The Reliability of Forensic Science – The Jury's Back

Introduction

This is the final article in a series of three examining the reliability of scientific evidence. The first article was a general review highlighting issues arising from the underlying science (or lack of), comparison evidence, low template DNA and quality standards. The article in the last issue of Acquitalk looked in some detail at the main quality standard to which all forensic science laboratories either are, or aspire to be, accredited; ISO 17025 (General requirements for the competence of testing and calibration laboratories). In essence ISO 17025 requires objective evidence of a competent organisation and valid methods. There is little in that standard concerning how results are evaluated and communicated to the Court which is of fundamental importance. This article describes a second quality standard which addresses those shortcomings, and others. That standard is the AFSP (Association of Forensic Science Providers www.afsp.org.uk) Standard, the full text of which may be found in *Science & Justice 2009 49:131-164*.

The AFSP Standard

The AFSP Standard is actually a coherent set of 16 standards covering all factors that impact on the evaluation of forensic evidence which results in an opinion of evidential weight. Forensic evidence complying with the Standard is characterised by balance, logic, robustness and transparency. Compliance makes a significant contribution to the reliability of scientific evidence.

In my opinion, and in the view of many forensic science practitioners and scholars, the concepts that underlie this Standard are revolutionary but like all revolutions it is really the product of a long evolution.

The full title of the AFSP Standard is 'A standard for the formulation of evaluative forensic science expert opinion'. Each term has been carefully chosen to clearly define scope and application. The gestation of this standard took many years probably beginning with *"A model for Case Assessment and Interpretation" R Cook and I W Evett Science & Justice 1998; 38(3): 151–156.* The AFSP cannot claim credit for the scholarship underlying and contributing towards the development of the Standard. It has many fathers but deserved of particular mention is Ian Evett formally of the Forensic Science Service. Here in New Zealand the following should be recognised; Bernard Robertson (editor of the NZ Law Journal) and Tony Vignaux (Emeritus Professor at Victoria University) authors of "*Interpreting Evidence: Evaluating Forensic Evidence in the Courtroom"* (Chichester: John Wiley and Sons, 1995) together with John Buckleton at the Institute of Environmental Science and Research. However, it was the AFSP that drew all the various strands together and codified the set of standards in a form suitable for implementation in a forensic science laboratory.

The AFSP includes organisations that provide the overwhelming majority of scientific evidence in the UK and Irish Republic. Therefore, publishing the Standard was a significant event and resulted in the members committing to fully implementing the Standard within their own organisations and promoting it elsewhere. The European Union is now funding a project aimed at developing the standard for implementation throughout Europe. *European Network of Forensic Science Institutes*

(ENFSI) Monopoly Project 2010 P1 "The development and implementation of an ENFSI standard for reporting evaluative forensic evidence".

Before proceeding it should be borne in mind that the AFSP Standard has been developed and codified by forensic science providers whose direct customer are the police. While such providers fully recognise that their ultimate responsibility is to the Court the Standard, for practical reasons, has been written from a prosecution perspective. However, that in no way biases the Standard in favour of the prosecution. Quite the contrary, compliance places a significant burden on investigators and prosecutors and in most cases non-compliance results in flawed or weakened forensic evidence incapable of supporting the prosecution case.

The Likelihood Ratio

At the heart of the Standard is a simple concept; the likelihood ratio. Its application involves an expert considering the likelihood of his or her findings, observations or results given a pair of competing and exclusive propositions; one that favours the prosecution case and one that favours the defence.

In more precise terms; the likelihood ratio is the likelihood of the evidence given the prosecution proposition divided by the likelihood of the evidence given the defence proposition.

There is no reason in logic why the ratio could not be inverted. However, the convention of the prosecution proposition being in the numerator has been established and accepted.

As a result of the convention a likelihood ratio (LR) of less than 1 means that the evidence is more supportive of the defence proposition and a LR of greater than 1 is more supportive of the prosecution proposition. Forensic science laboratories usually provide a table of verbal equivalents to aid understanding which is usually populated with DNA evidence in mind where LRs of over a million are not uncommon. Other forms of evidence can rarely generate such large LR values. A table of verbal equivalents taken from an Institute of Environmental Science and Research witness statement is reproduced below. There is much debate within the forensic science community about the terminology used in verbal equivalents.

LR	Verbal Equivalent
10 – 100	Supports (the prosecution proposition)
100 – 1,000	Strongly supports
1,000 – 1,000,000	Very strongly supports
Over 1,000,000	Extremely strongly supports

The likelihood ratio approach to the evaluation of forensic evidence has enormous power and, in my opinion, represents the 'final frontier' in forensic science.

This approach makes it crucial for experts to be informed as soon as possible of the defence proposition and to be advised of any new evidence which might influence the evaluation of the two propositions. The Standard requires the expert to make every effort to clearly identify the defence proposition at the earliest opportunity. Where this is not possible the Standard suggests that the expert;

- adopts on behalf of the defence an alternative proposition,
- considers an alternative range of explanations for the findings, or
- considers the likelihood of the findings given only the prosecution proposition – in this case an evaluative opinion cannot be offered.

The key point is that without a defence proposition the evidence cannot be evaluated and an evaluative opinion of evidential weight cannot be offered. When this unbalanced approach is employed it must be made clear in the statement or report.

Examples of the Evaluative Process

There are almost as many examples as there are cases involving forensic evidence but the following should demonstrate the concept.

Evidence – chemical analysis of illicit drugs

What is the likelihood of finding the degree of chemical resemblance observed given that the tablets

a) are from the same batch?
b) are from different batches?

Evidence – intracranial injuries

What is the likelihood of finding subdural hematoma and retinal haemorrhage given that

a) the baby was shaken?
b) the injuries were caused accidentally?

Evidence – inorganic gunshot residue particle

What is the likelihood of finding a particle on the hand of a suspect given that he or she

- discharged the weapon?
- was innocently contaminated?

I am sure the reader can provide a whole host of further examples based on experience.

In each case the expert calculates or estimates the likelihood ratio, the likelihood of the evidence (finding, measurement or observation) given each proposition, and provides an opinion of evidential weight based on the ratio. The jury (triers of fact) considers the weight of the forensic evidence along with all the other evidence presented and considers the likelihood of guilt given all the evidence adduced.

The calculation of the likelihood ratio, where relevant data exists, or its estimation, where it is lacking, has essentially one of three outcomes. The forensic evidence favours either;

- the defence (LR is less than 1),
- neither the defence or prosecution proposition i.e. neutral (LR approximately equal to 1), or
- the prosecution (LR is greater than 1).

Where the outcome is that the evidence is neutral legal questions of relevance and admissibility arise. The retrial of Barry George for the murder of Jill Dando is an example. The Court of Appeal in deciding on a retrial held that gunshot residue evidence (GSR) adduced in the original trial was neutral, see *R v George [2007] EWCA Crim 2722*. At the retrial the judge concluded that the prejudicial value of the neutral GSR evidence outweighed its probative value and ruled the evidence inadmissible. Barry George was acquitted.

Likelihood Ratio and Comparison Evidence

Evidence types relying on comparisons and involving some degree of subjective assessment are currently being closely scrutinised. These include fingerprints, footwear marks, firearms, tool marks and most recently voice and image comparisons.

The propositions considered for fingerprints and marks are along the following lines.

a) the mark was made by the same person/item as the reference or control mark
b) the mark was made by some other unknown person/item

For all these evidence types the issues being addressed are,

- what characteristics should be measured?
- what matching process/procedure should be used?
- what amount of data is needed to ensure a true result with acceptable precision?

While the Court still seems content to rely on evidence based on observation and the experience of the expert it should be born in mind that methods of evidence evaluation which are not based on data are merely speculative and have no logical basis. While the likelihood ratio approach has been criticised in some quarters it remains the only rational means of evaluating forensic evidence. The continuing acceptance by the Court of comparison evidence unsupported by appropriate data is of concern but remains a matter for the Court. However, the Court should be aware of the unease among many forensic scientists regarding the acceptance of speculative evidence.

Courts have tried to find a way around this difficulty by attempting to classify subjective comparison evidence as 'non-scientific' e.g. *Kumho Tyre v Carmichael [526 US 137 (1999)]* and more recently *R v T [2010] EWCA Crim 2439*. However, comparison evidence can be made significantly more objective and should be evaluated using the likelihood ratio approach on the grounds that such an approach is the only rational means of evaluating expert evidence.

Likelihood Ratio and the Scene of Crime

In common with most jurisdictions, the expert in New Zealand owes an overriding duty to the Court and a corollary of this and the application of a logical framework is that the expert must have knowledge of and bear in mind the defence proposition, or possible propositions, from his or her earliest involvement e.g. before forensic recovery at the crime scene. Failure to do so could well bias the investigation. An unbalanced approach to evidence collection could lead to crucial evidence being overlooked.

Forensic recovery should be conducted in a neutral and comprehensive manner. It is a benefit of the Standard not often recognised by law enforcement agencies that a balanced approach to crime

scene examination (i.e. giving consideration to the defence scenario) is as likely to benefit the prosecution as it is the defence and will ensure that experts can do their job effectively and exercise their duties and responsibilities with regard to the Court.

Likelihood Ratio and the Role of the Investigator

The Standard requires that the relevant defence and prosecution propositions are identified at the earliest opportunity. Now this can lead to a practical difficulty in that most forensic science providers have as their immediate customer law enforcement agencies. An early consideration of the (likely) defence case can be problematic. The police may not be content with the forensic scientist 'working on the defence case'. Nevertheless, as the overriding duty of the forensic scientist is to the Court, he or she must do so.

This also touches on the potential conflict between law enforcement science focussed on identifying and prosecuting offenders and forensic science focussed on a fair trial.

The early consideration of defence propositions essential to this logical approach can also be difficult to accomplish for the following reason. Such a central role for the forensic scientist from the earliest stage of the investigation does not fit the emerging 'production line' model for the delivery of forensic science where the broader expertise of the forensic scientist is not brought fully into play.

Likelihood Ratio and the Role of the Jury

The main judicial outcome from the use of likelihood ratio approach to the evaluation of scientific evidence is that such an approach ensures that the expert does not usurp the role of the jury, see *R v Doheny and Adams [1997] 1 Cr App R 369*. The expert is left to focus on matters properly within his or her province and the jury within theirs. These outcomes are considered further in the next section.

Likelihood Ratio and Logic

The logical fallacy of considering the likelihood of the proposition given the evidence (technically known as transposing the conditional) was first identified by Thompson WC and Shumann EL [1987]. *"Interpretation of Statistical Evidence in Criminal Trials" Law and Human Behavior (Springer)* ***II*** *(3): 167* , the impact on the court of usurping the role of the jury was highlighted in a report by Balding DJ and Donnelly P in *"The Prosecutor's Fallacy and DNA Evidence" [1994] CLR 711* and was the subject of the England &Wales Court of Appeal judgment *R v Doheny and Adams*. Although the Doheny and Adams ruling applied to DNA evidence the logic can, and should, be applied to most forms of expert evidence.

The consequence of the fallacy is the consideration of guilt (or non-guilt) by the expert given the evidence. This is the role of the jury and not the expert. Examples of the fallacy are; consideration by the expert of the likelihood that the suspect was the bomb maker given the finding of explosives on the hands or that the fingerprint was left by the suspect given the degree of resemblance observed. This logical trap is frequently fallen into by experts. It seems entirely reasonable for an expert to offer answers to questions of the type "Was it the shoe of the accused that left the mark at the crime scene?" which are of importance to the court. Nevertheless, it is illogical to do so and that type of question is for the trier of fact. Determining the likelihood of guilt versus not-guilt and

whether this exceeds a threshold such as “beyond a reasonable doubt” or “on the balance of probabilities” is the task of the jury (trier of fact).

Likelihood Ratio and Bias

The expert must not consider matters that lie in the province of the jury. To do so the expert would need to consider contextual information from a range of sources beyond the evidence being evaluated. Expert testimony could then be influenced by subjective conscious or unconscious opinion as to the guilt or innocence of the defendant. Human bias was a major concern in the US National Academy of Sciences report [2009] *“Strengthening forensic science in the United States: a path forward” pp. 4-9 to 4-11*. It is a major strength of this logical approach that it is resistant to influence from such sources of bias. In focussing solely on calculating or estimating a likelihood ratio the expert avoids both usurping the role of the jury and bias.

Likelihood Ratio and Value for Money

A particular selling point of the Standard for forensic science providers is that it saves money. The early identification of the defence and prosecution propositions ensures that only those forensic tests which address the propositions are carried out.

Summary

The AFSP Standard is a major development in forensic science. The implementation of this standard is considered one of the highest priorities by the major forensic science providers in the UK and Europe.

At the heart of the Standard is the calculation or estimation of a likelihood ratio and the principles of balance, logic, transparency and robustness. The approach leads to one of three outcomes; the evidence supports the prosecution, the defence or it supports neither.

Employing this logical framework ensures that the expert;

- does not usurp the role of the jury,
- is protected from logical fallacy, and
- is protected from bias.

Therefore, compliance with the Standard;

- enhances the reliability of forensic evidence,
- reduces costs, and
- contributes to a fair trial and a true verdict.

Compliance with the AFSP Standard is a challenge but is essential to avoid the consequences set out above. Compliance places a significant burden on investigators and prosecutors and in most cases non-compliance results in flawed or weakened forensic evidence incapable of supporting the prosecution case.

Concluding Remarks

While it is true to say that there is much work to be done in terms of implementing the AFSP Standard, considerable progress has already been made. The funding of a project to implement the standard throughout Europe and the adoption of the Standard by AFSP members demonstrates both

the commitment of directors and the importance of the Standard. It also bodes well for implementation. *R v T* and my own experience demonstrates that implementing the standard is a major challenge and gaining acceptance by the Court and many experts, particularly those in comparison fields, remains an up-hill struggle. However, in terms of the reliability of forensic evidence, compliance with the AFSP Standard is not optional, it is essential; it ensures that the strength of forensic evidence is logically assessed and accurately reported to the Court.

Finally, in addition to contributing to the reliability of forensic science it also ensures that the role of the jury is not usurped by the expert and, as the jury is a fundamental safeguard of justice, that is no bad thing.

Post script

In this series I hope I have made clear that forensic science has made commendable progress in terms of assuring scientific evidence placed before the Court is reliable. However, there is still work to be done as the standard for overall reliability is challenging. DNA evidence became the 'gold standard' through challenge in the Court. In my opinion, that is a process that much expert evidence should be subjected to.

Original versions of the preceding articles are available at the company web site www.linkedforensics.com. The original version of this latest article will appear there shortly.