

INCLUSION PROBABILITY IS A LIKELIHOOD RATIO: IMPLICATIONS FOR DNA MIXTURES

Mark W. Perlin, PhD, MD, PhD

Cybergenetics, Pittsburgh, PA USA

There has been much discussion recently amongst forensic scientists about the relative merits of inclusion and likelihood ratio (LR) methods for interpreting DNA mixtures. Advocates for the probability of inclusion (PI; also termed CPI, RNME or CPE) approach contend that it is a simpler statistic that is easier to explain in court. LR enthusiasts rejoinder that theirs is a more informative method that preserves more of the identification information present in the DNA data. The debate implicitly assumes that there is some essential difference between PI and LR, suggesting that each perspective should be understood and evaluated on its own merits.

In fact, there are many different LR statistics for DNA mixture interpretation. And PI happens to be just one of them. However, amongst all currently used LRs, the PI version does have a special distinction – it is the least informative.

Recognizing that PI is just another LR has important consequences for forensic science practice.

1. The current PI vs. LR controversy can be finally put to rest.
2. Inclusion efficacy can be measured in terms of how well it preserves the data's identification information. The logarithm of the LR is a standard information measure, and PI is a LR, so this assessment is easily accomplished.
3. The inclusion method can be supported in court based on its scientific status as a valid LR.
4. The PI statistic can be better understood through the inclusion likelihood function used in its LR construction.
5. The relevance of PI can be challenged on particular DNA evidence by examining the appropriateness of its (inclusion likelihood) modeling assumptions for that data.

In the paper, we show by construction that PI is a LR. We first describe the inclusion likelihood function, and see how it naturally explains binary allele data. We next use Bayes theorem to form the inclusion genotype, represented by its probability mass function (pmf). Using an easily understood form of the LR (genotype probability gain), we then insert the inclusion genotype pmf into this LR expression to obtain the standard PI statistic. Having thus derived the PI as a LR, we then discuss what this result means for DNA mixture interpretation.

The talk visually explains the underlying concepts to the forensic practitioner, and uses no mathematics besides basic probability. We focus primarily on the forensic science implications of PI actually being a LR. This proper scientific foundation for PI may invite a re-examination of some prevalent DNA mixture interpretation practices.