Comment on Proposed SWGDAM 2025 Guidelines for Internal PGS Validation

Mark W Perlin, PhD, MD, PhD
Chief Scientist
Cybergenetics
Pittsburgh, PA
14 February 2025

The proposed 2025 SWGDAM guidelines are centered on many issues that arise with validating one particular continuous probabilistic genotyping system (PGS), ESR's STRmix™. The continuous PGS approach was developed 25 years ago, describing a continuous likelihood function that connects data, genotype and mixture weight (1).

The proposed guidelines are certainly helpful for laboratories that elect to validate STRmix or its progeny, such as EuroForMix. However, SWGDAM does not discuss other validation approaches or reliability metrics. This brief note offers some comments.

1. SWGDAM nomenclature

Much of the proposed guideline's terminology comes from STRmix. For example, the ESR developers idiosyncratically call their genotype likelihoods "weights". This is nonstandard numerical usage, since "weight" is often reserved for a set of nonnegative numbers that add to one, as in "mixture weight" (1). Probabilities must add to one, but likelihoods need not do so.

2. Laboratory data independence

Some PGS software requires initial calibration on data from a laboratory's polymerase chain reaction (PCR) amplification and amplicon detection instruments. However, Cybergenetics' TrueAllele® results are independent of DNA laboratory, short tandem repeat (STR) process and genetic analyzer, or proficient human operator (2).

In most scientific fields, hardware and software do not depend on a particular laboratory or the examined data. For example, a microscope can be used in any hospital on any slide. An Excel spreadsheet can calculate statistics from any data source, without requiring site-specific calibration. Internal validation isn't needed.

3. Analytical data thresholds

Analytical thresholds strive to separate allele signal from background noise. This data reduction attempt may not be perfect. For systems using analytical thresholds, a sensitivity study for each case item at different threshold levels is advisable.

Advanced Bayesian modeling doesn't need or use a threshold. The computer can instead consider all the data, including noise peaks, and assign higher likelihood to genotype combinations that better explain the data (3).

4. PGS parameter settings

Some limited probability models may require external input for PCR stutter, allele drop out and drop in, variance parameters, etc. They may need initial parameter calibration before processing casework STR data. For systems requiring input parameter settings, a sensitivity study for each case item on different sets of parameter levels is advisable.

Advanced Bayesian models (3) use additional variables to derive nuisance parameters directly from DNA evidence data, without advance calibration. In TrueAllele, these parameters include PCR stutter and variance; drop in/out variables aren't needed.

5. Number of contributors

In general, an exact number of contributors (NOC) to a DNA mixture cannot be determined from the observed evidence data. The data's NOC is a probability distribution of contributor number possibilities. Visually examining the data gives a range of NOC values. Considering all reasonably feasible NOCs that might explain the data in a case is advisable.

6. Genotype LR information

The likelihood ratio (LR) quantifies the impact of DNA evidence on genotype (or match) probability. The log(LR) is a standard measure of information (4). One LR log unit is called a ban; a thousandth of a log unit is a milliban.

A probabilistic genotype is completely determined by its prior and posterior probabilities (or likelihood) at each STR locus (5). A useful way to describe the LR is as the ratio of posterior to prior (6) genotype probability for a reference genotype (7).

Generally, log(LR) information is proportional to log(DNA) amount (3, 8). Some data factors (e.g., allele sharing, peak saturation) are known to reduce identification information.

7. Genotype LR distributions

It is useful to examine the distribution of real-valued log(LR) numbers when a probabilistic evidence genotype is compared with all possible reference genotypes. Population-weighted references produce a Noncontributor specificity log(LR) distribution curve, centered left of zero. Weighting the references by posterior genotype probability produces a Contributor sensitivity distribution, centered right of zero. See Figure.

A genotype's distributions show its log(LR) outcome probabilities when the genotype is compared with a reference. The Noncontributor distribution predicts LR outcome when the reference didn't contribute their DNA to the evidence. The Contributor distribution predicts LR outcome when the reference resembles the DNA evidence contributor.

Mathematical convolution of locus probability functions can produce complete (10^{24} references considered) exact (milliban resolution) distributions in a fraction of a second (9). TrueAllele's Distribution module provides this convolution functionality.

8. LR error rates

LR error rate gives helpful context to a trier of fact assessing DNA evidence. The error rate (ER) depends on the LR value. A stronger LR generally has a lower ER.

An exact error rate can be computed directly from a LR distribution (9). The ER of an inclusionary LR is the right tail area of the genotype's Noncontributor distribution, evaluated at the LR. Conversely, the exclusionary LR's ER is the left tail area of the genotype's Contributor distribution at LR.

For inclusionary LR values (LR>1), a well-known upper bound on error rate is the LR reciprocal, or ER ≤ 1/LR. For exclusionary values (LR<1), the bound is ER ≤ LR. When no ER has been reported, these bounds can help experts explain LR results.

The ER provides a frequency context for the LR that indicates how often the LR may be misleading (10). It gives the chance that a subject with an inclusionary LR may not have contributed their DNA to the evidence. Or, conversely, the chance that an exclusionary LR corresponds to a true contributor. Reporting the ER for a LR is advisable.

9. Sensitivity and specificity validation

Sensitivity and specificity are two key PGS validation axes. A genotype's Contributor distribution contains everything knowable about the genotype's sensitivity and LR-dependent error rates (9). Similarly, the Noncontributor distribution fully explicates genotype specificity.

Some validation approaches sample small sets of LRs formed by comparing evidence genotypes with a set of reference genotypes. Comparing with randomly generated population references forms Noncontributor specificity histograms (11). Collating LRs from observed matches gives Contributor sensitivity histograms (12, 13). Internal validation studies typically calculate 10^3 to 10^6 LR values. This time consuming, labor intensive, approximate result is unnecessary when using exact composite distributions.

Exact LR composites are more complete than sampled LRs (9). The distribution combination maintains exact milliban resolution, takes seconds to build, and represents 10^{24} reference comparisons. The resulting sensitivity and specificity distributions consider all possible trillion-trillion references, not just a few thousand samples. TrueAllele's Distribution module provides this composite distribution functionality.

10. Casework validation

Some PGS validation methods rely on “ground truth” knowledge to determine error rates. Such external knowledge isn’t needed when accurate dense LR distributions are available. This was explained in a published TrueAllele casework data PGS validation study (11, p. 12) for inclusionary LR values:

> “Since the method’s high [Noncontributor] *specificity* assures identification hypothesis H with considerable certainty, we can safely examine the Pr{**X**=x | H} *sensitivity* distribution of [Contributor] positive log(LR) values.”

Exchanging Noncontributor and Contributor roles enables exclusionary LR validation.

An empirically derived genotype contains all the information about its Contributor and Noncontributor distributions, and its LR-dependent error rates. Nothing more is needed for casework applications. Validation is done on sets of genotypes and combines their derived LR distributions to obtain sensitivity and specificity information for the laboratory’s PGS process.

References

1. Perlin MW, Szabady B. Linear mixture analysis: a mathematical approach to resolving mixed DNA samples. J Forensic Sci. 2001;46(6):1372-7.
2. Allan WP, Bracamontes JM, Legler MM, Perlin JB, Perlin MW. Different TrueAllele® users, same DNA answer: a multi-center proficiency study. International Symposium on Human Identification; San Antonio, TX: Promega; 2024.
3. Perlin MW, Sinelnikov A. An Information Gap in DNA Evidence Interpretation. PLoS ONE. 2009;4(12):e8327.
4. Good IJ. Probability and the Weighing of Evidence. London: Griffin; 1950.
5. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, et al. Validating TrueAllele® DNA mixture interpretation. J Forensic Sci. 2011;56(6):1430-47.
6. Cover TM, Thomas JA. Elements of information theory. 2nd ed. Hoboken, N.J.: Wiley-Interscience; 2006.
7. Perlin MW. Explaining the likelihood ratio in DNA mixture interpretation. Promega's Twenty First International Symposium on Human Identification; San Antonio, TX2010.
8. Bauer DW, Butt N, Hornyak JM, Perlin MW. Validating TrueAllele® Interpretation of DNA Mixtures Containing up to Ten Unknown Contributors. J Forensic Sci. 2020;25(2):380-98.
9. Perlin MW. Efficient construction of match strength distributions for uncertain multi-locus genotypes. Heliyon. 2018;4(10):e00824.
10. Royall R. On the probability of observing misleading evidence. J Am Stat Assoc. 2000;95(451):760-8.
11. Perlin MW, Dormer K, Hornyak J, Schiermeier-Wood L, Greenspoon S. TrueAllele® Casework on Virginia DNA Mixture Evidence: Computer and Manual Interpretation in 72 Reported Criminal Cases. PLoS ONE. 2014;9(3):e92837.
12. Perlin MW, Hornyak J, Sugimoto G, Miller K. TrueAllele® Genotype Identification on DNA Mixtures Containing up to Five Unknown Contributors. J Forensic Sci. 2015;60(4):857-68.
13. Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework Validation Study. J Forensic Sci. 2013;58(6):1458-66.

Figure. *A genotype's LR distributions*. The distribution of Noncontributor log(LR) values is shown on the left. The Contributor distribution is on the right.