ANSI/ASB Standard 018, First Edition
2020

Standard for Validation of Probabilistic Genotyping Systems

Standard for Validation of Probabilistic Genotyping Systems

ASB Approved December 2019

ANSI Approved July 2020

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

Foreword

The validation of computer software systems used for the probabilistic evaluation and interpretation of genetic information from forensic casework is a critical component of the validation process any caseworking laboratory using such software undergoes. Validations of such systems provide the study results and conclusions necessary for customers of forensic science service providers to have confidence in the evidence provided.

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 planning to use it (i.e., by the manufacturer, developer, or other testing laboratory). 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.

If a laboratory will be incorporating a probabilistic genotyping system in its casework utilizing multiple sets of 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).

This document was revised, prepared, and finalized as a standard by the DNA Consensus Body of the AAFS Standards Board. The draft of this standard was developed by the Biological Data Interpretation and Reporting Subcommittee of the Organization of Scientific Area Committees (OSAC) for Forensic Science.

The AAFS Standards Board (ASB) is an ANSI-accredited Standards Developing Organization with the purpose of providing accessible, high quality science-based consensus forensic standards. The ASB is a wholly owned subsidiary of the American Academy of Forensic Sciences (AAFS), established in 2015 and accredited by the American National Standards Institute (ANSI) in 2016. The ASB consists of Consensus Bodies (CB), which are open to all materially interested and affected individuals, companies, and organizations; a Board of Directors; and Staff.

The following applies to all ASB documents:

the term **'shall'** indicates that a provision is mandatory, and can be audited for compliance

the term **'should'** indicates that a provision is not mandatory, but recommended as good practice.

All hyperlinks and web addresses shown in this document are current as of the publication date of this standard.

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

Table of Contents

Standard for Validation of Probabilistic Genotyping Systems

1 Scope

1.1 This standard sets forth the requirements to be used by laboratories for the validation of probabilistic genotyping systems related to interpreting autosomal STR results. Amelogenin is not covered by this standard.

1.2 Laboratories are advised to review their validation for compliance with this standard, supplement validation where necessary, and modify existing protocols accordingly.

2 Normative References

There are no normative reference documents, Annex B, Bibliography, contains informative references.

3 Terms and Definitions

For purposes of this document, the following definitions apply.

3.1
accuracy studies
Studies performed to assess the degree of conformity of a measured quantity to its actual (true) value. In probabilistic genotyping, these are studies performed to establish that the calculations made 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. However, profile results where the ground truth is not known are not suitable for accuracy studies.

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 PCR inhibition.

3.3
developmental validation
The accumulation of test data to demonstrate that established parameters, software settings, formulae, algorithms and mathematical functions perform as expected. Developmental validation should also demonstrate any known limitations of the system. Developmental validation may be conducted outside the laboratory planning to use it (i.e., by the manufacturer, developer, or other testing laboratory) and will precede any internal validations.

3.4
internal validation
The acquisition of test data within the laboratory 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.

3.5
performance check
A quality assurance measure to assess the functionality of the probabilistic genotyping software following a minor change such as reformatting of output reports. 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.

3.6
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. Probabilistic genotyping systems inherently do not produce the same exact numbers in repeated analysis. These studies should demonstrate the range of values that can be expected from multiple analyses of the same data.

3.7
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.

3.8
probabilistic genotyping system
Software, or software and hardware, which utilizes a probabilistic genotyping approach to infer genotypes and/or calculate likelihood ratios.

3.9
sensitivity studies
Studies performed to assess the ability of the probabilistic genotyping system to support the presence of a known contributor.

3.10
specificity studies
Studies performed to assess the ability of the probabilistic genotyping system to support the absence of true non-contributors. A true non-contributor is an individual who is known not to contribute.

3.11
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

NOTE Refer to Annex A, Requirements - Supporting Information, for additional information on the requirements in this section.

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. Developmental validation shall not replace internal validation.

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 that represent (in terms of number of contributors, mixture ratios, and total DNA template quantities) the range of scenarios that would likely be encountered in casework. Studies shall not be limited to pristine DNA samples but shall also include compromised DNA samples (e.g., low template, degraded, and inhibited samples).

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., low template, degraded, and inhibited samples). The internal validation shall not exceed the scope of the conditions tested in the developmental validation. Case type profiles that fall outside the range of conditions explored in developmental validation shall require additional developmental validation studies. See Annex A.

4.1.4 Internal validation studies shall include evaluating user input parameters that vary run to run. The effects of artifacts (e.g., stutter) and parameters that relate to the statistical algorithm (e.g., run time parameters for the software system that can vary from system to system) shall also be evaluated. 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 Internal validation studies shall also include the evaluation of multiple propositions for case type samples to aid in the development of propositions. Such studies shall also consider the effect of overestimating and underestimating the number of contributors.

4.1.6 For internal validation, the laboratory shall evaluate both the appropriate sample types (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 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 in peer-reviewed scientific journal(s).

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

4.4 Software modifications, changes to computing platform or changes to upstream analytical processes (i.e., amplification processes, detection platforms) that may impact the interpretation or reported result(s) shall be evaluated 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, changes to computing platform or changes to upstream analytical processes, the laboratory shall document the justification (e.g., software update simply enhances visual output or displays, therefore no performance check was conducted). See Annex A.

4.5 All validation and performance check studies conducted by the laboratory shall be documented and retained by the laboratory. See Annex A.

4.6 The laboratory shall have a mechanism to record the software settings that are used each time an analysis is performed. See Annex A.

4.7 Prior to implementation, the laboratory shall verify the functionality of its defined software settings and parameters utilizing different data sets than what were originally used to establish those settings and parameters. See Annex A.

5 Conformance

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

Annex A
(normative)

Requirements – Supporting Information

The following information is provided to aid personnel responsible for developing the validation and any personnel responsible for carrying out the validation. While each of the standards listed shall be addressed in the development and use of the laboratory validation protocol(s), the approaches used, the type of data evaluated, and the details of the protocols will vary between laboratories.

Requirement 4.1.3 - Repeated testing and data analysis are critical to the understanding of variability. While specific requirements for the minimum number of studies and sample sets used for validation studies are not detailed in this standard, the laboratory shall perform sufficient studies to address the variability inherent to the various aspects of DNA testing, data generation, analysis and interpretation of data and user input parameters.

Requirement 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. All computer programs are subject to code revisions, improvements and release cycles. As such, it is useful to have some concept of which changes made to the software by the developers are likely to have a fundamental impact, and equally how such changes can be recognized. A laboratory does not need to perform additional validation based solely upon changes to software version numbers or build numbers. Additional validation or a performance check shall be based on the list of documented changes provided by the developer that accompany each updated version of the software installed in the laboratory.

Requirement 4.5 - All internal validation and performance check studies shall be documented and retained by the laboratory. Any validation and performance check studies may take a significant amount of time and are likely to result in a considerable amount of documentation output material. It is incumbent upon any laboratory performing these studies to retain these results for the examination and evaluation by third parties. The results should be documented in such a way that the performance checks and validations can be reproduced and decisions made on the basis of these studies documented. Laboratories shall have a summary statement of the sample types of which the developer used to for their developmental validation.

Requirement 4.6 **-** The laboratory shall have a mechanism to record the software settings that are used each time an analysis is performed. Probabilistic genotyping software usually has a number of settings that are either specific to a laboratory, specific to a case, or specific to a run within a case. The latter may occur when a probabilistic genotyping analysis that incorporates elements of randomness is performed multiple times for the same evidentiary items. Settings may include input parameters specific to the algorithm (such as the probability of dropout, or the number of MCMC burn-in iterations), laboratory specific parameters (such as the distribution parameters for stutter peaks based on historical data from that laboratory), or run specific parameters (such as the number of contributors, or the number of MCMC iterations retained for inference). Any parameter/input in the system that the user can change should be recorded for examination, evaluation and reproduction. The recording of these settings will allow the system to be configured in an identical manner and allow a third party to achieve the same (or similar) outputs. The outputs will generally not be identical unless the same random number seed(s) is/are used.

Requirement 4.7 **-** Prior to implementation, the laboratory shall verify the functionality of its defined software settings and parameters utilizing a different data set than what was originally used to establish those settings and parameters. This serves to further verify the established software settings and parameters. Probabilistic genotyping software systems are calibrated using historical data ideally from the same laboratory in which that system is employed. It is therefore important to test the system by exposing it to data that it has not seen in the past. This, in turn, will provide the laboratory with a more realistic assessment of the readiness of the system for casework. The new data should be comprised of samples that represent the variety of casework handled within the validating laboratory.

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

2] Coble M.D., J. Buckleton, J.M. Butler, T. Egeland, R Fimmers, P. Gill, et al. "DNA Commission of the International Society for Forensic Genetics: Recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications." *Forensic Sci. Int.: Genetics 25*, 2016, pp. 191-197.

3] Cowell R.G. "Validation of an STR peak area model." *Forensic Sci. Int.: Genetics 3*, 2009, pp. 193-199.

4] Federal Bureau of Investigation, (latest version) *Quality Assurance Standards for Forensic DNA Testing Laboratories.* [1]

5] Gill P. L., Gusmão, H. Haned, W.R. Mayr, N. Morling, W. Parson, et al., "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*, 2012, PP. 679-688.

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

7] Institute of Electrical and Electronics Engineers. IEEE Std 1012-2012 - IEEE standard for system and software verification and validation, 2012;209; http://ieeexplore.ieee.org/document/6204026/.

8] Kelly H., J.-A. Bright, J.S. Buckleton, and J.M. Curran. "A comparison of statistical models for the analysis of complex forensic DNA profiles." *Science & Justice*, *54,* 2014, pp. 66-70.

9] Perlin M.W., M.M. Legler, C.E. Spencer, J.L. Smith, W.P. Allan, J.L. Bellrose, et al. "Validating TrueAllele DNA mixture interpretation." *J Forensic Sci 56*, 2011, pp. 1430–47.

10] Steele C.D. and D.J. Balding. "Statistical evaluation of forensic DNA profile evidence." *Annu. Rev. Stat. Appl. 1,* 2014, pp. 361-84.

[1] https://www.fbi.gov/file-repository/quality-assurance-standards-for-forensic-dna-testing-laboratories.pdf/view

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

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

13] Taylor D., J.S. Buckleton, and I. W. Evett. "Testing likelihood ratios produced from complex DNA profiles." *Forensic Sci. Int.: Genetics 16*, 2015, pp. 165-171.

14] Steele C.D., M. Greenhalgh, and D.J. Balding. "Verifying likelihoods for low template DNA profiles using multiple replicates." *Forensic Sci. Int.: Genetics 13*, 2014, pp. 82-89.

15] Scientific Working Group on DNA Analysis and Methods, (latest version) *Guidelines for the Validation of Probabilistic Genotyping Systems* [2].

16] Scientific Working Group on DNA Analysis and Methods, (latest version) *Validation Guidelines for DNA Analysis Methods* [2].

[2] Available at .[2] https://www.swgdam.org/publications

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

www.asbstandardsboard.org