2012-08-05

Response to: DR AS 5388.3 Forensic analysis - Part 3 - Interpretation

Response prepared by:

Prof David Balding

I am Professor of Statistical Genetics at University College London, and a member of the steering committee of UCL's Centre for Forensic Science. I hold the degrees of B Math (Hons class 1, University Medal) from U Newcastle (Aust) and D Phil from Oxford (UK). I am a Fellow of both the Royal Statistical Society and the Society of Biology. I have about 19 years experience of occasional expert witness work, predominantly in connection with the statistical interpretation of DNA evidence, mostly in the UK but also in Australia, Ireland and the USA. I have published relevant articles in genetics, statistics, law, and forensic science journals, as well as a monograph "Weight-of-evidence for forensic DNA profiles" (Wiley, 2005).

I have read the response to DR AS 5388.3 prepared by Dr Geoffrey Stewart Morrison of UNSW and endorsed by many distinguished international scholars. I am in broad agreement with that submission, and in particular its assertion that, if adopted, the draft Standard will support and encourage much bad practice. It will also stymie the progress towards better forensic evidence evaluation that the US National Research Council's 2009 report highlighted as so greatly needed. I concur with the suggestion of Morrison and co-signatories that having no written standard would be preferable to having this Standard as currently drafted.

Like those writers, I am surprised that the draft Standard appears to draw nothing from the revolution in forensic evidence interpretation that has followed from the advent of DNA evidence about 20 years ago, nor from the landmark US NRC report. In particular I highlight the sentence (Sec 4.1):

"The examiner should follow an approach that seeks to evaluate the null hypothesis."

The inadequacy of this approach for the analysis of DNA evidence has led to it being purged from most serious discussion of evidence interpretation in recent years. The forensic examiner needs to compare competing hypotheses, one corresponding to the prosecution position and the other to that of the defence. There may be many reasonable alternatives to the prosecution's allegation, and hence multiple defence hypotheses; it is their cumulative weight that matters, not the weight of any one of them (for example, if the DNA did not come from the defendant as alleged, it may have come from his brother, a cousin, or another man of similar ethnic origin; also different numbers of contributors to the crime sample may need to be considered). Moreover, it is crucial to evaluate the consistency of the evidence with the prosecution case: for degraded DNA samples and many other evidence types, there can be much variation in this. Only the likelihood ratio approach captures these crucial desiderata. Its success in the evaluation of DNA evidence has led to widespread support for this more rigorous and fair approach to be adopted for other evidence types. While the process of moving towards that goal is far from complete, it should be recognised as the standard that forensic science is aiming towards.

While reliance on professional judgment cannot be completely eliminated, I am alarmed at the emphasis given to it throughout the draft Standard. It forms essentially the only topic under the heading "The role of the examiner" (Sec 4.2). This is the bad old forensic science criticised by the US NRC and that we should be trying to move away from. The forensic examiner should strive to minimise the role of professional judgment and to use scientific methods instead: quantitative methods drawing on data from

appropriate databases and experimental results.

There follows comments on some other specific statements in the draft Standard. These should be regarded as illustrative and not an exhaustive list of the unsatisfactory statements in the draft.

Sec 4.1 “Interpretation includes answering the question as to whether or not ...” The role of the forensic examiner is not to answer these questions, but to evaluate evidential support for the corresponding assertions, relative to the support for their negations.

Sec 4.3 “In qualitative examinations, variation exists but is not usually susceptible to quantitative analysis that yields an uncertainty value.” By definition qualitative analysis is not quantitative so the statement has little content. The point should be that examiners should strive towards quantitative analysis, making the methods, assumptions and background data available to the court as far as possible. Qualitative analysis is sometimes unavoidable but its limitations should be stressed.

Sec 5.1 “(c) Consider the probability of error.” This phrase seems to imply that the examiner is making a yes/no decision, in which case it is relevant to talk of probability of error. But the examiner should avoid making decisions as far as possible, and instead evaluate evidential support for competing hypotheses that facilitates the court’s role in making the decisions. To this end s/he needs to consider measurement error and other sources of uncertainty, not the probability of error.

Sec 5.2 “Databases used to derive estimates for the likelihood of any hypothesis shall be validated.” This leaves open the question of what “validated” means. In DNA analysis such ideas have led to many spurious tests of “validity” that only give an aura of scientific respectability, without any real content. The term “validate” is widely used in forensic science but without definition it adds little to the discussion and generates confusion.

Sec 5.3 “If a common origin is supported, any differences observed shall be capable of being adequately explained.” This approach forms the basis of much bad forensic science, where the expert makes up her/his mind and then searches for reasons to explain away any discrepancies. Instead what is needed is a quantitative assessment of the plausibility of the evidence given the same-origin hypothesis, which needs to be compared with the corresponding assessment under defence hypothesis.

Sec 5.4 “The examiner shall also consider alternative explanations that may emerge during the investigation process.” Why is this important point made only in this specific context? It is central to all forensic evaluation.

Sec 6.1 “Where practical, the initial examiner should retain primary responsibility for the evaluation of the data.” This is a dangerous recommendation as it seems to preclude an important role for experts in forensic evaluation, who should ideally work with the subject-matter forensic expert (e.g. a chemist).

Appendix B. I question the emphasis given Type 1 and Type 2 error. As noted above, forensic experts should as far as possible avoid usurping the court’s role in making decisions; s/he should assist the court to make its decision. So measurement error and other sources of uncertainty are more important in forensic science than type 1 and type 2 errors.