NATIONAL COMMISSION ON FORENSIC SCIENCE

National Institute of Standards and Technology
U.S. Department of Commerce

Views of the Commission
Statistical Statements in Forensic Testimony

Subcommittee	
Reporting and Testimony	
Status	
Final Draft	

Date of Current Version	22/03/17
Approved by Subcommittee	27/03/17
Approved by Commission	[*dd/mm/yy*]

Note: This document reflects the views of the National Commission on Forensic Science and does not necessarily represent the views of the Department of Justice or the National Institute of Standards and Technology. The portion of the document directly labeled "Views of The Commission" represents the formal Views of the Commission. Information beyond that section is provided for context. Views documents do not request specific action by the Attorney General, and thus do not require further action by the Department of Justice upon their approval by the Commission. The National Commission on Forensic Science is a Federal Advisory Committee established by the Department of Justice. For more information, please visit: https://www.justice.gov/ncfs.

Overview

As forensic science moves forward, the Commission supports efforts to make the reporting and testimony of forensic analyses more overtly statistical and quantitative. This document focuses on statements and opinions that can, or should be made, to convey the accuracy of measurements or observations and the interpretation of those findings. We focus specifically on pattern, impression and trace evidence disciplines that presently lack an empirical foundation. The Commission supports continued research and support necessary to provide the requisite scientific data and develop much needed statistical databases. This document does not advocate that any one statistical model need be adopted universally for all purposes.

The Commission advances a preference for a statistical foundation for statements because mathematical analyses provide a useful framework for assessing and expressing uncertainty. Statistics is the science of making inferences and decisions from data when faced with multiple sources of variability and uncertainty. It supplies a set of principles, based on probability, 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 and forensic medicine, but the discussion that follows is of special relevance to pattern, impression, and trace evidence. An explicit statistical foundation for statements is necessary to enable forensic science and medicine providers to assess and express that uncertainty. 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 Views document presents background information and views on the following question: When forensic science and forensic medicine practitioners present the results of forensic science and forensic medicine examinations, tests, or measurements in reports or testimony, what types of quantitative or qualitative statements should they provide to indicate the accuracy of measurements or observations and the significance of these findings? This document refers to such statements as "statistical statements." These statistical statements may describe measurement precision, weight of evidence (the extent to which measurements or observations support specific hypotheses), or the probability of conclusions.

For many types of evidence, forensic science practitioners may not 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. Absolute certainty lies at one extreme of a spectrum of probability. At the other extreme of the probability spectrum is a negation of the defendant as the source of the latent print, known in statistics as the "null hypothesis." A statement that "it is unlikely that the print came from anyone else" suggests an extremely high likelihood that the print came from the defendant based on an assessment of the frequencies of similar features in fingerprints from the same individual and in prints from different individuals. Between absolute certainty and the null hypothesis lie the varying degrees of probability and uncertainty.

Statistical statements should be based on: (1) the existence of a defined relevant database describing characteristics, images, observed data, or experimental results; (2) a statistical model that accurately assesses the strength of the inference in question or describes the process that gives rise to the data linked to the question at hand; (3) information on variability and errors in measurements or in statistics or inferences derived from measurements; and (4) a statistical statement regarding the probative value of any comparisons done or calculations performed (e.g., how rare is an observed positive association when two items arise from the same source and when they arise from different sources?).

Trace, impression, or pattern evidence practitioners should follow a valid and reliable process to determine the extent to which evidence indicates a positive or negative (exclusion) association between the item in question (often called a "questioned" sample or specimen) and a sample whose source is known (such as a reference sample from the defendant). Reliability and external validity should be established via scientific studies that have been the subject of independent scientific scrutiny. See Views Document on Technical Merit Evaluation of Forensic Science Methods and Practices *(Adopted at NCFS Meeting #10 - June 21, 2016).* Principles of statistics hold that only when the reliability and validity of the process have been studied quantitatively can a statistical model for indicating the uncertainty in measurements and inferences be credible statistician.

Statistical models are most convincing when a scientific understanding of the physical process that generates the features exists. Sufficient knowledge of the process leads to 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. For other types of trace and pattern evidence, however, no widely accepted statistical models of the phenomena that give rise to the features are available. Consequently, most efforts to provide statistical statements about features and their degree of association often rest on the personal impressions of forensic science practitioners. This is supported by their subjective judgment developed through individual training and experience, or by reference to empirical studies of the reliability of the judgments of forensic science practitioners. Training and experience are important in applying valid techniques, but they are not a sufficient basis for establishing the uncertainty in measurements or inferences.

When forensic science practitioners do provide a statistical statement—with or without a numerical articulation of probability, odds, or likelihoods—such a statement must be supported by an empirical assessment of the underlying statistical model. Statistical calculations used in judicial proceedings should be replicable, given the data and statistical model; however, when observations are largely subjective or when different statistical models are in use, the quantitative summary of the significance of the findings may vary from forensic science practitioner to forensic science practitioner and from laboratory to laboratory. Consequently, an essential element of a forensic science practitioner's report is a statement of the measurements and the models or software programs applied to assist other experts in replicating the statistical quantities reported.

At the core of all of such statistical calculations, there must be data from a *relevant population* and one needs to report the extent to which the database represents this population. To be applicable to casework, empirical studies of the reliability and accuracy of forensic science practitioners' judgments must involve materials and comparisons that are representative of actual cases and rely on data from a relevant population. As noted below, the strength of evidence will depend in part on how common or rare a measured or observed degree of similarity is in the relevant population. Consequently, it is important that forensic science practitioners clearly specify the relevant population used to derive 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 decision regarding use of the statistical statement.

In comparing forensic evidence recovered from crime scenes or from victims or suspects with known samples, the forensic science practitioner primarily focuses on ascertaining corresponding features and, traditionally, in deciding whether there is a positive association (often referred to as a "match," an "inclusion," or "consistent" or "indistinguishable" features) or a negative one (an "exclusion" or inconsistency) to the known sample. But a "positive association" is not probative unless it is more probable when the items have a common source than when they originate from different sources. Indicating the statistical weight of the positive association therefore requires a statement of how common or rare the association is, based on a database or empirical data linked to the case at hand. For example, a positive association for the presence or absence of pigment in a hair cuticle is some evidence that the hairs could have a common origin, but the significance of this association is unknown without data from relevant populations.

The weight ascribed to any degree of association depends on (1) the probability of the degree of correspondence in the features, given that the samples came from the same source, and (2) the probability for the same measurement, given that the samples came from different sources.

Any recommendation on presenting explicit probabilities, however derived for specific forensic evidence, should distinguish between probabilities based on a statistical model and ones that characterize the forensic science practitioner's subjective sense of how probable the evidence is under alternative hypotheses. The latter are difficult to validate, but it also must be understood that statistical models are approximations, and, inevitably, there is some uncertainty in the selection of a model. In light of the limitations on both statistical modeling and more intuitive judgments of the significance of similarities, we offer the following views on the presentation of forensic science findings:

Views of the Commission

It is the view of the Commission that:

1. Forensic science practitioners, both in their reports and in testimony, should describe the features of the questioned and known samples (the data), and similarities and differences in those features as well as the process used to determine them.

2. No one form of statistical calculation or statement is most appropriate to all forensic evidence comparisons or other inference tasks. Thus, the forensic science practitioner must 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 one.

3. The forensic science practitioner should report the limitations and uncertainty associated with measurements and the inferences that could be drawn from them. This report might take the form of an interval for an estimated value, or of separate statement regarding errors and uncertainties associated with the analysis of the evidence. If the forensic science practitioner has no information on sources of error in measurements and inferences, the forensic science practitioner must state this fact.

4. Forensic science practitioners should not state that a specific individual or object is the source of the forensic science evidence and should make it clear that, even in circumstances involving extremely strong statistical evidence, it is possible that other individuals or objects could possess or have left a similar set of observed features. Forensic science practitioners should confine their evaluative statements to the support that the findings provide for the claim linked to the forensic evidence.

5. To explain the value of the data in addressing conclusions as to the source of a questioned sample, forensic science practitioners may:

   A. Refer to relative frequencies of individual features in a sample of individuals or objects in a relevant population (as sampled and then represented in a reference database). The forensic science practitioner should note the uncertainties in these frequencies as estimates of the frequencies of particular features in the population.

   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. The model may relate the probability of the combination to the probabilities of individual features.

   C. Present statistical model-generated results to determine the possible origin of the questioned sample. The forensic science practitioner should note the uncertainties in any such values.

   D. When the statistical statement is derived from an automated computer-based system for making classifications, the forensic science practitioner should present not only the classification but also the operating characteristics of the system (the sensitivity and specificity of the system as established in relevant experiments using data from a relevant population). If the forensic science practitioner has no information or only limited information about such operating characteristics, the expert must state this fact.

6. Not all forensic sub-disciplines currently can support a probabilistic or statistical statement. There may still be value to the factfinder in learning whatever comparisons the forensic science practitioner in those sub-disciplines has conducted. The absence of models and empirical evidence should be expressed both in testimony and written reports.