



Correspondence

Fairness in evaluating DNA mixtures



In a recent letter [1], Presciuttini and Egeland argue that a mixture likelihood ratio only makes sense if the hypotheses compared agree as to number of contributors (N_C). But the argument is circular; it proves nothing. The ISFG committee recommendation [2] with which they disagree is correct, though it has been widely ignored at great risk of unjust DNA verdicts. Therefore it is timely that [1] brings attention to the issue.

[1] adopts the Bayesian framework of assuming a prior probability, formulated as

$$\Pr(H_p | \mathbf{I}) / \Pr(H_d | \mathbf{I})$$

where " \mathbf{I} " stands for all information that is available about the case" other than DNA. For the present discussion I accept this framework provided that by "information" we mean facts that are known and agreed. Then it does seem right that \mathbf{I} is the same for H_p as for H_d . But where Presciuttini and Egeland next assert that N_C is included in \mathbf{I} , the argument given relies on the reader not paying attention. It confuses the words parameter and information. A parameter means a number that describes a real quantity, with no implication that its value is known to us. They claim that N_C is a parameter – i.e. that for every mixture there is some particular number N_C . Suppose so. But N_C isn't known, as the authors obviously agree since they refer to "estimating" it. Being unknown, it isn't information. But they nonetheless incorporate it into the information \mathbf{I} by mistake, baldly inserting a premise equivalent to the desired conclusion. That's circular reasoning.

That the argument in the letter is wrong doesn't prove that the conclusion is wrong, but it is. A moment's serious thought refutes the view that N_C is even a well-defined real-world number. We can formulate bright-line definitions: "The number of people who contributed at least one DNA molecule" – but that's not a practical

or useful meaning to ascribe to N_C . Practically speaking what we intend to count by N_C is something like "significant" contributors – but that's a vague definition and consequently it should not be a surprise that for some mixtures the number N_C varies depending on one's point of view – in particular as I plan in a forthcoming article to illustrate, often depending on whether the suspect is included or not. To be clear: it is not only "conceivable that the N_C value that maximizes the H_p likelihood is different from the value that maximizes the H_d likelihood" as [1] comments, but during development and testing of Mixture Solution I have run into many cases where it happens. For one example, comparing 4-person explanations gives good evidence supporting H_p , comparing 3-person explanations gives the opposite, powerful evidence supporting H_d , and the only fair comparison – 4-person H_p versus 3-person H_d – gives a resounding "inconclusive." It follows that most mixture programs, which are limited by the premise that H_p and H_d must agree on the N_C , risk a miscarriage of justice in such cases. From my experience such cases are common.

References

- [1] S. Presciuttini, T. Egeland, About the number of contributors to a forensic sample, *Forensic Sci. Int. Genet.* 2016 (25) (2016) e18–e19, doi:<http://dx.doi.org/10.1016/j.fsigen.2016.08.005>.
- [2] P. Gill, C.H. Brenner, et al., DNA commission of the International Society of Forensic Genetics: recommendations on the interpretation of mixtures, *Forensic Sci. Int.* 160 (2006) 90–101.

Charles H. Brenner
DNA-VIEW & UC Berkeley Human Rights Center, United States E-mail
address: charles@dna-view.com (C. Brenner).

Received 18 October 2016

Available online 5 December 2016