2021-S-0003
Standards for Determining Analytical and Stochastic Thresholds for Application to Forensic DNA Casework Using Electrophoresis Platforms

Human Forensic Biology Subcommittee
Biology Scientific Area Committee
Organization of Scientific Area Committees (OSAC) for Forensic Science

Draft OSAC Proposed Standard

2021-S-0003 Standards for Determining Analytical and Stochastic Thresholds for Application to Forensic DNA Casework Using Electrophoresis Platforms

Prepared by
Human Forensic Biology Subcommittee
Version 1.0
June 2022

Disclaimer:

This OSAC Proposed Standard was written by the Human Forensic Biology Subcommittee/Biology Scientific Area Committee of the Organization of Scientific Area Committees (OSAC) for Forensic Science following a process that includes an open comment period. This Proposed Standard will be submitted to a standards developing organization and is subject to change.

There may be references in an OSAC Proposed Standard to other publications under development by OSAC. The information in the Proposed Standard, and underlying concepts and methodologies, may be used by the forensic-science community before the completion of such companion publications.

Any identification of commercial equipment, instruments, or materials in the Proposed Standard is not a recommendation or endorsement by the U.S. Government and does not imply that the equipment, instruments, or materials are necessarily the best available for the purpose.

To be placed on the OSAC Registry, certain types of standards first must be reviewed by a Scientific and Technical Review Panel (STRP). The STRP process is vital to OSAC's mission of generating and recognizing scientifically sound standards for producing and interpreting forensic

science results. The STRP shall provide critical and knowledgeable reviews of draft standards or of proposed revisions of standards previously published by standards developing organizations (SDOs) to ensure that the published methods that practitioners employ are scientifically valid, and the resulting claims are trustworthy.

The STRP panel will consist of an independent and diverse panel, including subject matter experts, human factors scientists, quality assurance personnel, and legal experts, which will be tasked with evaluating the proposed standard based on a comprehensive list of science-based criteria.

For more information about this important process, please visit our website at: https://www.nist.gov/topics/organization-scientific-area-committees-forensic-science/scientific-technical-review-panels

Standards for Determining Analytical and Stochastic Thresholds for Application to Forensic DNA Casework Using Electrophoresis Platforms

Keywords: analytical threshold, stochastic threshold, DNA, validation, signal, artifact, noise

Foreword

Interpretation of short tandem repeat-based DNA profiles from electrophoresis platforms shall require determination of analytical and stochastic thresholds when interpretation will be performed without the use of probabilistic genotyping. Each of these will be defined in this document along with the individual minimum requirements for their determination and validation.

Such thresholds help to ensure confidence in the reliability of the data obtained, while clearly conveying assumptions under which data will be evaluated during downstream interpretation. The goal is for the laboratory to consistently produce reliable and reproducible designations of allelic data and potential allelic dropout that are supported by internal validation data and laboratory protocols.

If a laboratory, as part of its data analysis methods, makes binary determinations regarding the detection or non-detection of peaks for casework, analytical thresholds must be established. Similarly, if a laboratory, as part of its data analysis methods, makes binary determinations regarding the potential for allele drop-out in casework, stochastic thresholds must be established.

Whenever a threshold is applied, it is possible that a classification error may occur. Intrinsic to any analytical threshold is the expectation that non-reproducible noise will produce some peaks that are incorrectly classified as alleles because they exceed the threshold, and that some true alleles will be undetected because they produce peaks below the threshold. Intrinsic to any stochastic threshold is the expectation that some errors will occur in determining whether allelic drop-out may have occurred. Some heterozygous genotypes will incorrectly be classified as homozygotes because drop-out occurred, but the surviving peak is above the stochastic threshold, while some homozygotes will incorrectly be classified as possible heterozygotes because the homozygous peak is below the stochastic threshold. The advantage of determining thresholds based on statistical analysis of relevant empirical data, is that estimates can be made of the relative risk of these possible errors for a given threshold level. In setting thresholds, a statistically based approach must be employed by the laboratory to determine what proportion of these events are acceptable for the analysis of forensic casework.

The draft of this standard was developed by the Human Forensic Biology Subcommittee of the Organization of Scientific Area Committees for Forensic Science.

Table of Contents

Standards for Determining Analytical and Stochastic Thresholds for Application to Forensic DNA Casework Using Electrophoresis Platforms – 2021 Edition

1 Scope

These standards shall be used by forensic laboratories which, as part of their casework data analysis processes, are making determinations of: a) whether a peak in an electropherogram represents true signal or might be noise; and b) whether drop-out of a heterozygous sister allele to an observed peak either did not occur or might have occurred.

This standard is applicable to forensic STR DNA typing performed on electrophoresis platforms.

2 Normative References

The following reference is indispensable for the application of the standard. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ANSI/ASB Standard 038, First Edition 2020 – Standard for Internal Validation of Forensic DNA Analysis Methods

3 Terms and Definitions

For purposes of this document, the following definitions apply.

3.1
Allelic Peak
Signal distinguishable from noise arising from the amplification of the targeted DNA template.

3.2
Analytical Threshold
The minimum height requirement (in relative fluorescent units, RFUs, or equivalent) at and above which detected peaks on a STR DNA profile electropherogram can be reliably distinguished from instrument background noise; peaks above this threshold are generally not considered noise and are either artifacts or true alleles.

3.3
Artifact
Signal arising from the amplification of non-targeted DNA template, anomalies of the detection process, or by-products of primer synthesis.

3.4

Controls
Samples, of known types, run in parallel with experimental, reference, or evidence samples that are used to demonstrate that a procedure is working correctly.

3.5
Coverage factor (k)
Coverage factor (k factor) (Guide to Uncertainty of Measurement, GUM[1]): numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty.

3.6
Drop-out
(1) Failure of an otherwise amplifiable allele to produce a signal above analytical threshold because the allele was not present or was not present in sufficient quantity in the aliquot that underwent PCR amplification. (2) A hypothesis/postulate for the failure to observe one or more allelic peaks in an electropherogram that are expected for the assumed contributor(s) to a sample.

3.7
Empirical Data
Factual data that is based on actual measurement, observation, or direct sensory experience rather than on theory.

3.8
Internal Validation
In general, the accumulation of test data within the laboratory for developing standard operating procedures and demonstrating that the established protocols for the technical steps of the test and for data interpretation perform as expected in the laboratory.

3.9
Locus (loci)
Unique physical location(s) on the DNA molecule.

3.10
Noise
Meaningless output occurring in electronic equipment; it is random electronic variation that is generated by and intrinsic to the electronic circuitry. It ultimately establishes the smallest analytical signal that can be quantitatively measured with confidence. For DNA testing, see analytical threshold.

3.11

[1] NISTIR 6919, Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales, Val Miller, State Laboratory Program, Weights and Measures Division National Institute of Standards and Technology, Technology Administration, U.S. Department of Commerce

Signal
Meaningful output occurring in electronic equipment; nonrandom variation that can be distinguished from noise. For DNA testing, see analytical threshold.

3.12
Stochastic threshold
The peak height value (in relative fluorescent units, RFUs, or equivalent) in a DNA electrophoretic profile above which it is reasonable to assume that, at a given locus, allelic drop-out of a sister allele in a heterozygous pair has not occurred in a single source DNA sample; due to the possibility of shared alleles in mixed samples, the presence of allele peaks above the stochastic threshold is no guarantee that allele drop-out did not occur in mixed DNA sample profiles.

3.13
Stutter
An artifact of PCR amplification typically observed one or more repeat units smaller or larger than an STR allele in a DNA electrophoretic profile, may result from strand slippage during PCR amplification. A stutter peak is generally of lower RFU than the allele peak.

3.14
Validation
The process of performing and evaluating a set of experiments that establish the efficacy, reliability, and limitations of a method, procedure or modification thereof; establishing recorded documentation that provides a high degree of assurance that a specific process will consistently produce an outcome meeting its predetermined specifications and quality attributes. May include developmental and/or internal validation.

NOTE Variations from the FBI Quality Assurance Standards (QAS) definitions are to ensure consistency of the OSAC Lexicon across OSAC disciplines, but do not necessarily contradict or otherwise negate the QAS definitions.

4 Requirements

4.1 The laboratory shall have an analytical threshold[2] for each electrophoresis platform (*e.g.,* distinct CE models) used in casework that is established and tested using data generated during internal validation.

NOTE When multiple instruments of the same kind/model/platform are used for casework, data generated from each instrument should be considered due to potential variations in noise inherent to each instrument.

4.1.1 The laboratory shall determine and document the acceptable proportion of noise peaks that will exceed the analytical threshold (*e.g.*, as reflected by the number of standard

[2] This does not apply to the dye channel used for the internal size standard.

deviations above the noise mean [RFU]). This establishes the laboratory's predetermined expectation for acceptable performance of the analytical threshold.

NOTE As the number of standard deviations increases, the potential for allele non-detection also increases. Recognizing there is a tradeoff between the risk of allele non-detection and the risk of mistakenly labelling noise peaks, the analytical threshold should be set such that the probability that noise exceeds the analytical threshold is between 10^{-2} and 10^{-6} (*e.g.*, k=2 to k=5). See Table 9 (Mönich et al., page 115) relating k value to probability that a randomly generated noise peak exceeds the analytical threshold. [3]

4.1.2 The laboratory shall establish an analytical threshold based on internally generated empirical data acquired from the same electrophoresis platform, analysis software and DNA profiling chemistry utilized in casework.

4.1.3 Validation studies used to establish an analytical threshold shall include samples of known composition (*e.g.,* known genotype and negative controls[4]). Casework samples shall not be used to determine an analytical threshold.

4.1.4 Analytical thresholds shall be verified by confirming that the performance of the analytical threshold continues to meet the acceptable proportion of noise peaks that will exceed the analytical threshold whenever modifications to the instrument are made that have the potential to impact the noise output of the instrumentation (*e.g.*, performance check following change in laser and/or recalibration of the instrument).

4.1.5 Acceptable positions on an electropherogram to interrogate when establishing analytical thresholds are those that exclude possible allele or artifact peaks of known origin (*e.g.*, alleles and associated stutter products such as n-1, n-2, and n+1 positions, spectral pull-up peaks including those due to internal size standard, voltage spikes, unincorporated dye peaks).

4.1.6 The laboratory shall assess statistically-based analytical thresholds for each dye channel. A number of statistical methods to establish analytical threshold(s) have been described in the scientific literature.[5] Relevant references are provided in Annex A Bibliography. Laboratories employing a single global analytical threshold for all dye channels shall provide statistical support (*e.g.*, based on 1-way analysis of variance [ANOVA] showing no statistically significant differences [$p<0.05$] in noise across dye channels).

[3] Mönich, U.J., Duffy, K., Medard, M., Cadambe, V., Alfonse, L.E. and Grgicak C. "Probabilistic characterisation of baseline noise in STR profiles." Forensic Science International: Genetics 19 (2015): 107-122.

[4] Negative amplification and Reagent blank controls are acceptable sample types providing that they contain no indication of amplified product.

[5] Methods based on an extreme value calculation (e.g., 2X peak to trough difference) do not address the statistical confidence of a given analytical threshold. In addition, such methods can be easily skewed by outlier data and thus do not meet the requirements of this standard.

4.1.7 If the laboratory employs rounding (*e.g.*, to the nearest unit of 5 or 10 RFU), the implications of this rounding regarding the chance of mistaking noise for signal and the chance of not labeling a true allele in low template samples shall be documented.

4.2 The laboratory shall have a stochastic threshold[6] for each electrophoresis platform (*e.g.,* distinct CE models and DNA profiling chemistry) used in casework that is established and tested using data generated during internal validation.

NOTE When multiple instruments of the same kind/model/platform are used for casework, data generated from each instrument should be considered due to potential variations inherent to each instrument.

4.2.1 The laboratory shall determine and document the acceptable proportion of false homozygotes (drop-out) that will appear above the stochastic threshold (*e.g.*, as reflected by the number of standard deviations above the mean). This establishes the laboratory's predetermined expectation for acceptable performance of the stochastic threshold.

NOTE As the number of standard deviations increases, the potential for true homozygote detection decreases. Recognizing there is a tradeoff between detecting true homozygotes and the risk of mistakenly labelling a heterozygote with drop-out as a homozygote, the stochastic threshold should be set such that the probability that drop-out exceeds the stochastic threshold is between 10^{-2} and 10^{-6} (*e.g.*, k=2 to k=5).

4.2.2 The laboratory shall establish a stochastic threshold based on internally generated empirical data acquired from the same electrophoresis platform, analysis software, and DNA profiling chemistry utilized in casework.

NOTE Though laboratories using probabilistic genotyping systems are not required to establish or apply stochastic thresholds, they are still required to conduct validation studies that inform the laboratory of stochastic issues (*e.g.,* allele drop-out).

4.2.3 Validation studies of allelic drop-out used to establish a stochastic threshold shall include dilution series of single source samples of known genotype with a high level of heterozygosity and a range of differences in sister allele separation within each locus. The dilution series shall include DNA quantities around which allelic drop-out is likely to occur. Stochastic events are, by definition, random.

NOTE The use of larger data sets (*e.g.*, number of replicates) improves the accuracy of the stochastic threshold.

4.2.4 If processes are utilized to increase sensitivity (*e.g.*, increased amplification cycle number, increased injection time, and post-amplification purification or concentration of

[6] Determination of a stochastic threshold does not apply to the dye channel used for the internal size standard.

amplified products), the laboratory shall perform additional studies to determine the appropriate stochastic threshold(s) for the method(s) employed.

4.2.5 If processes are used to decrease sensitivity (e.g., reduced injection time, dilution of amplified product) that result in the interpretation of allelic peaks below the laboratory's stochastic threshold implemented for routine data analysis (i.e., data generated under methods that neither increase nor decrease sensitivity), the laboratory shall apply a stochastic threshold appropriate to the decreased sensitivity conditions.

NOTE For any profile generated using a reduced sensitivity method, where all interpreted peaks remain above the default stochastic threshold, the laboratory may evaluate whether or not the stochastic threshold implemented for routine data analysis is applicable to the decreased sensitivity method.

4.2.6 A number of methods to calculate a stochastic threshold have been described in the scientific literature. Relevant references are provided in Annex A Bibliography. The method selected must be supported by both the scientific literature and empirical data generated during internal validation by the laboratory.[7] The laboratory shall document the desired level of confidence (*e.g.*, as reflected by the number of standard deviations above the mean) for establishing a stochastic threshold.

4.3 Following the completion of the laboratory's internal validation study, all data and data analyses, calculations, and interpretations used to determine the analytical and stochastic thresholds must be documented in the final validation report.

4.3.1 The validation summary shall include the following information:

a) a record of predetermined specifications and quality attributes (*i.e.*, confidence level/error rate[8]) for accepting and implementing the thresholds(s) into operations.

b) a description of the samples, test methods, electropherograms and data used to calculate the threshold(s).

c) any formulae or theory applied to compute the thresholds.

d) reference literature as appropriate.

[7] Methods based on the largest surviving allele do not directly address the probability of allele drop-out at the stochastic threshold. Therefore, these methods are not recommended for determining a stochastic threshold. Thresholds shall be established based on statistical analysis, and skewed data must be appropriately transformed prior to further analysis.

[8] Type 1 Error, the rejection of a true null hypothesis (*e.g.*, a 99% confidence level has a 1% error rate). In the context of the analytical threshold, this represents the probability that an instrument noise peak will exceed the analytical threshold. In the context of the stochastic threshold, this represents the probability that a true heterozygous peak will exceed the stochastic threshold while the sister allele has dropped out.

4.3.2 The validation summary, all data and data analyses, calculations, and interpretations used to determine the analytical and stochastic thresholds shall be maintained by the laboratory.

5 Conformance

Documented conformance to these requirements needs to be: (1) approved by the laboratory's DNA Technical Leader or other appropriate personnel (2) communicated to all analysts during training, and (3) made readily available for review (*e.g.*, by auditors or inspectors, stakeholders who use reports generated by laboratory, etc.).

Annex A: (informative) Bibliography

Gilder, J., Doom, T., Inman, K. and Krane, D. Run-specific limits of detection and quantitation for STR-based DNA testing. J For. Sciences, 2007, 52:(1):97-101.

Bieber, F.R., Buckleton, J.S., Budowle, B., Butler, J.M. and Coble, M.D. "Evaluation of forensic DNA mixture evidence: protocol for evaluation, interpretation, and statistical calculations using the combined probability of inclusion." *BMC genetics* 17.1 (2016): 125.

Butts, E.L., Kline, M., Duewer, D., Hill, C., Butler, J.M. and Vallone, P. "NIST validation studies on the 3500 Genetic Analyzer." *Forensic Science International: Genetics Supplement Series* 3.1 (2011): e184-e185.

Gill, P., Puch-Solis, R. and Curran, J. "The low-template-DNA (stochastic) threshold—its determination relative to risk analysis for national DNA databases." *Forensic Science International: Genetics* 3.2 (2009): 104-111.

Mönich, U.J., Duffy, K., Medard, M., Cadambe, V., Alfonse, L.E. and Grgicak C. "Probabilistic characterisation of baseline noise in STR profiles." *Forensic Science International: Genetics* 19 (2015): 107-122.

Puch-Solis, R., Kirkham, A.J., Gill, P., Read, J., Watson, S. and Drew, D. "Practical determination of the low template DNA threshold." *Forensic Science International: Genetics* 5.5 (2011): 422-427.

Rowan, K.E., Wellner, G.A. and Grgicak, C.M. "Exploring the impacts of ordinary laboratory alterations during forensic DNA processing on peak height variation, thresholds, and probability of dropout." *Journal of forensic sciences* 61.1 (2016): 177-185.

Federal Bureau of Investigations Quality Assurance Standards (QAS) for Forensic DNA Testing Laboratories *https://www.fbi.gov/file-repository/quality-assurance-standards-for-forensic-dna-testing-laboratories.pdf*

NISTIR 6919, Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales, Val Miller, State Laboratory Program, Weights and Measures Division National Institute of Standards and Technology, Technology Administration, U.S. Department of Commerce.