A model for case assessment and interpretation

R COOK and IW EVETT*

Forensic Science Service, Metropolitan Laboratory, 109 Lambeth Rd, London SE1 7LP, United Kingdom

G JACKSON and PJ JONES

Forensic Science Service, Chorley Laboratory, Washington Hall, Euxton, Chorley PR7 6HJ, United Kingdom

and

JA LAMBERT

Forensic Science Service, Priory House, Gooch St North, Birmingham B5 6QQ, United Kingdom

Science & Justice 1998; 38: 151–156

Received 24 November 1997; accepted 28 January 1998

The authors describe a new approach to decision-making in an operational forensic science organization based on a model, embodying the principles of Bayesian inference, which has been developed through workshops run within the Forensic Science Service for forensic science practitioners. Issues which arise from the idea of pre-assessment of cases are explored by means of a case example.

Die Autoren beschreiben einen neuen Ansatz für die Entscheidungsfindung in einer operativen kriminaltechnischen Organisationseinheit. Der Ansatz geht von einem Modell aus, das die Grundsätze des Bayes Theorem berücksichtigt. Es wurde in Workshops des Forensic Science Service für kriminaltechnische Praktiker entwickelt. Fragen, die sich zum Arbeitsmodell der vorläufigen Falleinschätzung ergeben, werden an einem Fallbeispiel dargestellt.

Les auteurs décrivent une nouvelle approche quant à la prise de décision dans une organisation opérationnelle de sciences forensiques basée sur un modèle, incorporant les principes de l'inférence bayésienne, qui a été développé au travers d'ateliers pour praticiens qui se sont déroulés dans le Forensic Science Service de Grande-Bretagne. Les questions soulevées par l'idée d'une évaluation préalable des cas sont explorées à l'aide d'un exemple.

Los autores describen un nuevo enfoque a la toma de decisiones en una organización de ciencias forenses operacionales, basada en un modelo que se basa en principios de inferencia bayesiana y que se ha desarrollado a través de mesas redondas dentro del Forensic Science Service para profesionales de las ciencias forenses. Las posibilidades que surgen de la idea de prevaloración de los casos se estudian a través de casos prácticos.

Key Words: Forensic science; Management; Finance; Bayesian inference; Likelihood ratio; Probability; Statistics; Decision-making; Cost-effectiveness; Value for money.

*Corresponding author

Introduction

The essence of forensic science is the drawing of rational and balanced inferences from observations, test results and measurements. This process, which we know as *interpretation*, is experiencing a steady period of evolution, largely because of the growing body of scholarship based on application of the Bayesian paradigm. Recent books by Robertson and Vignaux [1] and Aitken [2] provide cogent overviews of the new discipline. There is a widespread tendency in the forensic science community to view interpretation as a stage which comes somewhere near to the end of a casework examination – at the time of preparing the formal report or statement. However, in this paper we attempt to show the advantages of broadening this view.

The Forensic Science Service (FSS) is committed to the idea of continuous improvement across the spectrum of its activities. In that spirit, the authors have recently been working with scientists of all disciplines on the establishment of a model for enhancing the cost-effectiveness of its casework activities. This work has been done through participative workshops and an initial sketchy idea has evolved, through discussion and casework exercises, into a formal model for interpretation and decision-making which spans the duration of each case; from initial contact with the customer, to post-delivery assessment of the extent to which his/her needs have been met.

This paper presents an overview of the model, together with an illustration of what is perhaps the most important phase – pre-assessment – by means of a simple case example. This will demonstrate that interpretation is not just something to be relegated to the final stage of examination. On the contrary, it should start at the moment of first contact with the customer. Future publications are planned that will enlarge on various issues.

Resources: a new culture

In 1991 the FSS became a government agency. It continues to be a part of the Home Office but with wide responsibilities for its management devolved to its Chief Executive. The FSS is not profit making, is not a business, but *is* run on business-like lines.

One of the early developments was to introduce direct charging for casework, so the FSS is now financed through the sale of its services. The immediate consequence of this is that the costs of forensic science are no longer invisible to operational policemen – who increasingly have come to be referred to as “customers”. This radical change in culture took place as police forces themselves were increasing the devolvement of financial responsibility to officers who are closer to day-to-day operations. Today, around 95% of the business of the FSS comes from British law enforcement agencies, including the London Metropolitan Police (following the merger between the FSS and the Metropolitan Police Forensic Science Laboratory in 1996). Other customers include private law firms and overseas organizations.

Recently, the FSS has introduced the notion of *products*. A product is defined in terms of: an activity; the time taken for it; the cost; the standards to which the activity adheres; the expected outcome; and the chargeable unit.

The aim of direct charging is to enable the customers to make better decisions about how to allocate their own resources, in turn creating a greater sense of value for money. The spirit of the new relationship was explained by Mr Blakey [3], the Chief Constable of West Mercia Constabulary:

> The police and the FSS and other providers have changed over the last decade. We all keep books and look at the figures in ways which would not have occurred to us then. Some people bemoan the need to do so, but it is inevitable and with scarce public money it is right. We can see the writing on the wall and are seeking to change awareness and methods towards finding more cost-effectiveness. The future is exciting and forensic science is firmly a part of it; it has the real capacity to provide the police with an unequivocal value for money service.

It is the belief of the FSS that the perception of such value for money should be that of the *customer*. The initiative that we describe in this paper has the aim of not only providing better value for money, but also of achieving improvement through a genuine partnership in which the customer has a greater participation than hitherto in decisions about what work is done in the laboratory.

Case Assessment and Interpretation Model

The objective of the model is:

> To enable decisions to be made which will deliver a value for money service meeting the needs of our direct customers and the Criminal Justice System.

The form of the model has evolved through three workshops involving experienced forensic scientists in management and advisory roles. At the time of writing, it is not yet operational but already it can be seen to be a strong basis for formulating guidelines for casework that provides a major step towards the objective. The model is not cast in stone and it will, no doubt, evolve in the light of casework experience.

The outline of the model is shown at Figure 1 and it can be seen that it embodies three interlinked phases: *customer requirement, case pre-assessment* and *service delivery*. The pairs of curved open arrows are not meant to connect particular boxes; rather, they are included to convey the impression of continual review and re-appraisal.

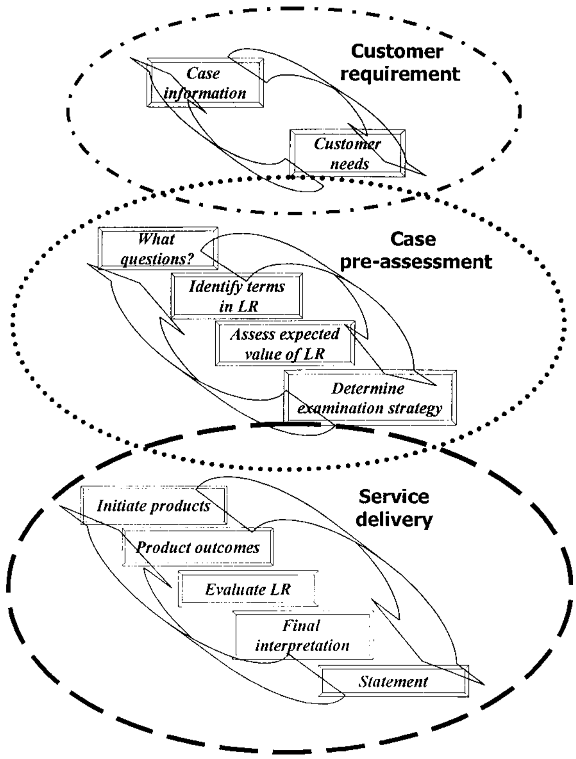

FIGURE 1 Outline of the model. (LR = Likelihood ratio)

Customer requirement

This first phase of a value for money examination clearly must be devoted to determining what the customer *needs*. The roles of the customer and scientist are complementary here and good communication is essential. The scientist should have an adequate appreciation of the case circumstances so that he/she can set up a concise framework for thinking about what kind of examinations may be carried out and what may be expected from them. Right at the outset, the scientist will be taking a balanced view of the case and will wish to learn everything that any suspects say to explain what has happened. It is a principle which is clarified by the Bayesian view of evidence, that it is not sensible for a scientist to attempt to concentrate on the validity of a particular proposition without considering at least one alternative. In the criminal field, it often happens that this situation can be distilled down to a prosecution proposition and a defence proposition. But the clear specification of those propositions is rarely a trivial matter. Suffice it to say, for the present, that in some cases the scientist might be able to address propositions which are quite close to the deliberations of a court such as "this is the person who murdered the victim"; in other cases it might be necessary to settle for propositions further removed from the ultimate issue such as "these fibres came from that garment". A paper which is presently in preparation will explain how there are three broad levels of propositions: level I relates to the *source* of trace material, such as the latter example in the previous sentence; level II relates to *activities*, such as "the suspect is the man who smashed the window"; and level III, the "Jury level" relates to *crimes*, such as the first example in the previous sentence. This can be described as a *hierarchy of propositions*.

Case assessment

The first phase evolves naturally into a consideration by the scientist of what might be expected from an examination. Many forensic scientists would follow this process informally but there is much to be gained by formalising the process and devoting more effort to it than hitherto. This phase requires the scientist to sharpen up the formulation of the propositions to be tested and, in the light of the circumstances, to think hard and, if possible, quantitatively, about what might be expected. This process will be illustrated in more detail by means of an example in the next section.

In this phase, it is desirable that the scientist should document his expectations in statements of the kind "if such and such a proposition were true, then I would expect to find appreciable quantities of transferred fibres". Such expectations not only pave the way for sound decision making during the service delivery phase, but also form the basis of the interpretation which is later made when the statement is written.

Later the scientist will interpret the evidence by calculating the likelihood ratio *(LR)* which is central to the Bayesian formulation of interpretation. It takes the form:

$$\frac{\textit{Probability of the evidence if prosecution proposition is true}}{\textit{Probability of the evidence if defence proposition is true}}$$

In this pre-assessment phase, the expectations can be turned into a probability distribution for the expected weight of evidence, given the circumstances. The example that follows will illustrate this in more detail and it is intended that it should be discussed extensively in a future publication.

Service delivery

This phase represents the main part of the examination. Products are commissioned according to assessments made in the middle phase and following appropriate consultation with the customer: note that the recursive nature of the entire process is emphasised by the use of feedback arrows. It is not suggested that the propositions formulated and the expectations laid down in the middle phase should be regarded as set in concrete: there are many reasons why both propositions and expectations might change as a result of unexpected developments during the examination. There should be a continuous process of review and, where necessary, further consultation with the customer.

By the time that it comes to the writing of the statement, all of the important thought processes should have been followed through and so this stage should not be particularly difficult. The fact that the scientist's expectations were formulated *before* the examination serves to make them more convincing and counters any accusations of *post hoc* rationalisation of findings.

Of these three phases, the middle one represents a marked change in emphasis and is possibly the most important part of the model. This is illustrated by means of an example.

An example

In a hypothetical burglary case, entry to the burgled premises was gained by a person who broke a window at chest height. The alarm was raised and the police were on the scene shortly afterwards. The suspect was apprehended in the next street about ten minutes later. He denied having been anywhere near to the scene. His upper clothing – a polyester "fleece" type garment – was taken for examination about one hour after the arrest (purely for simplicity, we omit any mention of the lower clothing from our discussion). Adequate control samples were taken from the broken window. The investigator requires a search of the garment for glass and comparison between any recovered glass and the control sample.

Considering the case now from the perspective of the scientist who receives the case, the model suggests the following steps:

1. Establish a prosecution proposition C to be evaluated.
2. Establish a defence proposition $\bar{C}$ to be tested against the first.
3. Consider what evidence is to be expected if C is true.
4. Consider what evidence is to be expected if $\bar{C}$ is true.
5. Evaluate what LR is to be expected if C is true.
6. Evaluate what LR is to be expected if $\bar{C}$ is true.
7. Discuss with the customer the expected costs and expected outcomes.
8. Proceed according to agreement with customer.

The determination of the propositions to test gives rise to several issues, which will be discussed in a future paper. For the purpose of the present discussion, let us proceed on the assumption that the scientist decides to address:

C the suspect is the person who committed the burglary

$\bar{C}$ the suspect has had no connection with the crime scene

It is now necessary for the scientist to think about what an examination can be expected to achieve, given that each in turn of these two propositions is really true. He will form these assessments in the light of: previous casework experience; knowledge of glass transfer from previous studies; knowledge of clothing surveys, such as that of McQuillan et al [4]; and the circumstances of the particular case. Those expectations may be expressed in terms of probability distributions for the number of glass fragments which would be found to match the control sample, given C and given $\bar{C}$ respectively. The problem may be simplified further by classifying the number of matching fragments into three classes: none; a few (one to three); several (more than three). This is an arbitrary classification and, in practice, a more flexible and better graduated system would be expected. However, the simple classification system is adequate for illustrating the basic principles.

Now the expert needs to think about three probabilities given C and three given $\bar{C}$, remembering that each set of three is constrained by the requirement that it must sum to one. It is not expected that the expert assign these probabilities with any precision – after all, it is not yet known how rare or common the glass in the broken window is. The scientist's expectations will be subject to review as the case proceeds – assuming the decision is made to continue the examination.

To assess the expected number of matching fragments if C is true, the scientist must take into account:

- The time of arrest and time lapse to taking of clothing
- The retentive properties of the garment
- The consideration that the offender gained entry through the broken window
- Experience from examination of previous cases.

Clearly, the probability is highly case sensitive but, for the sake of illustration, assume that the scientist assigns the values shown in the second column of Table 1. Note that each of the columns sum to one, but there are no constraints on the sums of the rows (the fact that the first row sums to one has no significance).

To assess the expected number of matching fragments if $\bar{C}$ is true, the scientist must turn aside from all observations relating to the crime because now the suspect must be regarded as a person who is unconnected with the crime. The scientist must take account of:

- The retentive properties of the garment.
- Whatever is known about the background of the suspect – particularly any features which might make him susceptible to acquiring glass on his clothing.
- Clothing surveys such as those of McQuillan [4]

Again purely for illustration, assume that the scientist assigns the values shown in the third column of Table 1. It should be emphasised that a glass examiner would almost certainly assign smaller values than the probabilities for "Few" and "Many" if $\bar{C}$ is the case. The values in the table have been chosen because they more clearly illustrate the points that are made later.

Assessing the expected weight of evidence

For any quantity of matching glass Q that is subsequently found, the scientist will calculate the LR:

$$\frac{Pr(Q\,|\,C)}{Pr(Q\,|\,\bar{C})} \qquad (1)$$

So, for this partition of Q into three categories, there are three possible *LR*s to be expected, depending on Q. They can simply be calculated by dividing each probability in column 2 of Table 1 by the corresponding value in column 3, to give the values shown in Table 2.

It is not the usual practice of forensic scientists to report numerical LRs (except, of course, in DNA cases). However, the LR leads to the idea of a verbal scale which incorporates the notion of *support* qualified by terms which are loosely equivalent to a numerical LR scale: the concept is explained in more detail by Robertson and Vignaux [1]. Assume that the scale assigns the qualifier "weak" to LRs in the range 1 to 10; "moderate" to LRs in the range 10 to 100; and "strong" to those in the range 100 to 1000. The scientist can now express his expectations in terms of probabilities for the possible weights of evidence which might ensue from a glass examination, by combining the probabilities in Table 1 with the verbal equivalents of the LRs in Table 2 to give Table 3.

TABLE 1 Probabilities of finding quantities of matching glass.

Quantity Q	*Probability $Pr(Q\|C)$*	*Probability $Pr(Q\|\ )$*
None	0.05	0.95$\bar{C}$
Few	0.30	0.04
Many	0.65	0.01

TABLE 2 Likelihood ratios for the three values of Q.

Quantity Q	*Likelihood Ratio*
None	$^1/_{19}$
Few	7.5
Many	65.0

TABLE 3 Probabilities of possible outcomes from an examination for matching glass.

Opinion expressed	*Probability of opinion if C is true*	*Probability of opinion if $\bar{C}$ is true*
Moderate support for $\bar{C}$	0.05	0.95
Weak support for C	0.30	0.04
Moderate support for C	0.65	0.01

So, on the basis of this assessment, if the suspect is indeed the person who committed the burglary, there is a 65% chance that the result of the examination will provide moderate support for that proposition; and a 30% chance that it will provide weak support. If, on the other hand, the suspect is truly *not* the offender then there is a 95% chance of moderate evidence to support his innocence although there is a 5% chance of evidence which will tend falsely to incriminate him.

It would be helpful if the investigator were given the opportunity of making a decision at this stage, having also been told the expected cost of the examination. Bearing in mind the other non-scientific evidence in the case, would an opinion of "moderate support" from the scientist be sufficient to justify a prosecution? This appears to be a good question to ask though it is one which is outside the domain of the scientist.

TABLE 4 Updated probabilities of finding quantities of matching glass, given the observation that the control glass is unusual.

Quantity Q	*Probability $Pr(Q\|C)$*	*Probability $Pr(Q\|\bar{C})$*
None	0.05	0.995
Few	0.30	0.004
Many	0.65	0.001

TABLE 5 Likelihood ratios for the three values of *Q*, given the observation that the control glass is unusual.

Quantity Q	*Likelihood Ratio*
None	$\sim{}^1/_{20}$
Few	75
Many	650

TABLE 6 Probabilities of possible outcomes from an examination for matching glass, given that the control glass is known to be unusual.

Opinion expressed	*Probability of opinion if C is true*	*Probability of opinion if $\bar{C}$ is true*
Moderate support for $\bar{C}$	0.05	0.995
Moderate support for C	0.30	0.004
Strong support for C	0.65	0.001

This example encapsulates the basic ideas of the pre-assessment phase. It is a way towards the value for money problem discussed by Mr Blakey [3]: let the customer decide whether the examination offers value for money *before* the examination is carried out.

The customer may decide, in this case, that he/she does not wish to proceed. But there is yet another alternative which could be offered – a phased approach to the examination. If the customer agreed to this then things would proceed as follows.

Following agreement with the customer, the scientist commissions a simple product: the measurement of the refractive index of the control sample and comparison of the measurement with a suitable database. Armed with the outcome of this he/she can then revisit Table 1. Imagine, for the sake of illustration, that the control glass is found to be a particularly unusual one. The probabilities in the second column of Table 1 will not be affected by this new information because the quantity of glass transferred if the suspect was indeed the offender should not be influenced by its refractive index. However, we need to reconsider those in the *third* column. Because the control glass is rare, the probabilities of finding glass fragments to match it on the suspect's garment if $\bar{C}$ is true will be much smaller. Imagine that Table 1 is modified by the scientist to Table 4.

Then the LRs corresponding to finding "few", or "many" will be correspondingly larger than originally expected, as shown in the update of Table 2 to create Table 5.

The scientist could then report back to the customer with a modified form of Table 3 shown on the previous page as Table 6.

Clearly, an examination of the clothing promises to be considerably more informative in the light of the new information. This is why the importance of the curved arrows in the model, showing a continual process of reassessment as the case develops, should be emphasised.

This phasing process could be extended another step by next commissioning a search of the garment and reporting back to the customer at that stage, but with the present example such a refinement might not be justified. However, the extension of the principle to more complex cases is obvious.

It should be repeated at this point that the example has been deliberately simplified. One source of complication which has not been mentioned is that of the possible presence of non-matching groups of glass fragments. If the suspect is a habitual offender then glass from other incidents can be expected on his clothing. If C is the case then the presence of non-matching groups will dilute the strength of evidence from any matching glass. It is to be expected that, in the future, information systems would support the scientist through databases assembled from previous cases: these would assist in assigning probabilities to the expected outcomes. Indeed, bearing in mind the complexities of real casework, some kind of expert system would seem desirable. This is far from being a pipedream – a prototype glass expert system which would enable such assessments was reported previously by Buckleton and Walsh [5].

Reporting the results

At the reporting stage, another difference with this approach becomes clear. The thinking with regard to interpretation should start right at the outset of the case and is far from a last minute activity. The statement would be written following guidelines based on principles which have been discussed in more detail elsewhere (see, for example, Evett [6]). The alternative propositions to be addressed and the scientist's expectations have already been stated: the important thinking has already been done. It can rightly be claimed that the expectations were not findings led, but came from a detached appraisal of the circumstances.

However, it is important to recognise that the alternatives addressed and the expectations will not necessarily remain constant through every case. The one used here is a simple one and deliberately constructed so that the observations do not lead to a reappraisal of the alternatives and expectations. There will be cases in which such a reappraisal *is* necessary and this is emphasised by the iterative process implied by the arrows in Figure 1. Such a reappraisal will clearly not be taken lightly and the issues that arise in such cases will be discussed in a future paper.

Conclusion

The process described here forms a framework for sound decision making in terms of impartiality, scientific rigour and the need to be cost-effective. Within this framework, further developments and refinements can take place to tackle new issues and challenges that the forensic scientist will face. The next paper will present further examples to illustrate the concept of the hierarchy of propositions.

References

1. Robertson B and Vignaux GA. Interpreting Evidence: Evaluating Forensic Science in the Courtroom. Chichester: John Wiley and Sons, 1995.
2. Aitken CGG. Statistics and the Evaluation of Evidence for Forensic Scientists. Chichester: John Wiley and Sons, 1995.
3. Blakey DC. Does forensic science give value for money? Science & Justice 1995; 35: 1–4.
4. McQuillan J and Edgar K. A survey of the distribution of glass on clothing. Journal of the Forensic Science Society 1992; 32: 333–348.
5. Buckleton J and Walsh KAJ. Knowledge based systems. In: Aitken CGG and Stoney DA, editors. The Use of Statistics in Forensic Science. Chichester: Ellis Horwood, 1991: 27–50.
6. Evett IW. Interpretation: a personal odyssey. In: Aitken CGG and Stoney DA, editors. The Use of Statistics in Forensic Science. Chichester: Ellis Horwood, 1991: 9–22.