Department of Forensic Biology
Charles S. Hirsch Center for Forensic Sciences
421 East 26th Street, New York, NY 10016
Telephone: 212-323-1200 Fax: 212-323-1590
Email: DNALab@ocme.nyc.gov
Official Website: www.nyc.gov/ocme

DATE: December 23, 2020

FROM: Craig O'Connor
Assistant Director / Technical Leader, Nuclear DNA Operations

SUBJECT: Risk Assessment: Underperformance of the HPD method in STRmix™

This Risk Assessment involves the disclosure by the developers of STRmix™ concerning the underperformance of the HPD method in the program. Several steps were taken by the Department of Forensic Biology to evaluate the impact, if any, on reported results and conclusions.

The OCME has been using the software STRmix™ since January of 2017 as an assistive tool for deconvolution of DNA results and calculating likelihood ratios to provide statistical weight to comparisons of reference samples. The likelihood ratio (LR) is a statistical measurement of the strength of support for one hypothesis over another. Generally, for DNA comparisons, the two hypotheses are related to the probability of the evidence given that a particular person of interest is included as a contributor to the sample, rather than the probability of the evidence given that the person of interest is not included as a contributor.

The LR (like any statistic) is an estimate, and the true value is unknown and unknowable. The LR is reliant on assumptions and multiple estimated values. This includes estimated values for number of contributors, the level of co-ancestry within racial sub-populations, sampling uncertainty of the allele frequency databases used, and variability within the STRmix™ deconvolution process. This inherent variability is due to the fact that STRmix™ uses a Markov chain Monte Carlo (MCMC) process to analyze data for deconvolution. MCMC is a well-established and widely used random re-iterative process which allows the software to consider millions of possibilities each time it runs to find the best fit that explains the data. At the end of the process, genotype weights for each of the contributor(s) are generated as an output. Because it is a random process, each time the same set of data is run through the software, the results of the genotype weights may be slightly different.

During the MCMC deconvolution process, the software calculates what is called an Effective Sample Size (ESS). This is an estimate of the number of independent iterations that were sampled within the overall run. To account for this variability within the MCMC process, a function called "MCMC uncertainty" can be turned on within the STRmix™ software during the likelihood ratio calculation. With this function turned on, the software will create distributions for each of the genotype weights using an effective count based on the ESS. That effective count is an estimate of how many independent iterations were spent considering that particular

genotype combination[1]. Generally, the higher the genotype weight (more certain) the less variability is seen, and the lower the genotype weight (less certain), the more variability may be seen. Using this distribution for each of the weights, a distribution of (1,000) likelihood ratios is calculated, and a lower bound for this distribution is what is output on the STRmix™ report. The OCME applied this function during the validation of the software before implementation and has continued to use this function in casework since 2017. Within the software, the Laboratory chose the 99% 1-sided Highest Posterior Density as its lower bound, meaning that of the 1,000 LRs calculated by the software, the 10th smallest LR is output on the STRmix™ report[2].

On June 30, 2020, a notification from the developers of the STRmix™ software was sent to STRmix™ users which stated the following:

> "We have tested the highest posterior density (HPD) method for applying a lower bound to the variation induced by the Monte Carlo effect in the probabilistic genotyping system, STRmix™ (not including allele frequency uncertainty or uncertainty in theta). Tests show that the approach is not giving the desired 99% coverage. The lowest coverage observed was 76%.
>
> Although less effective than desired, this method does provide a layer of conservatism that is additional to the other layers within the LR assignment including the population genetic model and values of theta. Overall, the LR produced will be strongly conservative."

This indicates that, while the calculations are programmed into the software as intended, in recent testing of the method that is used within the software to generate the distribution of genotype weights for calculation of the 1-sided 99% Highest Posterior Density process, is not providing the desired coverage. The previous expectation was that it was 99% probable that the true LR value for a comparison is greater than the reported 1-sided 99% HPD LR. While this disclosure indicates that the method is occasionally underperforming, the range of their observed coverage in their testing was between 76%-100%, so the STRmix™ software is still generating a likelihood ratio that is most often, lower than the true LR[3,4].

In addition to the MCMC uncertainty feature, multiple additional layers of conservativism are applied to the likelihood ratio calculation within the software and by OCME standard reporting protocols to give a conservative[5] lower bound best estimate of the LR. These include[1,6]:

---

[1] STRmix™ V2.4 User's Manual. Institute of Environmental Science and Research Limited. Issued to NYC OCME 31 March 2016.

[2] The OCME reports the lowest Unified LR from the four population sub-groups that are calculated. This value incorporates the 1-sided 99% highest posterior density method as described, and additionally accounts for potential relatives of the person of interest within the unknown population.

[3] Technical Report: The highest posterior density for the Monte Carlo effect in STRmix™. https://www.strmix.com/news/technical-report-2/

[4] Bright, JA et al. *Testing methods for quantifying Monte Carlo variation for categorical variables in Probabilistic Genotyping.*

[5] Conservative is defined in this context as an LR that tends to be lowered in support of a particular comparison to a person of interest.

[6] NYC OCME Standard Operating Procedures: STR Results Interpretation, PowerPlex® Fusion & STRmix™

1. Accounting for possible co-ancestry within racial sub-populations with the application of theta (θ) using the Balding-Nichols subpopulation model
2. Comparing the person of interest to all components of a DNA mixture and applying a factor in order to account for the possibility that the POI could be any of the potential contributors (and not just a single particular contributor, i.e. major/minor)
3. Generating a distribution of allele frequencies to account for allele sampling uncertainty
4. Reporting the lowest of four different racial populations calculated
5. Modeling for possible relatives of the person of interest within the population of unknown people within the alternate hypothesis

Because all of these layers of conservatism are applied to the likelihood ratio calculation, the number reported is still very likely to be lower than the unknown (and unknowable) true LR. All likelihood ratios for each of these layers are included within the STRmix™ report, and therefore within each casefile when an LR is calculated.

<u>The following steps were taken by the laboratory after notification from the developers:</u>

- A data review was performed on a set of samples from the original validation with the MCMC uncertainty function turned off and compared to previous results with the function turned on. The results demonstrated that applying the MCMC uncertainty function applies a layer of conservatism that will most often lead to the same or lower LRs for a given comparison. This data review will be shared with our customers.

  A memo was issued to customers which discussed the disclosure from the developers and the OCME response. The memo was also presented and discussed, along with a Microsoft® PowerPoint® presentation, with both prosecutors (July 2020) and defense attorneys (August 2020) during scheduled customer meetings with the laboratory management. As the labeling of the 1-sided 99% HPD on the STRmix™ report within the casefile is now known to be a misnomer, this memo has also been linked within the report appendix and has been placed on the NYC OCME public-facing website (https://www1.nyc.gov/assets/ocme/downloads/pdf/mcmc_uncertainty.pdf.).

- When a likelihood ratio is reported by the OCME, the value is stated as follows:

  > "The DNA mixture found on evidence sample is approximately [LR] times more probable if the sample originated from [hypothesis 1] than if it originated from [hypothesis 2]. Therefore, this supports that [reference DNA profile] is [included or excluded] as a contributor to this sample."

  The use of the word '***approximately***' within the reporting language accurately reflects that the likelihood ratio calculated is an <u>estimate</u>.

The OCME also includes an appendix within each issued report that describes various terms and testing procedures used by the laboratory. This appendix includes a definition of the LR:

> **Likelihood ratio (LR)** - A statistical measurement of the strength of support for one hypothesis over another. For example, this would be reported as "The DNA mixture found on evidence sample is approximately LR times more probable if the sample originated from hypothesis 1 than if it originated from hypothesis 2. Therefore, this supports that reference DNA profile is [included or excluded] as a contributor to this sample." The likelihood ratio is calculated using the STRmix™ software and the value reported is a 99% 1-sided highest posterior density (HPD). For an LR supporting inclusion, this means that it is 99% probable that the true LR value for a comparison is greater than the reported LR. For an LR supporting exclusion, this means that it is 99% probable that the true LR is less than the reported LR. The LR is reported in words and its respective exponential value.

After notification from the developers, the report appendix was updated as follows (highlighting added to indicate change):

> **Likelihood ratio (LR)** - A statistical measurement of the strength of support for one hypothesis over another. For example, this would be reported as "The DNA mixture found on evidence sample is approximately LR times more probable if the sample originated from hypothesis 1 than if it originated from hypothesis 2. Therefore, this supports that reference DNA profile is [included or excluded] as a contributor to this sample." The likelihood ratio is calculated using the STRmix™ software and the value reported is an estimate that accounts for possible co-ancestry between individuals within racial sub-populations, sampling uncertainty of the allele frequency databases, and variability within the STRmix™ deconvolution process. For further information in regards to the 1-sided 99% Highest Posterior Density, refer to https://www1.nyc.gov/assets/ocme/downloads/pdf/mcmc_uncertainty.pdf . The LR is reported in words and its respective exponential value.

The modification to the reporting appendix was made to more accurately reflect the estimate associated with the reported likelihood ratio, and to not over-state the level of specific coverage based on the recent disclosure from the developers.

- The protocol within the Laboratory "STRmix™ Glossary" that contained a definition of the 1-sided 99% HPD was revised to remove specific language on the coverage of the method.

- Laboratory staff were queried as to whether they had ever been asked in testimony to discuss the details of the 1-sided 99% HPD. None were identified. Laboratory staff were also notified of the disclosure from the developers, and the updated change to the report

appendix, and given advisement to review the memo and disclosures before any pre-trials or testimony.

- A joint statement from the members of the NY State Biology Technical Working Group (BIOTWG) that are using the STRmix™ software within casework, was presented to the NY State Crime Laboratory Advisory Committee (NYCLAC) and the DNA Sub-Committee in August 2020.

Before implementation at the OCME in 2017, developmental validation by the software developer[7], and internal validation by the Laboratory was performed[8]. The disclosure from the developers does not negate the results of these initial validation studies, nor the reliability of the STRmix™ software. The likelihood ratios produced by the software performed as expected within specificity, sensitivity, precision, and repeatability studies for both true and non-contributors. In addition, the software developers have indicated that there is no replacement method that they have yet identified to improve upon the current coverage for the MCMC uncertainty function[9].

After evaluating the potential impact of the disclosure and any potential risk involved, the OCME remains confident that the STRmix™ software is producing conservative, reliable estimates of likelihood ratios. As there is no currently identified replacement method to account for MCMC uncertainty within the software, the Laboratory is choosing to continue the use of the current function, as it has been shown to still provide a layer of conservativism to the calculation, despite not always achieving the desired 99% coverage. As new software versions are released by the developers, the Laboratory will consider upgrading if a new method is implemented within the program.

[7] Bright, JA et al. *Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles.* Forensic Science International: Genetics, Volume 23, July 2016, Pages 226-239.
[8] Internal Validation of STRmix™ V2.4 for Fusion NYC OCME. 18 November 2016. Updated 20 December 2019. https://www1.nyc.gov/assets/ocme/downloads/pdf/STRmix-V2-4-Fusion-5C-Validation%20Summary.pdf
[9] FAQs about HPD underperformance. STRmix™ Technical and Scientific Support Site.