International Statistical Institute, 57th Biennial Session, Durban, South Africa, 16-22 August 2009

The contribution statistics can make to "Strengthening Forensic Science"

Cooper, Antony K and Das, Sonali
Built Environment Unit, CSIR, PO Box 395, Pretoria, 0001, South Africa
Email: {acooper, sdas}@csir.co.za

Abstract

In February 2009, the National Academy of Sciences (NAS) released the pre-print version of its report, "*Strengthening Forensic Science in the United States: A Path Forward*". While the report focused on forensic science in the United States, it did draw on inputs from other countries and much of the report is relevant to forensic science in other countries. The report makes thirteen detailed recommendations, several of which will require statistics and statisticians for their implementation. In this paper, we highlight some of these recommendations and their statistical needs, specifically, Recommendations 3 (on accuracy, reliability and validity in forensic science), 5 (on human bias and error), 6 (on measurement, validation, reliability and proficiency testing) and 8 (on quality control). Underlying some of these is the need for classification, which we also discuss. We illustrate our discussion with examples from the forensic science of bloodstain pattern analysis.

Key words: Forensic science, Statistics, Classification, Bloodstain pattern analysis.

1. Background

In November 2005, the United States Congress directed the National Academy of Sciences (NAS) to conduct a study on forensic science. The Academy established the Committee on Identifying the Needs of the Forensic Sciences Community, which released the pre-print version of its report, "*Strengthening Forensic Science in the United States: A Path Forward*", in February 2009 [Committee on Identifying the Needs of the Forensic Sciences Community, 2009]. While the report focused on forensic science in the United States, it did draw on inputs from other countries and much of the report is relevant to forensic science in other countries.

The report makes thirteen detailed recommendations, several of which will require statistics and statisticians for their implementation. In this paper, we highlight some of these recommendations and their statistical needs.

Statistics has obviously been used before to assist the forensic sciences, particularly with the advent of DNA profiling, which changed the way the legal system views the use of quantitative data. The Royal Statistical Society defines "forensic statistics" as "the application of statistics to forensic science and the law" [The Royal Statistical Society, 2005]. Unfortunately, statistics can be misinterpreted in court (either deliberately or by mistake), with Thompson & Schumann [1987] identifying the two most common errors as being the *prosecutor's fallacy* and the *defence attorney's fallacy*, both involving the misunderstanding or misrepresentation of conditional probability, or neglecting the prior odds.

2. Bloodstain pattern analysis (BPA)

The first author was privileged to be invited to participate in the Fall 2008 and Spring 2009 meetings of the Scientific Working Group on Bloodstain Pattern Analysis (SWGSTAIN), on the basis of a short piece he wrote on categorising bloodstain patterns [Cooper 2003]. Established in 2000, SWGSTAIN is one of several Scientific Working Groups established by the Federal Bureau of Investigation (FBI) "to improve discipline practices and build consensus with our federal, state, and local forensic community partners" [Adams & Lothridge 2000]. SWGSTAIN has five Subcommittees, for Taxonomy and Terminology; Training and Education; Quality Assurance; Legal; and Research [SWGSTAIN 2009]. The Research Subcommittee has been compiling an annotated bibliography on bloodstain pattern analysis, which is due to be published soon. To date, they have found few references on quality assurance for bloodstain pattern analysis, for example [Michael Illes, *pers comm*, 2009].

With some knowledge of the processes around the use of bloodstain pattern analysis (but not being bloodstain pattern analysts ourselves), we will use this forensic science to illustrate aspects of the National Academy of Sciences report.

3. Accuracy, reliability and validity in forensic science

> ***Recommendation 3:***
> *Research is needed to address issues of accuracy, reliability, and validity in the forensic science disciplines. The National Institute of Forensic Science (NIFS) should competitively fund peer-reviewed research in the following areas:*
> *(a) Studies establishing the scientific bases demonstrating the validity of forensic methods.*
> *(b) The development and establishment of quantifiable measures of the reliability and accuracy of forensic analyses. Studies of the reliability and accuracy of forensic techniques should reflect actual practice on realistic case scenarios, averaged across a representative sample of forensic scientists and laboratories. Studies also should establish the limits of reliability and accuracy that analytic methods can be expected to achieve as the conditions of forensic evidence vary. The research by which measures of reliability and accuracy are determined should be peer reviewed and published in respected scientific journals.*
> *(c) The development of quantifiable measures of uncertainty in the conclusions of forensic analyses.*
> *(d) Automated techniques capable of enhancing forensic technologies* [Committee on Identifying the Needs of the Forensic Sciences Community, 2009].

Referring to Recommendation 3(a), the report highlights a number of situations where they feel that the interpretation of forensic evidence is not always based on scientific studies, such as partial fingerprint validation, and determining sources of bite marks and tool marks [Committee on Identifying the Needs of the Forensic Sciences Community, 2009]. Statistics is essential for addressing this recommendation, from experimental design for setting up the required scientific studies in the first place, through to analysing the results of these studies.

Experimental design can help us understand the relationships between the dependent and the independent variables, as well as the factors that contribute most to the variability of the dependent variable. For example, as discussed at the Spring 2009 meeting of the SWGSTAIN Legal Sub-committee, possible variables to consider for the formation of bloodstain patterns include the nature of the blood used (eg: controlled laboratory experiments might use expired human blood from a blood bank, but such blood most likely has been treated in ways that will affect how it flows and coagulates), the volume of blood, the nature of the target surface and environmental conditions (eg: heat and humidity might encourage moulds to grow on the bloodstains). Experiments can then be designed whereby the outcomes are the defined bloodstain types and the independent variables are these various factors that contribute to the characterisation of the stain.

SWGSTAIN has begun collecting examples of bloodstains (primarily as photographs) to create a reference library, which can be used for validating bloodstain pattern analysis and the training of analysts. However, underlying the analysis of the results of any experiments is the assumption that the data will be 'clean' and 'organised'. In other words, to work with data that come from different measuring and curating organisations and laboratories, there needs to be standards for how each bloodstain pattern is created and recorded, and there needs to be metadata (ie: documentation) recorded for each bloodstain pattern. Surprisingly, while the NAS report does mention the need for adequate documentation, it does not mention metadata at all [Committee on Identifying the Needs of the Forensic Sciences Community, 2009].

4. Human bias and error

> ***Recommendation 5:***
> *The National Institute of Forensic Science (NIFS) should encourage research programs on human observer bias and sources of human error in forensic examinations. Such programs might include studies to determine the effects of contextual bias in forensic practice (e.g., studies to determine whether and to what extent the results of forensic analyses are influenced by knowledge regarding the background of the suspect and the investigator's theory of the case). In addition, research on sources of human error should be closely linked with research conducted to quantify and characterize the amount of error. Based on the results of these studies, and in consultation with its advisory board, NIFS should develop standard operating procedures (that will lay the foundation for model protocols) to minimize, to the greatest extent reasonably possible, potential bias and sources of human error in forensic practice. These standard operating procedures should apply to all forensic analyses that may be used in litigation* [Committee on Identifying the Needs of the Forensic Sciences Community, 2009].

It is important to differentiate those aspects of a forensic science that are objective from those that are subjective. For example, with bloodstain pattern analysis, the identification of bloodstains (including determining that the stains actually are blood, and the species from which the blood originated, through chemical analysis), the identification of bloodstain patterns and the classification of the stains and patterns should be objective [Cooper 2003; Bevel & Gardner 2008]. When given the same crime scene to analyse, different bloodstain pattern analysts should produce the same objective results – there should not be dispute over them in a court hearing.

The subjective aspects of a forensic science relate to the interpretation of the objective aspects, particularly for reconstructing the sequence of events at the crime scene. For example, bloodstain patterns can be interpreted by an analyst to give their opinion on the likely position of a victim when they were bleeding, the nature of their wounds and the arc taken by a bloodied weapon when used to attack a victim.

Experimental design could be used to set up experiments to identify factors that can precipitate human observer bias or human error for both the objective and subjective aspects of the forensic sciences. Statistics can then be used to evaluate the results.

5. Measurement, validation, reliability and proficiency testing

> ***Recommendation 6:***
> *To facilitate the work of the National Institute of Forensic Science (NIFS), Congress should authorize and appropriate funds to NIFS to work with the National Institute of Standards and Technology (NIST), in conjunction with government laboratories, universities, and private laboratories, and in consultation with Scientific Working Groups, to develop tools for advancing measurement, validation, reliability, information sharing, and proficiency testing in forensic science and to establish protocols for forensic examinations, methods, and practices. Standards should reflect best practices and serve as accreditation tools for laboratories and as guides for the education, training, and certification of professionals. Upon completion of its work, NIST and its partners should report findings and recommendations to NIFS for further dissemination and implementation* [Committee on Identifying the Needs of the Forensic Sciences Community, 2009].

Recommendation 6 overlaps to some extent with Recommendation 3 (see above), though the key difference here is the emphasis on developing tools, protocols and standards. In addition, this recommendation also calls for proficiency testing. One way statistical methods could help in proficiency testing is to develop a classification method that helps to corroborate the trainees' abilities. For example, with bloodstain pattern analysis, some bloodstains and patterns, and their causes, will be identified by trainees with a high degree of accuracy across the objective and subjective methods, while some will not. The training set for developing a classification methodology thus should be updated continually with new data, and re-trained.

6. Quality control

> ***Recommendation 8:***
> *Forensic laboratories should establish routine quality assurance and quality control procedures to ensure the accuracy of forensic analyses and the work of forensic practitioners. Quality control procedures should be designed to identify mistakes, fraud, and bias; confirm the continued validity and reliability of standard operating procedures and protocols; ensure that best practices are being followed; and correct procedures and protocols that are found to need improvement* [Committee on Identifying the Needs of the Forensic Sciences Community, 2009].

Quality control processes can be implemented to test if the measurements fall within a certain degree of variability. Also, the possibilities of false-positive and false-negative tests can be implemented to identify possible mistakes.

SWGSTAIN has published guidelines for a laboratory to develop a quality assurance programme for bloodstain pattern analysis, but it is an outline of what is needed and lacks details for a specific implementation [SWGSTAIN 2008]. However, much of the detail is probably readily available from elsewhere, such as a documented health and safety programme.

7. Classification

As discussed above, a forensic science can have objective aspects (the facts discovered, identified and classified at a crime scene or on evidence) and subjective aspects, because of the need to provide expert witness in court (the analyst's opinion of the implications of the facts, such as for reconstructing the events at the time the crime took place). To enhance the scientific merit of the forensic science, the integrity of its objective aspects need to be protected and enhanced.

One way to do this is to have a standard classification or taxonomy that is robust, and hence defendable in court. In considering bloodstain pattern analysis, Cooper [2003] identified several issues to consider in developing a classification, including using qualitative characteristics rather than quantitative characteristics to differentiate categories; not overloading a category (ie: mixing up different types of characteristics to identify a category); and assuming there is only one classification (there might be several different '*views*' of the data). It is also not sufficient to just have a label (or term) to identify a category. One needs to understand that the category reflects an abstract concept (latent variable), and the label merely identifies the category uniquely. In addition, one also needs a formal definition of the category. However, as a definition is not necessarily sufficient to select the category correctly, Bevel & Gardner [2008] consider it essential to have a category selection mechanism as well, such as a decision map or a decision tree.

Secondly, with a classification one can also do backward conditional probability checks, in the sense that given that a certain bloodstain is a 'splash, what is the probability that it has, say, a certain 'irregular spines around the margin'? It might also perhaps be used to develop a guideline that requires a trainee to obtain expertise in classification up to a certain threshold level.

Of course, any classification needs to be tested in the field, including in court. A common mistake with developing a classification (or a standard, for that matter) is assuming that one can develop the perfect classification in a committee, and it is very easy to get into '*analysis paralysis*' and take too long. Invariably, the shortcomings of any classification are quickly revealed in the field. One must expect to revise any classification regularly [Cooper 2003].

8. Conclusions

We have reviewed the National Academy of Sciences, report, "*Strengthening Forensic Science in the United States: A Path Forward*", and discussed the statistical needs for implementing several of the recommendations in the report. We would suggest that there is a need for:

- A standardised data recording system;
- Designing experiments to understand the variability in such data;
- Quality check procedures;
- Developing classification trees for categorising different blood stains;
- Testing various conditional probabilities of classified groups; and
- Developing methods to reduce false-positive and false-negative classifications.

9. Acknowledgements

The first author would like to thank the Federal Bureau of Investigation, the Midwest Forensics Resources Center and the National Institute of Justice for the invitations and the funding to participate in the meetings of the Scientific Working Group on Bloodstain Pattern Analysis

(SWGSTAIN). He would also like to thank the CSIR for the funding support and the SWGSTAIN Members and Guests for being so welcoming to him at the SWGSTAIN meetings.

10. References

Adams DE & Lothridge KL, July 2000, *Scientific Working Groups*. Forensic Science Communications, Vol 2, No 3. Accessed 8 April 2009 at: http://www.fbi.gov/hq/lab/fsc/backissu/july2000/swgroups.htm

Bevel T & Gardner R, 2008, *Bloodstain Pattern Analysis with an Introduction to Crime Scene Reconstruction, Third Edition*. CRC Press, ISBN: 978-1-4200-5268-8, 440pp.

Committee on Identifying the Needs of the Forensic Sciences Community, 2009, *Strengthening Forensic Science in the United States: A Path Forward*, National Research Council of the National Academies. The National Academies Press, Washington DC, ISBN: 0-309-13131-6, 254 pp. Accessed 8 April 2009 at: http://www.nap.edu/catalog/12589.html

Cooper AK, 2 October 2003, *Thoughts on categorising bloodstain patterns*. CSIR ***i***comtek document number: 0442-0001-701-A1, 3 pp. Accessed 8 April 2009 at: http://researchspace.csir.co.za/dspace/handle/10204/2234

Scientific Working Group on Bloodstain Pattern Analysis (SWGSTAIN), January 2008. *Guidelines for a Quality Assurance Program in Bloodstain Pattern Analysis*. Forensic Science Communications, Vol 10, No 1. Accessed 8 April 2009 at: http://www.fbi.gov/hq/lab/fsc/backissu/jan2008/standards/2008_01_standards02.htm

Scientific Working Group on Bloodstain Pattern Analysis (SWGSTAIN), 2009. Home Page. Accessed 8 April 2009 at: http://www.swgstain.org/

The Royal Statistical Society, 2005, *A career as a … forensic statistician*. 6pp. Accessed 8 April 2009 at: http://www.rss.org.uk/main.asp?page=2241

Thompson WC & Schumann EL, September 1987, *Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy*. Law and Human Behavior, Vol 11, No 3, pp 167-187.