3B. Communicating Limitations

Introduction

Results of forensic science examinations are used for making decisions that have significant consequences. Therefore, it is important to understand that all methods of examination have limitations (e.g., measurement uncertainty, variability, possible sources of error). Science is a "work in progress." What is known today may only be applicable to a particular context or population, leaving unanswered questions. Science will continue to evolve using ever-increasing sophisticated technology, and this may change the nature of the limitations, but limitations will still exist. Understanding the limitations of reported results allows the decision-maker to assess the value of the information provided in a specific case. Information on the limitations of a method or a result will allow the comparison of results from different forensic science service providers (FSSPs), the comparison of a result to a legal specification, and the comparison of differing opinions, as well as help the decision maker identify what information or knowledge is not available to them.

Limitations

Although stated slightly differently, both the National Academy of Sciences and the National Commission on Forensic Science support identifying the limitations of methods and results in forensic science reports. Table 1 uses previous primer topics to provide examples of limitations communicated in forensic science reports.

Table 1: Examples of limitations communicated in forensic science reports

TOPIC	LIMITATION	EXAMPLE STATEMENT OR SCENARIO
Accreditation	A reference to accredited status can only be included when the examinations performed are within the scope of accreditation. An FSSP cannot give the impression that they are accredited for work outside the scope of accreditation.	Reference to accredited status in a report **Scenarios** Discipline: Accredited in the disciplines of biology, seized drugs, friction ridge, toxicology, and firearms but not accredited in the discipline of scene investigation Service: Accredited in the discipline of firearms but not accredited to perform trigger pull examination Technology: Accredited in the discipline of toxicology and the use of gas chromatography-mass spectrometry (GC-MS) but not accredited to use liquid chromatography-mass spectrometry (LC-MS)
Sample	Sample acceptance criteria for the method were not met.	**Scenarios** DNA: Quantity of sample is insufficient Latent Print: No print of value for comparison Toxicology: The volume of the sample is insufficient
Method	Quantitative: Limitations on lower and upper limits of quantitation. The compound of interest is present but at a concentration below the method's limit of quantitation (LOQ) or above the method's upper limit of quantitation (ULOQ). Qualitative: The compound of interest is present but at a concentration below the method's limit of detection (LOD) or the FSSP's reporting limit.	**Example** Item 1: Post-mortem blood sample Drug A: None detected at a reporting limit of 5 ng/mL. *Understanding: When tested, the instrument response for Item 1 was less than the instrument response for the calibrator at the decision point of 5 ng/mL.* *There is uncertainty in the instrument response. For a full understanding of the result, the range of possible instrument response values is necessary.* **Example** Item 2: Fire debris sample No ignitable liquid was identified. *Understanding: When tested, there was no instrument response above the method's limit of detection (LOD) for the ignitable liquids included in the method.* *Information on what ignitable liquids are included in the method should be available from the FSSP but may not be provided with the report.* *If there was no instrument response, the instrumentation used in the method may not be sensitive enough to identify an ignitable liquid that is present.*

TOPIC	LIMITATION	EXAMPLE STATEMENT OR SCENARIO
		If there was an instrument response, there is uncertainty in the instrument response. For a full understanding of the result, the range of possible values of instrument response is necessary.
	The method lacks selectivity; it is unable to resolve chemical isomers.	**Example** Item 3: Electronic cigarette cartridge containing light brown liquid Tetrahydrocannabinol (THC) was identified, but the method used in the examination of Item 3 cannot separate delta-8-THC from delta-10-THC. *Understanding: The instrument response for THC could be:* • *100% due to the presence of delta-8* • *100% due to the presence of delta-10* • *Any combination of delta-8 and delta-10*
Statistical Sampling	The confidence level and the proportion of the population to which the interference applies as well as the number of "negative" results that are acceptable (0 – 2) are used to determine the number of items in the population that will be selected and examined by an FSSP. There is a risk of misunderstanding the report statement and mixing up the confidence level and the inference level. Neither the confidence level nor the inference level can be confirmed without examining the entire population.	**Example** Item 1: 100 packages containing white powder Twenty-three packages were examined and found to contain Drug A. The 23 packages were selected using a hypergeometric sampling plan which allows an inference with a 95% confidence level that at least 90% of the 100 packages in Item 1 contain Drug A. *Understanding: There is an estimated 5% risk (95% confidence level) that fewer than 90 (90% inference level) of the packages in Item 1 will contain Drug A.*
Measurement Uncertainty (MU)	There is always MU. The MU may not always be included in a report of a measurement (i.e., weight, length, concentration, force). Accreditation requirements require the reporting of MU when it affects the evaluation of conformity with a specification. If the MU was not included in the report, obtaining the MU will be required to be able to calculate the range of possible results at a stated level of confidence. Different terminology may be used in the report statement: expanded uncertainty, coverage probability, coverage level, level of confidence, confidence interval, and confidence level. The terms confidence interval and confidence level are only appropriate to use when all uncertainty contributors are quantified using data from the measurement process. There is a risk of misunderstanding the report statement.	**Example** Item 1: Ante-mortem blood specimen The specimen was found to have an ethanol concentration of 0.082 g/dL +/- 0.004 g/dL at a 99.7% coverage probability. *Understanding: If one were to repeatedly perform this measurement 1000 times using this method, one would expect the true value to be included in 997 out of 1000 ranges of the measurement result +/- 0.004 g/dL. For this example measurement result, the range is 0.078 to 0.086 g/dL.* **Example** Item 2: Contains white powder The net weight of Item 2 is 28.55 g +/- 0.03 g at a level of confidence of approximately 95% (95.45%). *Understanding: If one were to repeatedly perform this measurement 100 times using this method, one would expect the true value to be included in 95 out of 100 ranges of the measurement result +/- 0.03 g. For this example measurement result, the range is 28.52 g to 28.58 g.*

TOPIC	LIMITATION	EXAMPLE STATEMENT OR SCENARIO
Frequency Data and Error Rate	It is not common for a report to include method performance frequency data or an error rate. Different terminology may be used: false-positive rate and false-negative rate, error rate, or average error rate. There is a risk of misunderstanding the applicability of frequency data or error rate provided.	**Example** This method has been shown to have an overall false positive rate (FPR) of 0.1% and an overall false negative rate (FNR) of 5.8%. *Understanding: If three or more results are possible from a method (e.g., same source, different source, inconclusive), more information will need to be provided on how each result type was handled in the calculation of FPR and FNR.* • *An average error rate may not apply to a case specific sample type or examiner. Understanding the method performance for that type of sample or for that examiner, will require obtaining detailed method performance frequency data.* • *Did the study protocol align with the quality assurance measures used by the FSSP (e.g., technical review, verification, or retesting)?* • *Accuracy can be affected by the type of item being tested. Did the study protocol include sufficient items similar to the case item type to support the calculation of an error rate?* • *Accuracy has been shown to vary between examiners. Did the study protocol include sufficient examiners with similar level of training and experience to the examiner in the case to support calculation of an error rate that applies to that case specific examiner?*
Likelihood Ratio (LR)	An LR evaluates the evidence from two perspectives (H_1 and H_2 propositions). The selection of these propositions is a limitation. The chosen propositions may not be correct. There may be multiple or changing propositions to evaluate different scenarios. The default defense proposition when pleading not guilty is that the prosecution proposition is not true. An LR may not always be provided in a report. An LR is provided in a report when an inclusionary statement is made. Statistical analysis is not required for exclusionary associations, comparisons between multiple questioned samples without a comparison to a known sample, nor application to inconclusive/uninterpretable results. Obtaining an LR in exclusionary associations may assist a decision maker to appropriately decide the weight to be given to the evidence. When an LR is provided, it is personal to the person providing it. It is based on their beliefs of what is important to consider. This basis may be robust data or experience and knowledge. An officer of the court will need to determine, either informally or formally with calculations, their own value for an LR. There is a risk that the LR number can lead to an incorrect interpretation of the evidence or misleading or inappropriate weight given to the evidence if the LR is misunderstood or misstated (prosecutor's or defense fallacy).	**Example** Item 1: DNA reference sample from (victim) Item 2: DNA reference sample from (suspect) Item 4: DNA sample from scene The genetic profile obtained from Item 4 is interpreted as a mixture of DNA from 3 contributors. Given this genetic profile, assuming Item 1 (victim) is a contributor, it is 1 million times more likely to observe this genetic profile if Item 1 (victim), Item 2 (suspect), and one unknown, unrelated individual are the contributors than if Item 1 (victim) and two unknown, unrelated individuals are the contributors. *Understanding:* *This example report statement explains the probability (Pr) of the evidence (E) and then provides the hypothesis (H). Pr (E|H).* $$LR = 1\ Million = \frac{(H_1)\ the\ contributors\ are\ the\ victim,\ the\ suspect,\ and\ one\ unknown,\ unrelated\ individual}{(H_2)\ the\ contributors\ are\ the\ victim,\ and\ two\ unknown,\ unrelated\ individuals}$$ *There is no single, correct LR. This provided LR of 1 million can vary between FSPs, FSSPs, or others by orders of magnitude depending on the information available, the assumptions made, and the model used for the calculation.* *The officer of the court will need to determine their own LR that aligns with their beliefs and the information that they have available.* *This statement does not (prosecutor's fallacy) mean H1 is 1 million times more likely to be true than H2 is.* *This statement does not mean (defense fallacy) that there are 323 people in the US (population 323 M) and 8000 people in the world (population 8 billion) that could be a contributor to the mixture, and this is reasonable doubt.*

Key Takeaways

1. Readers of reports should be aware that limitations exist both within the methods used and the results provided
2. Readers of reports must understand the report statements provided to ensure that they are applying any statistics correctly

Related Primers

Accreditation and Certification
Method Validation and Method Verification
Method Performance Statistics
Statistical Sampling
Population Statistics
Probability and Likelihood Ratios
Measurement Uncertainty, Frequency Data and Error Rates

Learn More

Likelihood Ratios for Lawyers

The Prosecutor's Fallacy and the Defense Attorney's Fallacy

Royal Statistical Society – Statistics and probability for advocates – Section 1.7

Part of NIST SP 1500-28
https://doi.org/10.6028/NIST.SP.1500-28