to distinguish between two types of probabilities: those based on a statistical model and supported by large empirical studies, and those that characterize the forensic science practitioner's subjective sense of how probable the evidence is under alternative hypotheses. The former probabilities are easier to validate, but it is important to recognize that statistical models are approximations, and, inevitably, there is some uncertainty in the selection of a model. These uncertainties can be difficult to quantify. In light of uncertainties associated with statistical

modeling and intuitive judgments of the significance of similarities, we offer the following views on the presentation of forensic science findings.

### Views of the ASA

1. The documents and testimony reporting the results of forensic science investigations should describe the features of the questioned and known samples (the data), and the process used to determine the level of similarity and/or dissimilarity between those features across multiple samples.
2. No one form of statistical calculation or statement is most appropriate to all forensic evidence comparisons or other inference tasks. Thus, we strongly recommend that forensic science practitioners be able to support, as part of a report and in testimony, the choice used in the specific analysis conducted and the assumptions upon which it was based. When the statistical calculation relies on a specific database, the report should specify which database.
3. A comprehensive report by the forensic scientist should report the limitations and uncertainty associated with measurements, and the inferences that could be drawn from them. This report might include a range of possible values for an estimated quantity, a separate statement regarding errors and uncertainties associated with the analysis of the evidence and resulting opinions, or empirical performance data associated with a chosen statistical model. If the forensic science practitioner has no information on sources of error in measurements and inferences, or has no validation data, the ASA recommends that this fact be stated.
4. The ASA strongly discourages statements to the effect that a specific individual or object is the source of the forensic science evidence. Instead, the ASA recommends that reports and testimony make clear that, even in circumstances involving extremely strong statistical evidence, it is possible that other individuals or objects may possess or have left a similar set of observed features. We also strongly advise forensic science practitioners to confine their evaluative statements to expressions of support for stated hypotheses: e.g., the support for the hypothesis that the samples originate from a common source and support for the hypothesis that they originate from different sources.
5. To explain the value of the data in addressing conclusions about the source of a questioned sample, forensic science reports and testimony containing statistical statements may:
  - a) Refer to relative frequencies of sets of features in a sample of individuals or objects in a relevant population (where the population is sampled and then represented in a reference database). We recommend that forensic science practitioners note the uncertainty associated with using the database frequencies as estimates of the population frequencies of particular features.
  - b) Present estimates of the relative frequency of an observed combination of features in a relevant population based on a probabilistic model that is well grounded in theory and data.
  - c) Present quantitative statements of the degree to which the evidence supports a particular hypothesis on the source of the questioned sample. We recommend that forensic science practitioners note the basis for the quantitative statements and characterize the uncertainties associated with the statements.
  - d) Present the operating characteristics of the system being used to make the classification, decision, or other inference. Examples include sensitivity and specificity as estimated from experiments using data from a relevant population or a designed study for estimating error rates

or calibrating likelihood ratios. If no information, or only limited information, about operating characteristics is known, we recommend that this fact be stated clearly.

6. Currently, not all forensic disciplines can support statistical statements. The trier of fact (or other interested parties) may still find value in knowing what comparisons were made by forensic science practitioners, what they concluded from them and how they reached their conclusions. The ASA recommends that the absence of models and empirical evidence be acknowledged both in testimony and in written reports.