Science and Justice 53 (2013) 251–252

Contents lists available at SciVerse ScienceDirect

Science and Justice

journal homepage: www.elsevier.com/locate/scijus

Editorial

Is forensic science the last bastion of resistance against statistics?

This editorial is based on a talk that I gave at the 6th European Academy of Forensic Sciences meeting where I attempted to encourage debate, reflection, and possibly change. Nothing I have written here is aimed at any one person. Over the last 15 years I have been involved in forensic research, education, and practice. This has given me considerable opportunities to see how the fruits of mine and my peers' research has affected, or changed, the way forensic scientists study, interpret, and present evidence. It is, therefore, rather disheartening to say that the impact appears to be low to negligible. I will briefly discuss my findings in three different fields with which I have had some interaction. These are: fingerprints; glass; and DNA.

How many of the following complaints sound familiar?

> "*It may well be possible to compile a statistical database for fingerprints, however there would need to be considerable work done to assess whether it is an achievable….*"

> "*A working analytical model will then need to be created to analyze the data and present it. This project may take years.*"

> "*The cost of research, training and then defending the new methodology will be significant.*"

> "*Then there is the question of who is asking for it?*"

All these statements came from a single article [1], in response to the R v T ruling. There has been significant, sustained, scientific effort in developing interpretation models, starting with Galton in 1892, up to the present day. There has been almost no uptake of these developments. I have communicated with Neumann and Champod, who are champions for advanced statistical methodology in this field. Champod (Pers. Comm.) pointed me towards significant interest in new models following the 2006 Mayfield review [2] and the 2011 UK Fingerprint Inquiry [3]. I am slightly cynical as interest in changing practice is substantially different from actually changing practice.

As with fingerprints, glass evidence is also a very mature evidence type. Unlike fingerprints, it is easily quantifiable through density, refractive index, or elemental composition measurements. Practitioners agree on standards and methods, and purpose-built equipment is commercially available. There are good statistical models and there has been considerable research not only on quantification, but also less tangible phenomena such as transfer and persistence. There is a long history of statistical interpretation of glass evidence, arguably starting with Evett's first paper on the topic [4]. What is the state of play in glass? In a recent collaborative testing exercise [5] 12.6% (14 out of 111) of laboratories presented the results of a Bayesian approach to evidence interpretation. Should I/we be happy that approximately 1 in 8 labs is using our work? Significant scientific efforts have been made to develop interpretation models. However, the predominant interpretation method has not progressed past a two sample *t*-test. There is minor uptake of the Bayesian approach, a methodology that has been well-developed and fairly stable for the last 10–15 years.

DNA is a comparatively new evidence type. DNA is often held up as the poster-child for statistical interpretation of evidence. For example the National Academies stated in the news release about its 2009 report, "*Nuclear DNA analysis has been subjected to more scrutiny than any other forensic discipline*" [6]. One would think that with such an endorsement that methods for DNA interpretation were beyond reproach. If this is the case, why then do I still encounter forensic scientists who: report, "*the evidence is consistent with the suspect being a contributor*"; report Random Man Not Excluded (RMNE) or Cumulative Probability of Inclusion (CPI) type calculations in (almost any) mixture case; ignore or deny the existence of population substructure when we know that both Hardy–Weinberg Equilibrium (HWE) and Linkage Equilibrium (LE) assumptions are demonstrably false; and refuse to report Low-Template (LT)-DNA cases when we know that issues such as contamination (drop in), allelic drop out, and stutter are a feature of standard case work where DNA is abundant. The assessment of DNA needs to be brutal. There has been a massive scientific effort in developing interpretation models and methods of interpretation and are routinely used in court. These methods have been subjected to more scrutiny, testing, and evaluation than those in any other field of forensic science. Why then do we have such hugely inconsistent uptake? many practitioners like RMNE/CPI because it is (supposedly) easy to explain to the court. As an expert presenting evidence to the court, I have an obligation to use the best scientific methods available to me, not the ones that are the easiest to explain. If you think the ease of explanation is acceptable practice, then I have question for you. Is explaining the likelihood ratio (or any other statistic) to the court harder than explaining how ICP-MS works? Is there a list of "acceptable black boxes" for the court?

I hope that I have convinced you that we might have something to discuss. I will, in turn, outline what I think are the causes behind our current state of affairs, and then discuss some potential solutions and ways forward. I share some of the blame myself, and on behalf of members of my own professional community. Many statisticians come across as simply too academic. That is, they are unable or unwilling to simplify explanation to facilitate better understanding. They may be unable or unwilling to find acceptable compromises to the "best" solution, and they may have different objectives to the forensic scientist. It is easy to be too disengaged from the problem of interest, or to far from the practical concerns and realities of casework. To be an effective and useful statistician requires one to engage with the specialists in that discipline, take time to understand the nature of the problem, and the issues that may affect the interpretation of the results. Statisticians who have not

1355-0306/$ – see front matter
http://dx.doi.org/10.1016/j.scijus.2013.07.001

testified, or who have testified under restricted circumstances, are generally unaware of court room proceedings and complexities. This leads to proposals of solutions which may not be allowed by the court.

Readers familiar with the latest research on the interpretation of DNA and fingerprints will know that statistical models for evidence interpretation are becoming increasingly complex, and computationally demanding. The literature describing these models is often written for completeness or validation, or for other experts. These methods are only useful if you have the software. Software for interpretation can be very expensive. There are reasons for this: it is extremely expensive to produce, or it may be the main source of income for some individuals/companies. However, the cost of a software license pales into insignificance when compared to a new ICP-MS. The cost of the license is only one part of the equation. When you factor in the number of caseworker hours spent doing tedious calculations, and the associated peer checking, then a piece of software that removes this becomes cheap by comparison. There is quite a lot of freeware, and there is an initiative by the International Society for Forensic Genetics to promote and support these efforts. There is some (legitimate) concern that freeware can be unsupported, unvalidated, and hard to use. This is true, but we could view this as a community effort. A good example of this is the operating system Linux which is the World's most dominant internet server platform, trusted for any number of e-Commerce transactions daily, without question or concern.

There have been several high-profile, and potentially damaging, rulings which criticise, or rule against, the use of Bayes' Theorem in court (R v Doheney and G. Adams [1997] 1 Cr App R. 369, R v D. Adams [1996] 2 Cr App R 467, R v D. Adams [1998] 1 Cr App R 377, R v T [2010] EWCA Crim 2439). Such rulings provide fruitful grounds for challenge. This means continued allegations of using biased, flawed, or discredited methods, continued ignorance/misunderstanding about the probabilistic nature of evidence, and continued reliance on dangerously incorrect inferences. Members of our own forensic community will see this as a reason not to improve methods of evidence interpretation, a justification to adhere to the match/non-match paradigm, a reason to not use, or to misuse, databases, and to continue reporting relative frequencies in populations, or *P*-values.

Readers might have thought that at this point all blame had been apportioned. Forensic practitioners are as much a part of the problem as they need to be part of the solution. Again the reasons are varied. There is the fundamental human dislike of change. Forensic scientists are very often constrained by their standard operating procedures, and by their management. Guiding this is the management fear of exposure to litigation over historical case work. Refusing to examine your practices will not diminish this exposure. We also need to acknowledge that there is a fundamental fear of statistics, both in the general public, and in the scientific community. This is completely understandable. Forensic agencies do need to recognize this issue, as forensic science trainees need to recognise that they must use statistics as part of the job. Defensive attitudes of current practice are contrary to the ideas of scientific progress and are extremely damaging.

The "*CSI effect*" has led to a dramatic increase in forensic science educational programmes. This may be perceived as a way to revitalize failing programmes by glamorizing selected scientific disciplines in the way that only the entertainment industry can. One might expect a decline in quality with such a rapid increase. There is a small number of publications relating to forensic science education programmes [7,8,10,9,11]. Of these, several contain information specifically about statistics. Furton et al. [10] surveyed crime lab directors on their statistical education requirements for various forensic positions, Almirall and Furton [9] recorded the number of semesters of mathematics and statistics new graduates had taken as part of their education, and Tregar and Proni reported that 55% of institutions required statistics as part of their curriculum. My suspicion, and I am happy to be corrected, is that the statistics requirement in many of these programmes is no more than the statistics requirement of any science degree. That is, students receive some elementary training in basic data analysis techniques such as hypothesis testing, regression, and maybe basic multivariate analysis. There are two main uses for statistics in forensic science which are in fundamental research and development, and in evidence interpretation. What is lacking is specialist training in the latter, even if it is just a brief overview of the likelihood ratio approach. Both are important, but have different focuses. It is the latter rather than the former that I have concentrated on today. In forensic science education, it would be a very positive step to have dedicated statistically trained faculty staff (preferably with court-going experience) to teach both topics. The judiciary and legal community need to be educated too. I have often been impressed by the scholarship judges apply to their work. We do need to accept however that many judges (and lawyers) are numerically challenged. Education of the judiciary is essential because once we have their acceptance we will no longer have our scientific progress hampered by legal precedence.

For any solution to be effective there has to be the desire to change. Without this nothing will change. The first steps must be about personal change. I plan to spend more time with forensic researchers. I want to spend time in specific communities towards improving methods of interpretation without the expectation that they go to the full Bayesian solution immediately. And finally, I wish to continue making my own work more accessible through software. The forensic statistics community needs to be bigger. Forensic statisticians need to also spend some time in the laboratory. You cannot understand a set of evidential measurements unless you have had some experience of taking those measurements.

I would like to reiterate that my overall aim here is to inspire thought, reflection, and hopefully change. Given that this is an editorial, I will save an extended discussion and bibliography for another piece in the future. Statistical interpretation is a vital part of a modern forensic scientist's toolbox. It is incumbent upon us, as a community, to make sure that we have the best tool set available and that everyone knows how to use it.

References

[1] W. Morris, Can fingerprinting adopt a statistical methodology? Fingerprint Whorld 37 (142) (2011) 85–89.

[2] US DOJ and OIG Oversight and Review Division, A review of the FBI's handling of the Brandon Mayfield case (unclassified and redacted), Technical Report, United States Department of Justice and Office of the Inspector General Oversight and Review Division, 2006, (www.justice.gov/oig/special/s0601/exec.pdf).

[3] A. Campbell, The fingerprint inquiry report, Technical Report, APS Group Scotland, 2011, (www.thefingerprintinquiryscotland.org.uk/inquiry/files/TheFingerprintInquiryReport_Low_res.pdf).

[4] I.W. Evett, The interpretation of refractive index measurements, Forensic Science 9 (1977) 209–217.

[5] Collaborative Testing Services, Glass analysis test no. 11-548 summary report, Technical Report, Collaborative Testing Services, 2011, (Downloaded from www.ctsforensics.com/assets/news/3148_Web.pdf. Last accessed: 07-Jun-2013).

[6] Office of News and Public Information, National Academies. News release, http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=12589, (Last accessed: 10-Jun-2013).

[7] K.M. Higgins, C.M. Selavka, Do forensic science graduate programs fulfill the needs of the forensic science community? Journal of Forensic Sciences 33 (1988) 1015–1021.

[8] J.A. Siegel, The appropriate educational background for entry-level forensic scientists: a survey of practitioners, Journal of Forensic Sciences 33 (1988) 1065–1068.

[9] J.R. Almirall, K.G. Furton, Trends in forensic science education: expansion and increased accountability, Journal of Analytical and Bioanalytical Chemistry 376 (2003) 1156–1158.

[10] K.G. Furton, Y.L. Hsu, M.D. Cole, What educational background is required by crime laboratory directors? Journal of Forensic Sciences 44 (1999) 128–132.

[11] K.L. Tregar, G. Proni, A review of forensic science higher education programmes in the United States: bachelor's and master's degrees, Journal of Forensic Sciences 55 (2010) 1488–1493.

James M. Curran

Dept. of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand

E-mail address: j.curran@auckland.ac.nz.