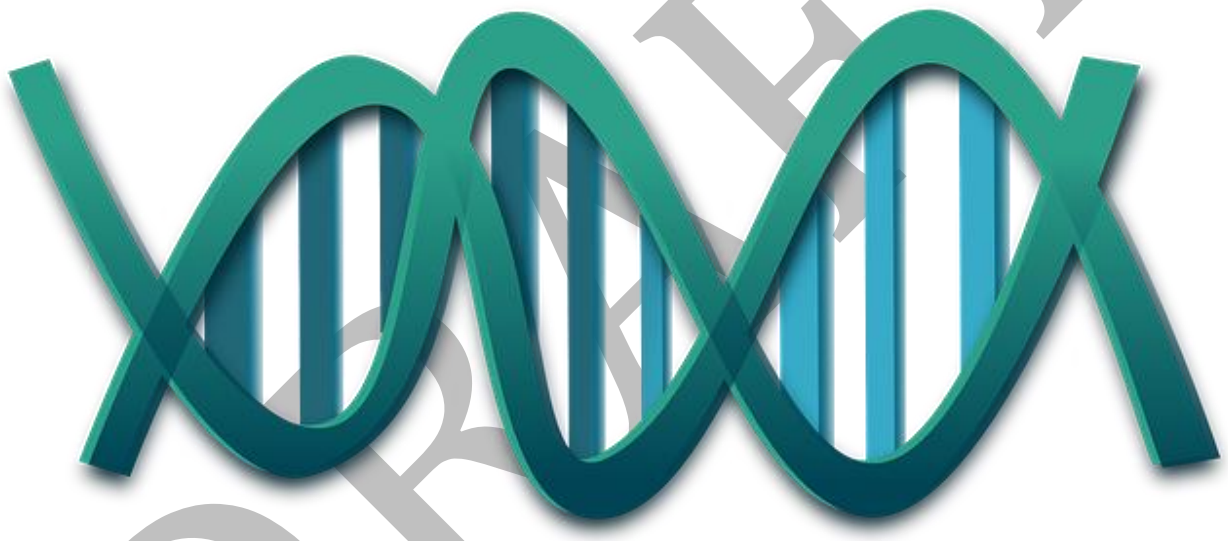
This document has been accepted by the Academy Standards Board (ASB) for development as an American National Standard (ANS). For information about ASB and their process please refer to asb.aafs.org. This document is being made available at this stage of the process so that the forensic science community and interested stakeholders can be more fully aware of the efforts and work products of the Organization of Scientific Area Committees for Forensic Science (OSAC). The documents were prepared with input from OSAC Legal Resource Committee, Quality Infrastructure Committee, and Human Factors Committees, as well as the relevant Scientific Area Committee. The content of the documents listed below is subject to change during the standards development process within ASB, and may not represent the contents of the final published standard. All stakeholder groups or individuals, are strongly encouraged to submit technical comments on this draft document during the ASB's open comment period. Technical comments will not be accepted if submitted to the OSAC Scientific Area Committee or Subcommittees.

VALIDATION STANDARDS FOR PROBABILISTIC GENOTYPING SYSTEMS
DRAFT

DRAFT DOCUMENT

ASB Standard 00XX
2016

Validation Standards for Probabilistic Genotyping Systems



ASB Standard 00XX
2016

Validation Standards for Probabilistic Genotyping Systems

Keywords: *validation, probabilistic genotyping, statistics, likelihood ratio*

Abstract
These standards were designed to provide direction and guidance to laboratories for the validation of probabilistic genotyping systems as they relate to the interpretation of autosomal short tandem repeat analysis.

Foreword

Validation of a new methodology is typically defined as developmental and internal, and each will be defined in this document along with their individual minimum requirements as it relates to probabilistic genotyping. Developmental validation may be conducted outside the laboratory (i.e., by the manufacturer, developer, or other testing laboratory) planning to use it. In these instances, the laboratory validating the system may choose to adopt and reference these studies already performed. However, developmental validation is not meant to replace internal validation. Instead, depending on the particular functions and applications of the system and its planned use in the laboratory, each laboratory will need to perform internal studies to demonstrate the reliability of the software and any potential limitations. With multi-laboratory systems using a common protocol, internal validation may be shared by all locations.

If a laboratory will be incorporating a probabilistic genotyping system in its casework utilizing different DNA typing parameters and protocols , the software and individual interpretation protocols will need to be validated with each method (e.g., standard and enhanced detection methods). The individuals designing and evaluating the validation studies shall possess the appropriate foundational knowledge in the calculation and explanation of likelihood ratios.

The Biology DNA Interpretation and Reporting Sub-Committee is under the Biology/DNA Scientific Area Committee (SAC) and is involved in developing and vetting standards and guidelines related to forensic DNA interpretation. A high priority for this sub-committee was to define the necessary requirements for validating probabilistic genotyping systems. A task group was empaneled and standards were developed for laboratories to follow when validating these systems.

Acknowledgements

Editor:

Deputy Editor(s):

Draft Working Group Members:

Joel Sutton, Defense Forensic Science Center
Steven Myers, California Department of Justice Jan Bashinski DNA Laboratory
Michael Coble, Ph.D., NIST
Simone Gittelson, Ph.D., NIST
Sandy Zabell, Ph.D., Northwestern University
Kathleen Corrado, Onondaga County Center for Forensic Sciences
Jeffrey Nye, Michigan State Police Forensic Science Division
Susan Greenspoon, Ph.D., Virginia Department of Forensic Science
Michael Adamowicz, Ph.D., University of New Haven
James Curran, Ph.D., University of Auckland
Adele Michell, Ph.D., Merck
Mark Powell, San Francisco Police Department Crime Lab

Consensus Group Members:

Table of Contents

Validation Standards for Probabilistic Genotyping Systems – 2016 Edition

1 Scope

1.1 These standards shall be used by laboratories for the validation of probabilistic genotyping systems related to interpreting autosomal STR results.

1.2 These standards are not meant to be applied to probabilistic genotyping systems which have been previously validated.

1.3 For probabilistic genotyping systems currently in use, laboratories are advised to review their previous validation relative to these standards.

2 Normative References

The following referenced documents are indispensable for the application of this document.

Federal Bureau of Investigation, (2011) *Quality Assurance Standards for Forensic DNA Testing Laboratories*, available at http://www.fbi.gov/about-us/lab/codis/qas-standards-for-forensic-dna-testing-laboratories-effective-9-1-2011

Scientific Working Group on DNA Analysis and Methods, (2015) *Guidelines for the Validation of Probabilistic Genotyping Systems*, available at http://www.swgdam.org

Scientific Working Group on DNA Analysis and Methods, (2012) *Validation Guidelines for DNA Analysis Methods*, available at www.swgdam.org

3 Terms and Definitions

For purposes of this document, the following definitions apply.

3.1
accuracy studies
Specifically defined in terms of developmental and internal validation.

3.1.1
developmental
Studies performed to establish that the calculations performed by the probabilistic genotyping system are correctly executed, and that the results obtained produce the expected likelihood ratio for situations where the calculations can be performed manually or with an alternate software program or application. Such situations include profile results from single source samples, 2-person mixtures with unambiguous major and minor contributors, and 2-person mixtures with equal mixture proportions.

3.1.2
internal
Studies performed to establish the range of sample and profile parameters for which the probabilistic genotyping system performs as expected such that the results are reasonable and consistent with expectations based on non-probabilistic mixture analysis methods.

3.2
case-type profiles
Data exhibiting features that are representative of a plausible range of casework conditions for mixtures and single-source samples. These features include masked/shared alleles and stutter, degradation (including different degradation levels for different contributors to a mixture), allele and locus drop-out, and inhibition.

3.3
developmental validation
The acquisition of test data to verify the functionality of the system, the accuracy of statistical parameters, the appropriateness of analytical and statistical parameters, and the determination of limitations of the system.

3.4
fully-continuous model
A statistical model and accompanying method that evaluates DNA profiles using peak height information to assign weights to the observed peak heights for different combinations of contributor genotypes at all tested loci. By modeling peak heights in this manner, this method does not typically rely on certain traditional parameters for interpretation (i.e., stutter, stochastic and peak height ratio thresholds; and mixture proportion ranges). Allele drop-out and/or drop-in may or may not be explicitly considered.

3.5
internal validation
The accumulation of test data within the laboratory to demonstrate that established parameters, software settings, formulae, algorithms and mathematical functions perform as expected; and that the information/results/data obtained is correct and consistent with expected values.

3.6
performance check
A quality assurance measure to assess the functionality of the probabilistic genotyping software following a material modification (e.g., a change in the computational core). This would typically involve functional testing of the software verifying it is performing tasks as expected and comparing results to previously validated versions of the software using the same data or sample set where possible.

3.7
precision studies
Studies performed to evaluate the variation in likelihood ratios calculated from repeated software analyses of the same input data using the same set of conditions/parameters. Some probabilistic genotyping systems may inherently not produce the same statistical calculation from repeated analysis. Where applicable, studies should demonstrate the range of values that can be expected from multiple analyses of the same data, and are the basis for establishing an acceptable amount of variation in the statistical calculations.

3.8
probabilistic genotyping
The use of biological modeling (i.e., statistical modeling informed by biological data), statistical theory, computer algorithms, and/or probability distributions to infer genotypes and/or calculate likelihood ratios.

3.9
probabilistic genotyping system
Software, or software and hardware, which utilizes a probabilistic genotyping approach to infer genotypes and/or calculate likelihood ratios.

3.10
semi-continuous model
A statistical model and associative method that evaluates DNA profiles by assigning weights (i.e., probabilities between 0 and 1) for the observed data assessing the presence or absence of allelic peaks for different contributor genotypes. These models rely on rules or pre-defined thresholds for initial interpretation (i.e., stutter, probability of drop-out and/or drop-in) and do not take peak height information into consideration when the software assigns weights to these genotypes in the final calculation of the likelihood ratio.

3.11
sensitivity studies
Studies performed to assess the ability of the probabilistic genotyping system to support the presence of a known contributor's or multiple known contributors' DNA over a broad variety of evidentiary typing results to include mixtures and low-level DNA template quantities where stochastic effects are likely to be present.

3.12
specificity studies
Studies performed to assess the ability of the probabilistic genotyping system to support true negatives over a broad variety of evidentiary typing results to include mixtures and low-level DNA template quantities. True negatives would correctly indicate the absence of an individual who is known not to contribute.

3.13
validation
The process of performing a set of experiments that establish the efficacy, reliability, and limitations of a method, procedure or modification thereof. Establishing recorded documentation that provides assurance based on empirical data that a specific process will consistently produce an outcome meeting its predetermined specifications and quality attributes.

4 Requirements

4.1 The laboratory shall validate a probabilistic genotyping system prior to its use for casework samples in the laboratory.

4.1.1 Validations shall include both developmental and internal studies. Developmental validation may be conducted by the manufacturer/developer of the application or another laboratory/agency.

4.1.2 Developmental validation studies shall address the following: accuracy, sensitivity, specificity, and precision. These studies shall include case-type profiles of known composition.

4.1.3 Internal validation studies shall address the following: accuracy, sensitivity, specificity, and precision. These studies shall include internally generated case-type profiles of known composition that represent (in terms of number of contributors, mixture ratios, and total DNA template quantities) the range of actual casework samples intended for analysis with the system at the

laboratory. Studies shall not be limited to pristine DNA samples but shall also include compromised DNA samples (e.g., partial, degraded, and inhibited samples).

4.1.4 Internal validation studies shall include evaluating user input parameters that vary run to run such as number of contributors, alternate hypothesis testing, and artifact considerations (e.g., stutter). The parameters may vary depending upon the approach or intended use of the software. Therefore, the specific parameters to be tested shall be determined by the laboratory.

4.1.5 For internal validation, the laboratory shall evaluate both the appropriate sample type (i.e., number of contributors, mixture ratios, and template quantities) and the number of samples within each type to demonstrate the potential limitations and reliability of the software. The laboratory shall base this evaluation on the function and intended application of the software

4.2 The underlying scientific principle(s) of the probabilistic genotyping model and associative method and software including the mathematical basis and underlying algorithms shall be published or accepted for publication in peer-reviewed scientific journal(s).

4.3 Quality assurance parameters, analytical procedures, and interpretation guidelines shall be derived from internal validation studies. Developmental and manufacturer recommendations may be used in addition to internal validation studies but should not replace internal validation.

4.4 Software modifications that may impact the analytical process, interpretation, or reported result(s) shall be evaluated as to the extent of the impact to determine whether a validation or performance check is required prior to implementation. Such modifications shall require a validation or performance check of the affected software component. If neither is conducted after a software modification, the laboratory shall document the justification (e.g., software update simply enhances visual output or displays, therefore no performance check was conducted).

4.5 All internal validation and performance check studies shall be documented and retained by the laboratory.

4.6 The laboratory shall have a mechanism to verify that validated system settings are used each time an analysis is performed.

4.7 Prior to implementation, the laboratory shall validate its interpretation procedures and guidelines using a different data set than what were initially used to validate the software. This serves to further verify the established software parameters.

5 Conformance

Documentation demonstrating conformance with the standards described in this document will be signed and dated by the laboratory's DNA technical leader and will be made readily available in hard copy and/or electronic form for review by the assessor.

Annex A
(informative)

Bibliography

This is not meant to be an all-inclusive list as the group recognizes other publications on this subject may exist. At the time these standards were drafted, these were the publications available to the working group members for reference. Additionally, any mention of a particular software tool or vendor as part of this bibliography is purely incidental, and any inclusion does not imply endorsement by the authors of this document.

1] Gill P., Gusmão L., Haned H., Mayr W.R., Morling N., Parson W., et al., (2012) DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods, Forensic Sci. Int.: Genetics 6: 679-688.

2] Kelly H., Bright J.-A., Buckleton J.S., Curran J.M., (2014) A comparison of statistical models for the analysis of complex forensic DNA profiles, Science & Justice 54: 66-70.

3] Steele C.D. and Balding D.J., (2014) Statistical evaluation of forensic DNA profile evidence, Annu. Rev. Stat. Appl. 1:361-84.

4] Bright J.-A., Evett I.W., Taylor D., Curran J.M., Buckleton J.S., (2015) A series of recommended tests when validating probabilistic DNA profile interpretation software, Forensic Sci. Int.: Genetics 14: 125-131.

5] Taylor D., Bright J.-A., Buckleton J.S., (2013) The interpretation of single source and mixed DNA profiles, Forensic Sci. Int.: Genetics 7: 516-528.

6] Perlin M.W., Legler M.M., Spencer C.E., Smith J.L., Allan W.P., Bellrose J.L., et al., (2011) Validating TrueAllele DNA mixture interpretation, J Forensic Sci 56:1430–47.

7] Taylor D., Buckleton J.S., Evett I.W., (2015) Testing likelihood ratios produced from complex DNA profiles. Forensic Sci. Int.: Genetics 16: 165-171.

8] Greenspoon S.A., Schiermeier-Wood L., Jenkins B.C., (2015) Establishing the limits of TrueAllele® casework: A validation study, J Forensic Sci., doi: 10.1111/1556-4029.12810. [Epub ahead of print].

9] Cowell R.G., (2009) Validation of an STR peak area model. Forensic Sci. Int.: Genetics 3: 193-199.

10] Steele C.D., Greenhalgh M., Balding D.J., (2014) Verifying likelihoods for low template DNA profiles using multiple replicates. Forensic Sci. Int.: Genetics 13: 82-89.

11] Taylor D., (2014) Using continuous DNA interpretation methods to revisit likelihood ratio behavior. Forensic Sci. Int.: Genetics 11: 144-153.

Academy Standards Board, AAFS
410 North 21st Street
Colorado Springs, CO 80904

http://asb.aafs.org/