

## **Unintended Catalyst: the Effects of 1999 and 2001 FBI STR Population Data Corrections on an Evaluation of DNA Mixture Interpretation in Texas**

### 1. FBI Data Corrections: What Do They Mean?

In May 2015, the Federal Bureau of Investigation (“FBI”) notified all CODIS laboratories it had identified minor discrepancies in its 1999 and 2001 STR Population Database. Laboratories across the country have used this database since 1999 to calculate DNA match statistics in criminal cases and other types of human identification. The FBI attributed the discrepancies to two main causes: (a) human error, typically due to manual data editing and recording; and (b) technological limitations (*e.g.*, insufficient resolution for distinguishing microvariants using polyacrylamide gel electrophoresis), both of which were known limitations of the technology. The FBI has provided corrected allele frequency data to all CODIS laboratories.

In May and June 2015, Texas laboratories notified stakeholders (including prosecutors, the criminal defense bar and the Texas Forensic Science Commission) that the FBI allele frequency data discrepancies were corrected. The immediate and obvious question for the criminal justice community was whether these discrepancies could have impacted the outcome of any criminal cases. The widely accepted consensus among forensic DNA experts is the database corrections have *no impact* on the threshold question of whether a victim or defendant was *included or excluded* in any result. The next questions were whether and to what extent the probabilities associated with any particular inclusion changed because of the database errors.

The FBI conducted empirical testing to assess the statistical impact of the corrected data. This testing concluded the difference between profile probabilities using the original data and the corrected data is less than a two-fold difference in a full and partial profile. Testing performed by Texas laboratories also supports the conclusion the difference is less than two-fold. For example, in an assessment performed by one Texas laboratory, the maximum factor was determined to be 1.2 fold. In other words, after recalculating cases using the amended data, the case with the *most substantially affected* Combined Probability of Inclusion/Exclusion (“CPI”)<sup>1</sup> statistical calculation (evaluated for a mixed sample) changed from a 1 in 260,900,000 expression of probability to a 1 in 225,300,000 expression of probability.

Amended allele frequency tables are publicly available for anyone to compare the calculations made using the previously published data and the amended allele frequencies, though expert assistance may be required to ensure effective use of the tables.<sup>2</sup>

### 2. The Impact of FBI Database Errors on DNA Mixture Interpretation Using CPI

As part of their ongoing commitment to accuracy, integrity and transparency, many Texas laboratories offered to issue amended reports to any stakeholder requesting a report using the corrected FBI allele frequency data. Some prosecutors have submitted such requests to laboratories, particularly for pending criminal cases. As expected, the FBI corrected data have not had an impact exceeding the

---

<sup>1</sup> The Combined Probability of Inclusion/Exclusion is commonly referred to as either “CPI” or “CPE.” They are referred to jointly in this document as “CPI” for ease of reference.

<sup>2</sup> <https://www.fbi.gov/about-us/lab/biometric-analysis/codis/amended-fbi-str-final-6-16-15.pdf>

two-fold difference discussed above. However, because analysts must issue *signed amended reports* with the new corrected data, they may only issue such reports if they believe *the analyses and conclusions in the report comply with laboratory standard operating procedures*. For cases involving DNA mixtures, many laboratories have changed their interpretation protocols and related procedures using CPI. To reiterate, changes in mixture interpretation protocols are unrelated to the FBI allele frequency data corrections discussed above. However, when issuing new reports requested because of the FBI data corrections, the laboratory's use of current mixture protocols may lead to different results if the laboratory had a different protocol in place when the report was originally issued. Changes in mixture interpretation have occurred primarily over the last 5-10 years and were prompted by several factors, including but not limited to mixture interpretation guidance issued in 2010 by the Scientific Working Group on DNA Analysis ("SWGDM").

The forensic DNA community has been aware of substantial variance in mixture interpretation among laboratories since at least 2005 when the National Institute of Standards and Technology ("NIST") first described the issue in an international study called MIX05. Though NIST did not expressly flag which interpretation approaches were considered scientifically acceptable and which were not as a result of the study, it has made significant efforts to improve the integrity and reliability of DNA mixture interpretation through various national training initiatives. These efforts have ultimately worked their way into revised standard operating procedures at laboratories, including laboratories in Texas. Based on the MIX05 study, we know there is variation among laboratories in Texas and nationwide, including differences in standards for calculation of CPI that could be considered scientifically acceptable. However, we also know based on a recent audit of the Department of Forensic Sciences ("DFS") in Washington, DC that some of the "variation" simply does not fall within the range of scientifically acceptable interpretation. This finding does not mean laboratories or individual analysts did anything wrong intentionally or even knew the approaches fell outside the bounds of scientific acceptability, but rather the community has progressed over time in its ability to understand and implement this complex area of DNA interpretation appropriately.

While in many cases the changed protocols may have no effect, it is also possible changes to results may be considered material by the criminal justice system, either in terms of revisions to the population statistics associated with the case or to the determination of inclusion, exclusion or an inconclusive result. The potential range of interpretive issues has yet to be assessed, but the potential impact on criminal cases raises concerns for both scientists and lawyers. We therefore recommend any prosecutor, defendant or defense attorney with a currently pending case involving a DNA mixture in which the results could impact the conviction consider requesting confirmation that CPI was calculated by the laboratory using current and proper mixture interpretation protocols. If the laboratory is unable to confirm the use of currently accepted protocols for the results provided, counsel should consider requesting a re-analysis of CPI.

The Texas Forensic Science Commission is currently in the process of assembling a panel of experts and criminal justice stakeholders to determine what *guidance and support* may be provided to assist Texas laboratories in addressing the challenging area of DNA mixture interpretation. In particular, a distinction must be made between acceptable variance in laboratory interpretation policies and protocols and those approaches that do not meet scientifically acceptable standards. An emphasis on statewide collaboration and stakeholder involvement will be critical if Texas is to continue to lead the nation in tackling challenging forensic problems such as those inherent in DNA mixture interpretation.