DNA mixture evidence and the need for accurate match statistics

Mark W. Perlin, PhD, MD, PhD, Cybergenetics, Pittsburgh, PA USA

ABSTRACT

Criminal justice relies on DNA evidence to convict or acquit the accused, and to protect the public from crime. In forensic science, DNA enjoys an unparalleled reputation for infallibility. But when DNA data is incorrectly interpreted, the resulting match statistics can be inaccurate.

DNA evidence is usually a mixture of two or more people. The molecules can be degraded or present in small amounts. Resulting laboratory data may require modern statistical analysis for accurate interpretation. Unvalidated statistical methods need not be reliable. And unreliable DNA reporting of forensic data can lead to unjust outcomes.

Forensic guidelines do not require crime laboratories to validate their DNA mixture interpretation. Laboratory analysts often apply "thresholds" that discard data, but the accuracy of threshold procedures has not been scientifically proven. Altering signals before entry into statistical software can lead to inaccurate results. By omitting informative DNA data, an "inconclusive" report can deny courts evidence that could implicate the guilty or exonerate the innocent.

Adjusting laboratory data can introduce human subjectivity. There is a danger that contextual bias (such as inadvertently assuming guilt) can yield a DNA analysis that is not impartial. Some mixture interpretation protocols do not use all the DNA data. Data selection can overstate the probative value of a match, which can mislead juries.

Simplifying complex data can cause DNA interpretation errors. Simple methods are appropriate for simple DNA data. But their application to more challenging samples must be empirically justified before they can relied upon. Without supporting validation data, an unsubstantiated interpretation method can taint DNA evidence in criminal cases.

Ten years ago NIST and others warned forensic practitioners about mixture interpretation issues. Since then, a decade of unsophisticated data analysis has led to hundreds of thousands of mixtures with inaccurate match statistics. This realization has recently shut down crime laboratories (e.g., Washington, DC) and necessitated extensive DNA evidence review (e.g., 24 thousand cases in Texas).

The CPI workhorse is a subjective one-sided match statistic, unrelated to identification information, raising doubt about DNA infallibility.

Victims and defendants need DNA justice. Mixture evidence, past and present, must be reviewed in an unbiased and scientifically valid way. Accurate DNA match statistics ensure conviction integrity, and maintain public trust in criminal justice.

MIXTURE CRISIS

Scientists and statisticians write about the DNA mixture failure. They contend that thresholds lack a scientific foundation. They find that *combined probability of inclusion* (CPI) statistics for low-level mixtures with little DNA could be unfair to defendants. They question whether CPI even makes any sense as a match statistic.

There is concern about *human bias* in the CPI method, and producing *subjective results* that were suspect-centric or pro-prosecution. A human analyst first adjusts the data (applying thresholds, removing apparent stutter, etc.), and then looks at the defendant's genotype to decide if the person is included in the mixture. Only after first changing the data and assuming inclusion does the analyst then run CPI software to calculate a match statistic, a number often used in court to help establish guilt. Assuming guilt to establish guilt is circular reasoning.

There is *bias when an analyst subjectively picks data* by choosing loci after first looking at the defendant's genotype. One report showed how analysts could justify including any "Tom, Dick or Harry" who was not actually in the DNA evidence. In another study, analysts who had the "potentially biasing context" that their corroborating DNA evidence "was essential to the prosecution" did not exclude a defendant from a mixture; however, without such context, only 1 of 17 other DNA examiners agreed, while 16 "reached a different and conflicting conclusion" (12 exclude, 4 inconclusive). Most mixture interpretation software requires an analyst to prepare the input by first selecting a subset of their data.

In their oft cited 2009 *"cartoon" paper*, the FBI proposed a solution – since one threshold failed, use two thresholds. They introduced a second "stochastic threshold" at a higher level to discard data that might have too much variation. There *was no statistical theory or empirical data* to support this unfounded proposal, just cartoon drawings. *No validation studies* were done to establish accuracy. Sophisticated mathematics can model data variation, but applying another simplistic threshold simply discards more data.

Regardless, the FBI's SWGDAM 2010 guidelines imposed stochastic thresholds on crime laboratories, making the cartoon paper *de facto* national policy. The labs compliantly determined these thresholds, and applied them to mixture evidence. The second threshold greatly decreased their match statistics and increased inconclusive outcomes, eliminating needed DNA information.

FORENSIC FAILURE

A 2011 TrueAllele® validation study conducted jointly with the NYSP DNA lab (Albany, NY) showed that *CPI vastly underreported DNA's probative value*. Whenever the lab was able to report a CPI statistic, their number was (on average) a million times less than the true match statistic on the same data. CPI analysis removed considerable DNA information.

A 2013 NYSP validation study examined how human mixture analysis performed on data where the TrueAllele computer produced a match result. TrueAllele's median match statistic was around a quadrillion. When TrueAllele gave a result, 70% of the time thresholds failed to report any match statistic. Human review was silent about most DNA evidence, *incorrectly concluding that informative items were inconclusive*.

In 2013, NIST conducted a MIX13 inter-laboratory study. The hope was that the new stochastic threshold procedure had adequately addressed natural data variation. Their hope went unrealized when a hundred participants examined a three person mixture that did *not* contain a particular suspect. Seventy groups *incorrectly included* this suspect, whose DNA was not present in the mixture (70% false match rate), giving irrelevant DNA match statistics that ranged from 9 to 344,000. Twenty four labs found the comparison inconclusive. Only six correctly excluded the suspect (6% accuracy rate), with one of them using Cybergenetics TrueAllele method.

A 2014 TrueAllele validation paper conducted on 72 Virginia mixture cases showed *the extent of CPI's lost information*. On a hundred DNA comparisons, the average TrueAllele match statistic of a hundred billion (10^{11}) dropped to only millions (10^{6}) when a threshold was applied and CPI calculated. Applying a second (stochastic) threshold to the same mixture data further reduced the modified CPI statistic to just hundreds (10^{2}). Moreover, the SWGDAM 2010 procedure did not eliminate all false matches.

In 2015, comparison of inclusion probability with TrueAllele match information showed that *CPI is a one-sided random number generator, uncorrelated with identification information*. The subjective CPI statistic depends on the number of loci tested, not on the probative value of the DNA evidence. That is why (using all loci) CPI always gave the same answer – around a million – regardless of the data. After an analyst first decides that a defendant's DNA is in a mixture (viewed as guilt by a jury), CPI can afterwards provide an impressive statistic that only restates a human judgment.

SCIENCE AND THE LAW

DNA holds considerable prejudicial sway over a jury. In a courtroom, the three letters can seem to abbreviate "Do Not Acquit." When DNA match statistics are routinely wrong or lack probative value, it is hard to justify introducing them in criminal trials. The Federal Rules of Evidence (FRE) provides *legal mechanisms for excluding harmful DNA evidence* from court.

FRE Rule 403 permits a court to "exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence."

A CPI match statistic essentially counts up the number of loci deemed an "inclusion" by a human analyst. The statistic is cumulative evidence that reframes an analyst's subjective conclusions as an objective-sounding match number that can mislead a jury. Since CPI is uncorrelated with identification information, it has little probative value. *Mixture statistics that are more prejudicial than probative can be challenged* in a pretrial hearing to keep the jury from hearing unfair DNA results.

Rule 702 guides who can testify as an expert witness to render a scientific opinion about DNA evidence. The expert's testimony must be based on reliably applying a reliable method to sufficient data. After a pretrial hearing, a judge can exercise their gatekeeper role to protect the jury from hearing unreliable scientific evidence.

Unreliable DNA match statistics are susceptible to challenge. A judge may rule that inaccurate or insufficiently validated DNA mixture statistics are not admissible. *Challenging unreliable DNA interpretation can keep out bad evidence*, even when there is good underlying data.

In *Brady v. Maryland* the Supreme Court of the United States held "that the suppression by the prosecution of evidence favorable to an accused upon request violates due process where the evidence is material either to guilt or to punishment." This ruling applies "irrespective of the good faith or bad faith of the prosecution" because "society wins not only when the guilty are convicted, but when criminal trials are fair."

When mixture interpretation fails, no DNA match statistic is reported. *The absence of a report can hide potential exculpatory DNA evidence*. But if a defendant requests all data from all laboratory testing, Brady requires the government to provide that data. Effective interpretation of the government's DNA data by an independent expert might exonerate the accused, or implicate another person.

RECOMMENDATIONS

The following recommendations may help society move beyond mixture interpretation failure, and enjoy consistently more reliable DNA evidence:

1. *Open DNA data to public scrutiny*.

The crime labs have failed to produce reliable match statistics for over fifteen years. The solution is open access to all DNA data, so that impartial scientists can publicly reassess crime lab results in every case.

2. *Revisit all past DNA mixture cases*.

Hundreds of thousands of DNA mixtures have been improperly interpreted. Only an unbiased, accurate software review of all this evidence can rectify the problem.

3. *Educate trial attorneys and judges*.

Law attracts many who would rather not study science or mathematics. But lawyers need to understand the evidence they attack or defend. Appropriate education is needed to teach them DNA statistics.

4. *Fully automate mixture interpretation*.

Human analysts are trained to remove DNA data from the input to their interpretation software, which introduces bias and error. Automated computing can help eliminate such human decision-making.

5. *Extensively validate DNA interpretation*.

Most mixture statistics have not been validated for their intended use. No method, whether done by man or machine, should ever be introduced as evidence without supporting validation.

6. *Keep methods within their limits*.

Defense vigilance helps ensure that crime labs stay within the bounds of their validated interpretation methods. Without this DNA pressure, false positives may falsely identify or convict innocent people.

7. *Go beyond laboratory limits*.

Better interpretation methods can solve DNA mixtures that crime labs cannot. Independent groups should interpret these data. Otherwise false negatives may fail to identify, withholding potentially exculpatory evidence.

CONCLUSION

Unscientific, untested "statistical" analysis of DNA mixtures has led to incorrect results on hundreds of thousands of evidence items. When thresholds give an "inconclusive" result on mixtures with data, that silent non-answer is usually wrong. When CPI match statistics are reported, again the answer is usually wrong.

Innocent people remain in prison because informative DNA wasn't used in their defense. Defendants are wrongfully convicted when misinterpreted DNA can't identify the true culprit. Perpetrators go free when DNA evidence is failed by forensic statistics. Freed criminals then commit more crime, which DNA should have prevented, needlessly harming innocent victims. This is not the fairest justice that DNA science can provide.

Modern genotyping programs use probability to help interpret DNA mixtures. TrueAllele has a fully Bayesian model that considers all data and all solutions. Less thorough programs remove data to simplify the problem using thresholds, dropout parameters or peak filters.

Subjective programs let a human operator choose input data and parameters to overcome software limitations. While crime labs have started adopting better match statistic software, validation studies are needed to determine their range of applicability.

Unfounded DNA statistics have inflicted considerable injustice on defendants, crime victims, and society. Every case that involved inconclusive DNA mixtures or unfounded match statistics should be revisited. It is time to rectify two decades of forensic failure with accurate, objective, and validated DNA interpretation.

WHITE PAPER

Perlin, M.W. "Failing to interpret DNA mixture evidence." Cybergenetics White Paper, July, 2016.

This 4,000 word report cites 48 references and is freely available from Cybergenetics website.

https://www.cybgen.com/information/report/2016/CYB/Perlin-Failing-to-interpret-DNA-mixture-evidence/page.shtml

The report is essential reading for:
- prosecutors using DNA mixture evidence
- defenders blocking unfounded DNA science
- judges deciding on DNA reliability
- police needing accurate DNA information
- journalists writing about forensic science
- scientists reporting on DNA mixtures

Cybergenetics