Dawid, A. P. (2008). Statistics and the law. In Evidence, edited by Andrew Bell, John Swenson-Wright and Karin Tybjerg. Cambridge University Press, 119-148.
DOI:10.2277/0521839610

STATISTICS AND THE LAW

Philip Dawid

Department of Statistical Science, University College London

The disciplines of Statistics and Law seem light-years apart, their spheres of interest disjoint and their practitioners inhabitants of different planets. Certainly for much of my own professional life as an academic statistician I gave no thought whatsoever to legal issues, until William Twining, Professor of Jurisprudence at University College London, opened my eyes to the fact — obvious in retrospect — that both disciplines share the same fundamental concern: making sense of evidence. Our ensuing interaction has resulted in my involvement, both as an academic and as an expert witness, in a number of issues arising at the interface between Statistics and Law. It turned out, very much to my own initial surprise, that many of these are intellectually fascinating and challenging, as well as vitally important for the fair administration of justice.

In this chapter I first discuss in some detail a number of statistical and logical issues arising from recent high profile cases involving multiple infant deaths. The issues are subtle, and common sense a lamentably poor guide. I then address similar features of forensic DNA identification, before going on to describe some formal tools, based on graphical representations, that have been found useful for structuring and handling evidence in complex cases. These also offer great promise for application to many other fields of enquiry.

Sally Clark[1]

The case of Sally Clark is the most celebrated of a number of recent cases in the UK in which mothers have been accused of murdering their babies, the evidence against them being wholly circumstantial. Cherie Booth has also discussed the case elsewhere in this volume[2]. It is worth considering at some length, because influential views on both sides of the case have been based on dubious statistical arguments.

One evening in December 1996 Sally was alone in her house with her firstborn son Christopher, aged 2 1/2 months, who until then had seemed a healthy child. Two hours after his

[1] [2003] EWCA Crim 1020

[2] *****Editors: Please insert cross-reference*****

Research Report No.244, Department of Statistical Science, University College London
Date: May 2004

early evening feed his mother found him apparently dead in his bouncy chair in his parents' bedroom, and called an ambulance. Resuscitation was attempted but was unsuccessful. At post-mortem the pathologist found signs of bruising to the legs and damage to the frænulum inside the mouth, as well as some abnormalities in the lungs. His conclusion was that death was due to a lower respiratory tract infection.

In January 1998, Sally's second child Harry died aged 2 months in almost identical circumstances, both parents being this time in the house. Because of Christopher's previous unexpected death Harry had been subject to intensive medical monitoring. Again, he had seemed in good health right up to his death. The post-mortem examination found signs of recent bleeding at the back of the eyes and in the spinal cord.

Alerted by the pathologist, the police initially suspected both parents of actively bringing about their children's deaths, but soon focused on Sally alone. She was accused of having murdered her two children by smothering.

At committal and trial the principal evidence for the prosecution came from medical experts. One of these was Sir Roy Meadow, Professor of Paediatrics and Child Health at Leeds University, who had examined the medical evidence. He testified on a number of medical issues, in which he could reasonably profess some expertise; and also on some statistical issues, in which he could not (in the words of Hermann Bondi[3], "Unhappily, the understanding that statistics is a difficult subject is not widespread, even among distinguished paediatricians"). The main thrust of his statistical evidence was the extreme rarity of two infant deaths occurring from unexplained natural causes (Sudden Infant Death Syndrome, or "SIDS" — otherwise known as "cot death") in a family such as the Clarks. Using figures from an epidemiological study of the incidence of SIDS, he claimed that the overall incidence of SIDS was 1 in 1300 births, falling to about 1 in 8500 if one took into account various characteristics of the Clark family (that they were non-smokers, had waged income, and the mother was over 26). He further testified that the chance of a repeat occurrence, once a first SIDS death has happened, would be essentially the same as for the first. That would imply that the probability of two SIDS deaths, in a family like the Clarks, could be calculated by multiplying 1 in 8500 by itself, leading to an overall rate for double SIDS deaths of about 1 in 73 million such families.

There was no witness at trial with any qualification in Statistics, and no serious cross-examination of the above argument. Summing up, Mr Justice Harrison said "However telling

---

[3] *Nature* **428**, 2004, 799

you may find those statistics to be, we do not convict people in these courts on statistics" but "it may be part of the evidence to which you attach some significance". On 9 November 1999 the jury at Chester Crown Court found Sally Clark guilty, by a 10–2 majority, of smothering her two babies, and she was sentenced to prison for life. The Press immediately seized on the figure of "1 in 73 million" as incontrovertible evidence that Sally Clark was a wicked woman who thoroughly deserved to be locked away.

In January 2000 the British Medical Journal published an Editorial "Conviction by mathematical error?" by Stephen Watkins, an epidemiologist. He claimed that Meadow's presentation of the figure of 1 in 73 million was based on a serious misunderstanding of probability theory. While, as discussed below, this may well have been so, Watkins's analysis was also fundamentally flawed. He argued that we should entirely disregard the first death (Christopher's), because this event had already been taken into account in drawing attention to Sally Clark in the first place — a principle that, if accepted more generally, would make it impossible ever to convict in any case where there was no evidence beyond that which led to arrest. In any case it was the second death (Harry's) that had aroused suspicion of foul play.

The effect of Watkins's recommendation would be to replace the squared figure, 1 in 73 million, by the much larger (and hence less incriminating) single-death figure, 1 in 8500 — or even, according to some studies of the rate of a second SIDS death, 1 in 1700. Meadow's eventual response (in January 2002) was that the statistical argument was never anything but a minor diversion from the medical evidence.

Meanwhile there had been a growing chorus of unease about Sally Clark's conviction, and a flurry of newspaper articles and radio and TV programmes arguing her innocence. Many of these again homed in on the 1 in 73 million statistic, but now aiming, like Watkins, to discredit it, or replace it with a larger and less incriminating figure.

The statistical issues

My own involvement in the case began when I was asked to contribute to the "Dispatches" programme about the case on Channel 4 television on 27 April 2000, and briefed on the background. Later (October 2000) I was engaged as an expert statistical witness for the defence at Sally's appeal hearing. Both in the TV programme and in my expert report I made two main points. The former has been widely echoed, but the latter, which is by far the more significant, much less so.

Point 1: Independence? To calculate the probability of two SIDS deaths in Sally's family, we must not, without further justification, simply square the rate for a single death, as Meadow did. We can do so when we can argue convincingly for the *independence* of those two deaths: i.e. that, after identifying the appropriate SIDS rate to apply to the death of Christopher, the same figure would apply to the death of Harry, even after taking the fact of Christopher's death into account. But on purely commonsense grounds this is implausible. Even after taking into consideration the specific features (no smoker, etc.) of Sally's family used to home in on the figure of 1 in 8500, the two children must have shared many further characteristics, both known and unknown — most obviously, shared genes and domestic environment. The very fact of Christopher's death then gives some reason to believe that there might have been some potentiating factor, which could affect both brothers; and this in turn would increase the chance that Harry, too, would be affected. (As already mentioned, there is in any case some epidemiological evidence that death by SIDS is much commoner after a previous SIDS death in the family.)

Many similar criticisms of Meadow's independence assumption have been aired, to the discredit of the infamous 1 in 73 million statistic. In a widely quoted sound bite in a BBC Radio 5 documentary broadcast in July 2000, Peter Donnelly, Professor of Statistical Science at Oxford, said "Unless the independence has been established, it's wrong. In that sense it's not rigorous, it's just wrong."

Point 2: What should we be looking at anyway? Suppose however we could all agree on a figure for the probability of two deaths by SIDS in a family such as Sally's. Why should this number, of itself, automatically be regarded as interesting and relevant?

An obvious response to this question is that a rare event can reasonably be supposed not to have happened — all the more so, the smaller its initial probability. If we accept this reasoning, the tiny probability for naturally occurring deaths in this case constitutes strong evidence for the alternative hypothesis of Sally's guilt. We can plausibly surmise that the trial jury, like the Press, could have been open to and swayed by such an interpretation of Meadow's statistical evidence, if only subconsciously — particularly in the light of the impressive sounding 1 in 73 million figure.

A more quantitative version of this qualitative reasoning has been dubbed "the prosecutor's fallacy". This involves regarding the figure of 1 in 73 million — actually a measure of the initial rarity of the event "two SIDS deaths" — as the appropriate measure of the probability that that event has happened in this case. This reasoning would rate the probability that Sally is innocent at an entirely negligible 1 in 73 million — so providing overwhelming proof of her guilt. It is clear why such an argument would appeal to prosecutors! There are abundant examples of this reasonable-sounding but in fact totally fallacious argument being accepted and applied, often implicitly and indeed unconsciously, by judges, juries, journalists and the man on the Clapham omnibus — and of consequent miscarriages of justice.

But even in its weaker qualitative form, the argument, while superficially appealling, is just plain wrong. For we could use it to deduce that Christopher and Harry are still alive — since the initial probability of their both dying from any cause at all was surely very small. The evident absurdity of this conclusion demonstrates the illogicality of the reasoning.

So what to do?

At trial it was already known that the very rare event of two infant deaths had happened — that was never in question. Rather, what was at issue was to decide between two possible versions of this event, both initially extremely rare: had Sally's children died of natural, or of unnatural, causes? In particular, as observed by Cherie Booth, from this perspective we should immediately appreciate that the probability that both babies would be murdered, which had never even been considered by the court, must have at least as much relevance to the case as did Meadow's probability that they would both die of SIDS. Applying Meadow's own multiplication approach to relevant official statistics yields a figure for the probability that two babies in one family will both be murdered of about 1 in 2 billion — although this specific calculation is at least as spurious as that of Meadow, and is merely proffered as an illustration of what needs to be considered.

Now whichever of these two competing versions we consider (both babies die of SIDS, or both die of murder), its initial probability is certainly very small indeed — and both these tiny probabilities are equally relevant to deciding the case. While one can readily envisage prosecution and defence brandishing their respective tiny statistics in adversarial combat, it is important to note that, from the statistical viewpoint, there is a unique correct way of proceeding, which involves combining, rather than contrasting, these two numbers. And when this is done, it turns out, surprisingly, that the small absolute values of these probabilities are simply not, after all, of any intrinsic interest or importance. Rather, it is the ratio of these probabilities — the relative odds for comparing the two alternative stories —-that alone carries the essential information needed to choose between them.

Using fictitious figures to illustrate the form of the argument, suppose that, for a family like the Clarks, the probability of two infant deaths from SIDS is taken to be 1 in 5 million, and that for two infant murders 1 in 15 million. The all-important *ratio* of these probabilities is 3: double infant death by SIDS is three times more likely than double infant death by murder. Now in Sally's case we have in fact observed a double death. Supposing we can exclude any other possible causes than SIDS and murder — and ignoring all other evidence in the case — we now know that one of these must be the cause of the observed event, so that their probabilities add to 1: so, to respect the previously calculated odds of 3:1 between the two causes, the probability that the babies both died from SIDS must be three quarters, and the probability that they were both murdered is one quarter.

To elaborate on the above reasoning, consider a hypothetical population containing (say) 150 million families, essentially identical with that of Sally Clark. Out of these we would expect to find about 30 in all (1 in 5 million) in which both babies died of SIDS; and about 10 families (1 in 15 million) in which both babies were murdered — a total of 40 families in which two babies died. We know that, in Sally Clark's family, both babies died: it is one of these 40 families. Since the infant deaths were due to SIDS in 30 of the 40 families, and since, in the absence of any further evidence, we have no reason to consider Sally's family as different in any relevant way from the other 39, the probability that Sally Clark's babies died of SIDS is 30/40, i.e. three quarters; and correspondingly the probability that they were murdered is 10/40, i.e. one quarter.

Applying the above logic to the pair of figures (both admittedly highly suspect) in the actual case — 1 in 73 million for two SIDS deaths and 1 in 2 billion for two murder deaths — their ratio (1/2 billion)/(1/73 million) = 0.0365 would give the odds on Sally's guilt given the evidence of the two deaths. This corresponds to a guilt probability of only 3.5%. Note how different this conclusion is from that resulting from the prosecutor's fallacy, which would put the probability of guilt at 1 minus (1 in 73 million), i.e. as close to 1 as makes no difference. In contrast to such seemingly overwhelming evidence, the conclusion of the correct analysis would certainly not be enough to dispel "reasonable doubt".

Even though truly appropriate figures in such a case may be hard to specify or agree on, the general thrust of the correct argument, combined with "ballpark estimates" of probabilities, is enough to undermine completely the seductive face-value message of the statistical evidence.

There was of course other, medical, evidence in this case, which should not be ignored. The statistical approach to incorporating such additional evidence will be discussed in connexion with the Adams case below.

Aftermath

The other statistical expert witness for the defence, Ian Evett of the Forensic Science Service, raised similar issues in his written report for the appeal, including an exposition of the prosecutor's fallacy. However, when Defence Counsel asked leave to call his two statistical experts to testify orally, Lord Justice Henry replied "We don't need to hear them — it would only be argumentative. After all, it's hardly rocket science". (Actually, it is. Rocket guidance systems such as those of the Apollo space shots are based on statistical control theory.) So we were denied our day in court, and Sally Clark was denied the opportunity to have the serious logical inadequacies of the prosecution's statistical evidence properly exposed. It was clear from the appeal judgment eventually handed down that much of our written evidence had received only the most cursory attention. Their Lordships claimed that the error in the prosecutor's fallacy was so obvious that it did not even need to be drawn to their attention, and that it could not possibly have had any influence on the trial jury's verdict. The conclusion of the court was that "any error in the way in which statistical evidence was treated at trial was of minimal significance". The appeal was dismissed.

Sally Clark was eventually allowed a second appeal, but in this the statistical issues were firmly relegated to the background. It revolved around newly discovered medical evidence of possible bacterial infection in Harry, previously observed but undisclosed by the pathologist, Dr Williams. Although the court did now accept in passing that the original statistical evidence had been misleading, this was done without any new argument, and with frankly questionable logic. At any rate, Sally's conviction was declared unsafe, her second appeal allowed, and her conviction finally quashed on 29 January 2003.

Soon after this two similar cases hit the headlines. In each, a mother was accused of murdering her babies, the principal prosecution evidence, again proffered by Sir Roy Meadow, being the statistical improbability of such deaths occurring by natural causes. Trupti Patel was acquitted at trial on 11 June 2003; Angela Cannings, convicted in April 2002, was freed on appeal on 10 December 2003. In both cases the defence brought evidence of similar unexplained infant deaths among the mother's extended family, aiming to establish a possibility that her own children's deaths could have occurred as a result of some genetically heritable trait.

In all three cases there was specific medical or genetic evidence to provide a plausible alternative explanation for the deaths. It is not clear whether, without this, an attack on the statistical evidence alone would have been enough to outweigh its initial impact. But certainly in their aftermath the pendulum has swung to the other extreme. Statistical evidence of the kind presented by Meadow has been thoroughly rubbished in the courts and the Press — but again with little evidence of any logical understanding of the real issues — and Meadow is being publicly and professionally hounded for allegedly perverting the course of justice. An urgent review is currently under way of all 298 cases from the past ten years in which a parent or carer was found guilty of murdering a child, as well as of thousands of family law cases where a mother suspected of harming her child has had a child taken into care. At the time of writing appeals or referrals to the Criminal Cases Review Commission have been recommended for 5 of the 97 completed reviews of criminal cases.

It now seems unlikely that any prosecutions for causing infant deaths, on the basis of "naked statistical evidence", will be brought to law for the foreseeable future. While this may well be commendable for all sorts of legal and procedural reasons, it would be a great pity if a consequence were to be the banishment of all statistical argument. For example, the impact of medical or genetic evidence could be usefully measured by an estimate of its effect on the all-important ratio between the probabilities of the deaths under natural and unnatural causes.

Identification evidence

One of the major current areas where Statistics impinges on the Law involves the interpretation of identification evidence. In such a case in criminal law, the principal uncertainty is not whether a crime has been committed (the issue for Sally Clark), but rather whether the accused is the culprit. Closely related issues arise in civil law, for example in cases of disputed paternity.

Evidence would now be brought specifically to address the issue of identity. Although this might be, for example, eyewitness evidence, the most incisive identification evidence comes from forensic examination of the crime scene, resulting in the discovery of "trace evidence" — for example, footprints, fingerprints, cartridge cases, DNA — and its matching to similar information obtained from a suspect person or object. With modern advances in DNA technology, DNA profiling has become by far the most common form of forensic identification evidence brought, and I shall focus on it here. Many (though by no means all) of the logical issues arising from DNA identification are essentially the same as for other kinds of identification evidence. An authoritative account of DNA profiling technology and its forensic applications has been given by its inventor, Sir Alec Jeffreys, in his Darwin College Lecture last year.[4]

Initially, a match based on DNA profiling was treated as essentially incontrovertible proof of identity. After all, everyone knows that no two individuals (with the exception of identical siblings or clones) have exactly the same genome, whereas two different samples from the same individual will be genetically identical. However, this ignores serious limitations of both the technology and its naïve interpretation. The forensic scientist could never have the immense scientific, financial and time resources required to construct a full genetic map from a DNA sample, and instead contents herself with making measurements on a limited number of markers (these being identifiable segments of “junk DNA” that display significant variation from person to person). But at this more limited level it is no longer impossible for two distinct individuals to share the same DNA profile. This means that a DNA match is no longer complete proof of identity, so that the principal question now becomes: “What is the strength of the DNA evidence concerning identity?”. More generally, how are we to take proper account of the DNA (or similar) identification evidence in the overall context of the case at hand?

One important special feature of DNA identification evidence is the availability of large databases from which the frequency, in the population at large, of the observed measurement on each marker can be estimated with some precision. On top of this, accepted genetic theory justifies us in multiplying such frequencies, across all the measured markers, to calculate the overall “match probability” — the frequency with which that profile can be expected in the population at large. This step will often result in incredibly small values: figures such as 1 in 1 (US) billion are now routine.

At this point it could be argued that we are back to the initial state of play. Such incredible rarity appears tantamount to impossibility; so that, when we see a match between DNA profiles obtained from two sources, the only remaining viable explanation must be that they originated from the same individual. But once again, such seemingly obvious arguments from tiny probabilities can be grossly misleading.

I shall illustrate some of the issues involved by reference to real cases.

---

[4] *DNA: Changing Science and Society*, Chapter 2,Cambridge University Press, 2004

Denis Adams[5]

In January 1995 Denis John Adams, who lived in the area of the crime, was tried on a charge of sexual assault. The prosecution case rested on expert evidence of a DNA match between Adams and semen, accepted as being that of the culprit, extracted from the victim. No other incriminating evidence was presented. The defence relied on the fact that the victim did not identify Adams at an identification parade and said that he did not look like the man who had raped her. In addition Adams's girlfriend testified that he had been with her at the time of the crime.

The prosecution's forensic expert testified that the match probability attached to the DNA evidence was 1 in 200 million. The defence tried to argue that a figure of 1 in 2 million could not be ruled out. It might be thought that, once we enter the realms of such tiny numbers, arguing about just how tiny they are is the equivalent of counting angels on pinheads — but as we shall see, such argument is not entirely pointless. For the moment, for the sake of argument, we work with the figure of 1 in 2 million.

The logically incorrect but dangerously plausible "prosecutor's fallacy" is particularly tempting in cases involving identification evidence. It would consist here in misinterpreting the match probability — in fact, *the probability of obtaining a DNA match to the crime sample*, had the culprit been some one other than Adams — as *the probability*, on the basis of the DNA match evidence, *that the culprit was not Adams*. (In fact this argument, though common enough in other cases, was carefully avoided by the prosecution in the case of Adams — though we cannot rule out the possibility that the jury nonetheless misunderstood the meaning and impact of the match probability.) More accurately but less memorably, replacing one of these logically quite distinct concepts by the other is also called "transposing the conditional".

The first of the two probabilities above refers to the biological and physical processes generating the two DNA profiles found; the second relates specifically to Adams's culpability. Confusion between these quantities is particularly hard to avoid given that each can be expressed as "the probability that the crime sample came from some one other than Adams" — a form of words that sounds as if it means something definite but in fact is highly ambiguous. Indeed, natural language sometimes seems to have been carefully crafted to facilitate just this confusion. Robert Matthews, one of the few journalists who understands these issues clearly and takes great care in choosing the right words to express them, tells me that his articles are

[5] [1996] 2 Cr.App.R. 467; [1996] Crim. L. Rev. 898; [1998] 1 Cr. App. R. 377

subjected after submission to “minor editing to improve readability” — usually perverting their meaning.

As an analogy to help clarify and escape this common and seductive confusion, consider the difference between “the probability of having spots, if you have measles” — which is close to 1 — and “the probability of having measles, if you have spots” — which, in the light of the many alternative possible explanations for spots, is much smaller.

Application of the prosecutor’s fallacious argument would here have yielded the conclusion that, given the DNA evidence, the probability that Adams is guilty is 1 minus (1 in 2 million) — a number so incredibly close to 1 that there could be no reasonable doubt as to his guilt.

Likelihood ratio

So how should we go about making sense of the DNA evidence?

The task before the jury is to compare two different hypotheses: on the one hand, the *Prosecution Hypothesis*, that the perpetrator of the crime was in fact Adams; and, on the other, the *Defence Hypothesis*, that the perpetrator was someone else. For the sake of argument, we shall assume that under the defence hypothesis we can regard the unknown perpetrator as a random member of some relevant population. (It will often be appropriate to modify or refine this defence hypothesis: for example there may be specific alternative suspects, or one might want to consider the possibility that a — possibly unidentified — relative of the accused was the true culprit. Although such refinements complicate the analysis, they do not affect its overall logic.)

Now statistical theory has given a great deal of careful attention to the general problem of comparing two hypotheses on the basis of evidence obtained. Although there are a number of schools of thought, with different starting-points, arguments and emphases, all are in agreement that the impact of the evidence can be isolated in a quantity called the *Likelihood Ratio*. This is defined as the ratio of the two probabilities assigned to the given evidence, calculated under the two rival hypotheses. Note that each term in this ratio can be regarded as measuring how well the associated hypothesis explained the data actually obtained: there is at least a superficial resemblance to the philosophical doctrine of “inference to the best explanation” expounded by Peter Lipton elsewhere in this volume[6].

In our context the likelihood ratio arising from the DNA evidence can be expressed as:

$$\frac{\text{The probability of obtaining the DNA match, if Adams is guilty}}{\text{The probability of obtaining the DNA match, if Adams is not guilty}}.$$

This is a number that measures the strength of the DNA evidence in favour of the hypothesis of guilt, as against that of innocence. Larger values of the likelihood ratio constitute stronger evidence in favour of guilt. One might consider that the value unity is entirely neutral between the hypotheses, that larger values favour guilt, and that smaller values favour innocence. Although there is a sense in which this is correct, it is a subtle one: as will be discussed below, one must be wary of over-simplistic direct interpretation of numerical value of the likelihood ratio, which can only be sensibly considered in conjunction with other information.

For interpreting and calculating this expression a number of background suppositions and items of evidence are typically required. We may suppose that DNA has been taken from the victim that can be assumed to originate from the culprit, and that its profile has been measured and taken into evidence. We also assume that, under the defence hypothesis, the culprit is unrelated to Adams.

The top line of the likelihood ratio is unity: under the assumptions made, and applying current genetic understandings, if Adams is guilty then there has to be a DNA match. As for the bottom line, under the assumptions made this is just the match probability, which we are taking as 1 in 2 million. Thus the likelihood ratio evaluates to *2 million*.

It seems natural to interpret this very large number as very strong evidence in favour of guilt — an argument distinct from that of the discredited "prosecutor's fallacy", but tending to much the same conclusion. However, we shall see that matters are not so simple.

Other evidence

Quite apart from its logically misleading nature, a serious problem with the prosecutor's fallacy is that it cannot take account of other, non-DNA, evidence in the case. Whether from a logical, a legal or a common-sense point of view, this is clearly unsatisfactory. In the Adams case all the other evidence pointed towards Adams's innocence, and to ignore it would have been highly prejudicial.

Even in cases where no other evidence is explicitly presented — and this is now common in DNA identification cases — this very fact is significant and should be taken into account.

[6] *****Editors: Please insert cross-reference*****

Before introducing any evidence, the accused should be considered as no more likely to be guilty than any other "random" member of the appropriate population. This might be taken as a mathematical translation of the legal "presumption of innocence".

The Adams case was unusual in that both sides took these problems seriously, and attempted to instruct the jury on the reasoning processes needed to address them. In the course of this, numerical values were suggested for probabilities that were not amenable to strict scientific quantification, on the understanding that these values were illustrative, and that each juror should replace them by his or her own assessments before applying the logic. Although this introduces an irreducibly subjective element, there should be reasonable agreement on the order of magnitude of the inputs and the corresponding outputs.

Bayes's theorem

Before any explicit evidence is presented, it might be reasonable to suppose that the culprit is a male aged between about 18 and 60 who is likely to live locally. There were about 150,000 such, and we could expand this to say 200,000 to allow some possibility of a non-local culprit. All that is known about Adams at this point is that he matches these characteristics. Thus the prior probability of his guilt is about 1 in 200,000.

We now face the task of combining this prior assessment with the DNA evidence (we shall consider below the further incorporation of the defence evidence). Fortunately, probability theory tells us exactly how to do this: we have to apply a general result known as *Bayes's Theorem*. In our context this can be expressed as:

$$\textit{posterior odds} = \textit{likelihood ratio} \times \textit{prior odds},$$

where *prior* [*posterior*] refers to the uncertainty *before* [*after*] incorporating the evidence whose effect is measured by the likelihood ratio. In particular, this shows the very specific status and relevance of the value of the likelihood ratio: it is to effect the journey from prior to posterior uncertainty, rather than, as might be thought, to describe the final destination — which must also depend on the starting point, the prior uncertainty.

We have taken the prior probability of Adams's guilt as 1 in 200,000: this is 1 chance for guilt to every 199,999 against, equivalent to prior odds of 1/199,999 on guilt. The DNA likelihood ratio has already been calculated to be 2 million. Substituting these figures into the right-hand side of Bayes's formula, we calculate the posterior odds on guilt as $2{,}000{,}000 \times 1/199{,}999 = 10$ (to 4 decimal places). That is, 10 chances of guilt to every 1 chance of innocence, or 10 chances out of a total of 11: a posterior probability for guilt of $10/11 = 91\%$. While high, it would be hard to argue that this is proof "beyond a reasonable doubt". In any event, there is a striking difference from the corresponding answer, 1 minus (1 in 2 million), resulting from the prosecutor's fallacy — which would be approximately correct when the prior probability of guilt was 50%, rather than the more appropriate 1 in 200,000 used here.

And now for the defence evidence. This has two separate components:

i. The inability of the victim to identify Adams as her assailant.

ii. The alibi provided by Adams's girlfriend.

The defence's statistical expert, Peter Donnelly, explained how a juror might go about evaluating likelihood ratios based on these items of evidence, using indicative figures for clarity but pointing out that these should be replaced by the juror's own assessments.

The probability of obtaining the nonidentification evidence (i) would be low — say around 10% — if Adams were truly guilty; and higher — say around 90% — if he were innocent. Taking the ratio of these two values produces a likelihood ratio of $1/9$ in favour of guilt (which, being smaller than 1, is in fact evidence against guilt).

As for the alibi evidence (ii), we might reasonably expect the girlfriend to produce this — say with probability 25% — even if Adams were guilty; but it would be still more probable — say 50% — if he were truly innocent. This yields a likelihood ratio of $1/2$ (recall that it does not matter whether the specific probability figures we have assessed are correct — only their ratio is of importance).

Under assumptions that are reasonable in this context, we can multiply together the above component likelihood ratios. The overall defence evidence likelihood ratio becomes $1/18$.

Finally, we again use Bayes's formula to update our earlier uncertainty in the light of this further evidence. Multiplying the previously calculated odds of 10 (which, though posterior to the DNA evidence, are prior to the new defence evidence) by the defence's likelihood ratio of 1/18, we find the final odds on guilt, now posterior to all the evidence in the case, to be 5/9: that is, 5 chances for guilt to 9 chances against, or 5 in a total of 14, for a final guilt probability of $5/14 = 35\%$. If there was cause for reasonable doubt before the defence evidence, after it there can be absolutely no case for conviction.

So far we have taken the match probability to be 1 in 2 million. Table 1 shows the effect of varying this while keeping all other ingredients unchanged. We see that argument about the number of noughts in the match probability cannot be dismissed as nitpicking. If it can be shown to be as small as 1 in 200 million, the resulting posterior probability becomes 98%, which might be regarded as beyond reasonable doubt, even after having factored in the low prior probability and the defence evidence. But the larger values of the match probability are much less convincing.

The above Bayesian argument was presented and accepted without objection at trial, but may have well been lost on the jury, who — perhaps subconsciously swayed by a “prosecutor’s fallacy” interpretation of the match probability — convicted. On appeal a retrial was granted on the basis that the judge had not properly instructed the jury on what to do if it did not wish to follow the Bayesian argument. At retrial the defence again presented the argument, this time against prosecution objection, the jury again convicted, and Adams again appealed. This second time the appeal was dismissed — and with it the whole Bayesian argument, on the grounds that it “usurps the function of jury” which “must apply its common sense”. Noble sentiments perhaps; but problematic when common sense can be such a poor guide to handling statistical evidence. Although not a legally binding precedent, this judgment has undoubtedly hampered the presentation of rational statistical argument in the courts.

Hanratty

New issues of calculation and interpretation arise when for some reason it is impossible to obtain a DNA from a suspect. Recourse might then be made to profiling his or her relatives. Because DNA from related individuals will share some features — in a random but well-understood way described by Mendel’s laws of genetic inheritance — such indirect profiling is clearly of some relevance: but how, exactly?

In the infamous “A6 murder” case, James Hanratty was found guilty of murder and rape and hanged on 4 April 1962, going to his death strongly protesting his innocence. From the beginning the verdict was strongly contested, the ensuing disquiet being instrumental in bringing about, in 1965, the abolition of the death penalty in the UK.

Certain items of evidence from the original trial — in particular a handkerchief found wrapped around the murder gun and knickers from the rape victim Valerie Storie — had ever since been retained by the police. In 1998 it was decided to apply modern DNA profiling technology to re-examine these items. A DNA profile, taken to be from the culprit, was found on both items. Its frequency in the population at large was calculated at around 1 in 2.5 million.

Even though Hanratty was dead and buried and so could not supply a DNA profile for comparison, the popular Press greeted the news of this finding in terms such as the following:

"There is a 1 in 2.5 million chance that Hanratty was not the A6 killer"

"The DNA is 2.5 million times more likely to belong to Hanratty than anyone else"

The first quotation here has clearly fallen prey to the prosecutor's fallacy. The second might charitably be interpreted as a description of the likelihood ratio, but is more likely to be have been meant, and certain to be interpreted, as the posterior odds. But, far more crucially, both have completely missed the point that — with no DNA available from Hanratty — the new DNA evidence from the crime exhibits could be no more incriminating against him than against any one else!

In an attempt to prove his innocence, Hanratty's mother and brother now offered their own DNA for profiling. Although a full match in these circumstances was not to be expected even had Hanratty been guilty, if there had been some marker at which the crime profile did not overlap at all with that of Hanratty's mother that would have proved that the crime DNA could not have come from him.

In the event there was no such exclusion. This was widely regarded as tantamount to a full match with Hanratty, so justifying use — and misuse! — of the match probability figure of 1 in 2.5 million. Indeed the forensic expert report did essentially this, referring to a hypothetical suspect whose DNA provided a full match — so erecting a prejudicial smokescreen in front of the inconvenient fact that this was simply not the case for Hanratty.

However, while this indirect evidence based on his relatives' DNA clearly points towards rather than away from Hanratty's guilt, measuring its strength is by no means routine or logically straightforward. At one point the forensic report mentions a likelihood ratio figure of 440 (based on the "STR" component of the DNA). Although no details of this calculation are available, this figure does appear more plausible than the 2.5 million that would be appropriate for a full match. While still providing evidence going towards Hanratty's guilt, it is very much weaker than it was widely, and incorrectly, being taken to be — especially when we remember that a likelihood ratio is not by itself a measure of certainty in the light of the evidence, but has to be combined with suitable prior odds.

Most of these subtle considerations became redundant when in March 2001 Hanratty's body was exhumed, and it was found that his DNA did indeed provide a full match to the crime profile. With this new evidence the likelihood ratio does now become 2.5 million. Although the defence attempted to attribute the match to contamination during the many years for which the crime items had been stored, they failed to persuade the court, and it would seem that the case is finally closed.

Disputed paternity

Problems of DNA testing for disputed paternity can be regarded as involving indirect matching: of the putative father with the (necessarily unavailable) true father. Given DNA profiles from the mother, child and putative father, the likelihood ratio in favour of paternity can be obtained by calculations well known to forensic geneticists (although these are frequently misinterpreted as supplying the posterior odds). As with all cases of indirect matching, the values so obtained are nothing like the stellar figures typically associated with a direct DNA match, although when based on many markers they can still constitute strong evidence.

A still greater degree of indirectness occurs when the putative father's profile is itself unavailable — perhaps he has fled the country. In that case indirect information about his DNA might be obtained from profiles taken from one or more of his relatives. For example, in one real case DNA profiles were obtained from two full brothers and an undisputed child of the missing putative father, as well as from that child's (different) mother.

Without the principles of probability theory for guidance, forensic scientists have been very unclear as to how to interpret such evidence. With that guidance, attention again properly focuses on the likelihood ratio: we have to compare the two probabilities attached to the totality of the observed evidence — whatever its nature — under the competing hypotheses of paternity and non-paternity. But, although this may clarify the logic, calculation of the required probabilities can still be daunting.

Probabilistic expert systems

The modern technology of *probabilistic expert systems* (PES — also known as *Bayesian networks*) has proved invaluable for solving such problems. A PES is a computer software system that allows one to build a graphical representation of a problem, describe the probabilistic relationships between the variables involved, enter evidence on some of them, and "propagate" this to obtain revised probabilities for other variables.

Figure 1 shows a PES network for the indirect disputed paternity case described above. There is a similar network for each DNA marker measured. The white nodes represent unobserved individual genes, the red nodes observed genotypes, and the black node the query

“Is the putative father the true father or not?”. The arrows indicate probabilistic dependencies, specified numerically elsewhere in the system. On entering and propagating the observed evidence at a given marker, the likelihood ratio this generates can easily be extracted from the query node. The overall likelihood ratio, based on all markers, is obtained by multiplying together all such contributions from individual markers.

For the specific case at hand 12 genetic markers were used, the resulting single-marker likelihood ratios in favour of paternity ranging from 0.25 to 6.04. While any single one of these is only weak evidence of paternity (and some, being less than 1, even point in the opposite direction), on combining them we obtain a likelihood ratio value of 1303: that is, the overall DNA evidence (on 12 markers for 6 measured individuals) was 1303 times more probable under the hypothesis of paternity than under that of non-paternity.

The final step of converting this to a posterior odds or probability cannot be taken without the further input of a prior probability, based on other evidence in the case. If this were, say, 5%, the resulting posterior probability of paternity would be nearly 99%.

There are many other variations on the basic theme of DNA profiling, where both the logical and the computational difficulties of interpretation are magnified still further. These include issues such as: multiple perpetrators and/or stains; mixed crime stains (as in rape, or scuffle); database search to identify a suspect; mutation; contamination; laboratory errors; etc. Some of the purely computational problems arising can again be handled using PES. As an example, Figure 2 shows a PES network that can account for the possible disturbing effect of genetic mutation on attributions of paternity — as well as supplying estimates of mutation rates based on nuclear family data when the possibility of non-paternity has to be allowed for. Here white, red and black nodes are much as before, yellow nodes model the mutation process, blue nodes the overall possibly unknown mutation rate, and green nodes various relationships of genetic compatibility or incompatibility among the family profiles.

Mixed evidence

Graphical representations, such as a PES, can also be invaluable at a purely qualitative level, in helping us to organise and comprehend complex webs of relationships between multiple items of evidence from a variety of sources. The following fictitious but realistic example[7] combines eyewitness, fibre and blood evidence.

> An unknown number of offenders entered commercial premises late at night through a hole, which they cut in a metal grille. Inside, they were confronted by a security guard who was able to set off an alarm before one of the intruders punched him in the face, causing his nose to bleed.
>
> The intruders left from the front of the building just as a police patrol car was arriving and they dispersed on foot, their getaway car having made off at the first sound of the alarm. The security guard said that there were four men but the light was too poor for him to describe them and he was confused because of the blow he had received. The police in the patrol car saw the offenders only from a considerable distance away. They searched the surrounding area and, about 10 minutes later, one of them found the suspect trying to "hot wire" a car in an alley about a quarter of a mile from the incident.
>
> At the scene, a tuft of red fibres was found on the jagged end of one of the cut edges of the grille. Blood samples were taken from the guard and the suspect. The suspect denied having anything to do with the offence. He was wearing a jumper and jeans, which were taken for examination.
>
> A spray pattern of blood was found on the front and right sleeve of the suspect's jumper. The blood type was different from that of the suspect, but the same as that from the security guard. The tuft from the scene was found to be red acrylic. The suspect's jumper was red acrylic. The tuft was indistinguishable from the fibres of the jumper by eye, microspectrofluorimetry and thin layer chromatography. The jumper was well worn and had several holes, though none could clearly be said to be a possible origin for the tuft.

Figure 3 captures the salient features of this story and their probabilistic causal relationships. As in the DNA networks above, it has been helpful to introduce some unobserved nodes, here N, A and B, as well as the "query node" C. The network could be elaborated by numerical specification of the probabilistic dependencies indicated by the arrows. But even without this it is possible to read off from the network, by well-defined rules, certain purely qualitative relationships. For example, it can be deduced that once we know N (the number of offenders) and A (who left the fibres on the grille), any further information carried by G_1 (what the guard said about N) and Y_1 (the properties of the fibre tuft) is entirely independent of that carried by R (the pattern of blood on the jumper), and totally uninformative about B (the identity of the person who punched the guard). Such relationships between items of evidence not only are of interest in themselves, but also can be used to simplify the expression and calculation of relevant likelihood ratios.

Even in more straightforward cases, a full and careful analysis will typically require considerable elaboration of the hypotheses and the evidence and of the relationships between

[7] A. P. Dawid and I. W. Evett, *Journal of Forensic Sciences* **42**, 1997, 226–231

them. Thus evidence of Sally Clark's opportunity to commit murder is not of itself evidence of exclusive opportunity (in particular Stephen Clark had also been suspected of causing at least one of the deaths), still less of her criminal intent to murder — a necessary legal component of her guilt, which could itself be separated into separate charges over each death; while the alternative hypothesis of SIDS was never really anything other than a "straw man", and (as transpired at second appeal) by no means exhausts all possible medical explanations. Similarly, DNA evidence of sexual contact can never in itself be evidence of rape, still less, in Hanratty's case, of having murdered some one else; more generally, identification evidence which places an individual at a crime scene does not in itself entail criminal behaviour or guilt. Such hierarchies of propositions and their relationships with the evidence can be handled algebraically in simple cases, but any degree of complexity is more helpfully represented and manipulated using graphical aids.

Wigmore chart

The idea of constructing graphical representations to structure and simplify complex problems has a long history in Law. A seminal contribution was the "chart method" introduced by the American jurist John Henry Wigmore in 1913[8] as an aid to trial lawyers in preparing for and acting in court. Although ignored for many years, its usefulness and generality are now better appreciated, thanks in particular to the efforts of William Twining, Terence Anderson and David Schum.

In this approach the problem is first broken down into known and unknown atomic propositions, specified in a "key list". These are then represented as nodes of a graphical network, and joined by arrows describing various qualitative forms and strengths of evidential relations between them. Wigmore used a wide variety of shapes and other annotations to describe the various forms of evidential relationship: an example from his book is reproduced as Figure 4. Modern versions operate with a reduced collection.

There are both similarities and differences between the PES and Wigmorean charting methods. As for similarities, the Wigmorean approach explicitly aims to represent a specific individual's standpoint rather than "objective truth"; typically a PES can helpfully be considered as serving the same purpose. Both approaches organise many disparate items of evidence and their relationships, focus attention on required inputs, and support coherent

[8] *Illinois Law Review* **8**, 1913, 77

narrative and argumentation.

A PES works with variables and questions, distinguishes between causal and inferential relationships, answers relevance queries, supports efficient computation and simplifies expression of likelihood ratios and posterior odds. A Wigmore chart works with propositions, focuses on inference towards some ultimate probandum, emphasises the distinction between occurrence and report of an event, pays particular attention to the many links in a chain of reasoning, and assists qualitative analysis and synthesis. Current work is exploring these similarities and difference with the aim of developing a system combining the best features of both approaches.

As a comparative case study, David Schum, in as yet unpublished work, has constructed a Wigmore chart for the fictitious robbery case previously modelled by a PES in Figure 3. Working from the standpoint of the prosecutor, he takes as the ultimate probandum:

U: Harold S. unlawfully and intentionally assaulted and injured a security guard Willard R. during a break-in at the Blackbread Brewery premises, 27 Orchardson St., London NW8 in the early morning hours of 1 May 2003.

This is dissected into penultimate probanda:

$\mathbf{P_1}$: In the early morning hours of 1 May 2003, four men unlawfully broke into the premises of the Blackbread Brewery, located at 27 Orchardson St., London NW8.

$\mathbf{P_2}$: Harold S. was one of the four men who broke into the premises of the Blackbread Brewery in the early morning hours of 1 May 2003.

$\mathbf{P_3}$: A security guard at the Blackbread Brewery, Willard R., was assaulted and injured during the break-in at the Blackbread Brewery on 1 May 2003.

$\mathbf{P_4}$: It was Harold S. who intentionally assaulted and injured Willard R. during the break-in at Blackbread Brewery on 1 May 2003.

Schum’s full key-list contains 97 propositions, which number could have been further expanded by explicit incorporation of the generalizations used to warrant evidential links between other items. Table 2 contains the subset of these propositions relating to penultimate probandum $\mathbf{P_2}$: this subset is charted (using leaner modern symbolism) in Figure 5.

A general approach to evidence

I have discussed statistical reasoning, probabilistic expert systems and Wigmore charts in the specific context of Law; but in fact these are completely general formal tools for representing, manipulating and interpreting evidential relationships, with a vast array of potential applications. There is no reason why they could not be fruitfully applied much more widely. As one example of this broader viewpoint, Figure 6 shows part of a Wigmore chart constructed by Terence Anderson to address a historical query raised by Mark Geller: When did the ability to read cuneiform script disappear?

Still more broadly, such extensions suggest that it should be valuable to try and identify general logical principles underlying the interpretation of evidence across all fields of human enquiry, together with general tools for applying them. It is remarkable that, while the need for such an approach to evidential analysis is at least as old as the need for Aristotelian logic, and arguably even more pressing, neither the ancient Greeks nor their modern counterparts have seen fit to pay it the same degree of attention.

The perception of this need has provided the impetus for an interdisciplinary research programme "Evidence, Inference and Enquiry" which has recently been established at University College London with the support of the Leverhulme Trust and the Economic and Social Research Council. This is involving participants with a wide range of disciplinary backgrounds and affiliations, including Statistics, Law, Crime Science, Psychology, Economics, Philosophy of Science, Geography, Medicine, Ancient History, Computer Science and Education. Topics being addressed include: subject- and substance-blind approaches to evidence; inference, explanation and causality; recurrent patterns of evidence; narrative, argumentation, analysis and synthesis; cognitive biases; formal rules; decision aids; and interdisciplinary comparisons. There is much ground to be covered, but the journey has begun.

Acknowledgments

This work was supported by the Leverhulme Trust and the ESRC. I am grateful to Terence Anderson, Ian Evett, Richard Leary, Julia Mortera, David Schum and William Twining for their invaluable past present and future inputs and collaborations.

Further reading

Terence J. Anderson and William L. Twining, *Analysis of Evidence*, London: Weidenfeld and Nicholson, 1991. A modern exposition of Wigmore’s method, with an Appendix on “Probability and Proof” by A. Philip Dawid.

Sally Clark Supporters (http://www.sallyclark.org.uk/). Website maintained by Darwin College Fellow David J. C. MacKay, with links to much relevant information about the case.

A. Philip Dawid, ‘Bayes’s theorem and weighing evidence by juries’, Proceedings of the British Academy **113**, 2002, 71–90. More on Sally Clark and identification evidence.

A. Philip Dawid, Julia Mortera, Vincenzo L. Pascali and Daniel W. van Boxel, ‘Probabilistic expert systems for forensic inference from genetic markers’, Scandinavian Journal of Statistics **29**, 2002, 577–595. How to use PES to solve complex DNA identifications problems.

David A. Schum, *The Evidential Foundations of Probabilistic Reasoning*, New York: Wiley, 1994. Original and authoritative approach to the rational study of evidence — all human life is there.

John Henry Wigmore, *The Science of Judicial Proof*, Third edition, Boston: Little, Brown, 1937. Evidence charting for the trial lawyer, by its originator.

William L. Twining and Iain Hampsher-Monk, *Evidence and Inference in History and Law: Interdisciplinary Dialogues*, Northwestern University Press, 2003. Many interesting and wide-ranging contributions, including a three-chapter dialogue between Mark Geller and Terence Anderson on Wigmorean analysis of “The last wedge”.

Table 1. The Adams case: Dependence of posterior probability of guilt on match probability

		1 in:	
Match probability:	200 million	20 million	2 million
Posterior probability of guilt:	98%	85%	35%

Table 2. Robbery. Key list for penultimate probandum P_2

29) The intruders' car left immediately at the first sound of the alarm leaving the intruders stranded.

30) Willard R. testimony to 29).

31) The intruders dispersed from the Blackbread Brewery premises on foot.

32) Willard R. testimony to 31).

33) The four intruders went their separate ways.

34) In a search of the area surrounding the Blackbread Brewery premises, police apprehended Harold S. trying to "hot wire" a car in an alley about 1/4 mile from the Blackbread Brewery premises.

35) DI Leary testimony to 34).

36) A photo of Harold S. taken shortly after his apprehension to be shown at trial.

37) The photo shown at trial is the same one police took of Harold R. shortly after his arrest.

38) The car Harold S. was trying to "hot wire" did not belong to him.

39) Harold S. was one of the four intruders fleeing the Blackbread Brewery premises.

40) During the police investigation a short time after the intrusion, the police found a tuft of red fibres on a jagged end of one of the cut edges of the metal grille on the Blackbread premises.

41) DI Leary testimony to 40).

42) The tuft of fibres to be shown at trial.

43) The tuft of fibres shown at trial is the same one that police found on a jagged end of one of the cut edges of the metal grille on the Blackbread premises.

44) The tuft of the fibres found on the metal grille on the Blackbread Brewery premises is red acrylic.

45) DI Leary testimony to 44).

46) The tuft of red acrylic fibres found on the metal grille came from an article of clothing.

47) The article of clothing the fibres came from was being worn at the time of the break-in at the Blackbread Brewery.

48) Harold S. was wearing a jumper and jeans at the time of his apprehension.

49) DI Leary testimony to 48).

50) The jumper and jeans to be shown at trial.

51) The jumper and jeans to be shown at trial are the same ones the police took from Harold S. after his apprehension.

52) Harold S's jumper is made of red acrylic.

53) DI Leary testimony to 52).

54) Harold S. was wearing this red acrylic jumper at the time of the break-in at Blackbread Brewery.

55) The tuft of red fibres found on the metal grille on the Blackbread Brewery premises is visually indistinguishable from the fibres on Harold S's jumper.

56) DI Leary testimony to 55).

57) The tuft of fibres and the jumper to be shown together at trial.

58) The tuft of fibres and the jumper are the same ones police obtained during their investigation of the break-in at the Blackbread Brewery.

59) The tuft of red fibres found on the metal grille on the Blackbread Brewery premises is indistinguishable from the fibres on Harold S's jumper as shown by a microspectroflourimetry analysis.

60) DI Leary testimony.

61) Microspectroflourimetry analysis result to be shown at trial.

62) The microspectroflourimetry results shown at trial are the same ones police obtained from the forensic scientist ["boffin"] who performed the analysis.

63) The tuft of red fibres found on the metal grille on the Blackbread Brewery premises is indistinguishable from the fibres on Harold S's jumper as shown by a thin layer chromatography analysis.

64) DI Leary testimony to 63).

65) The results of the thin layer chromatography analysis. to be shown at trial.

66) The thin layer chromatography results shown at trial are the same ones police obtained from the forensic scientist who performed the analysis.

67) The jumper belonging to Harold S. is well worn and has several holes in it.

68) DI Leary testimony to 67.

69) None of holes in Harold S's jumper can be clearly identified as a possible source of the tuft found on the metal grille on the Blackbread Premises.

70) DI Leary testimony to 69).

71) Matching of tufts to holes in fabrics is very difficult.

72) The jumper worn by Harold S. on 1 May, 2003 was torn on a hole in the metal grille at the Blackbread premises.

73) Harold S. was wearing the article of clothing that produced the tuft of red acrylic found on a jagged end of the hole cut into the metal grille at the Blackbread Brewery premises on 1 May, 2003.

74) Testimonial denial by Harold S. of P2, that he was one of four men who broke into the premises of the Blackbread Brewery in the early morning hours of 1 May, 2003.

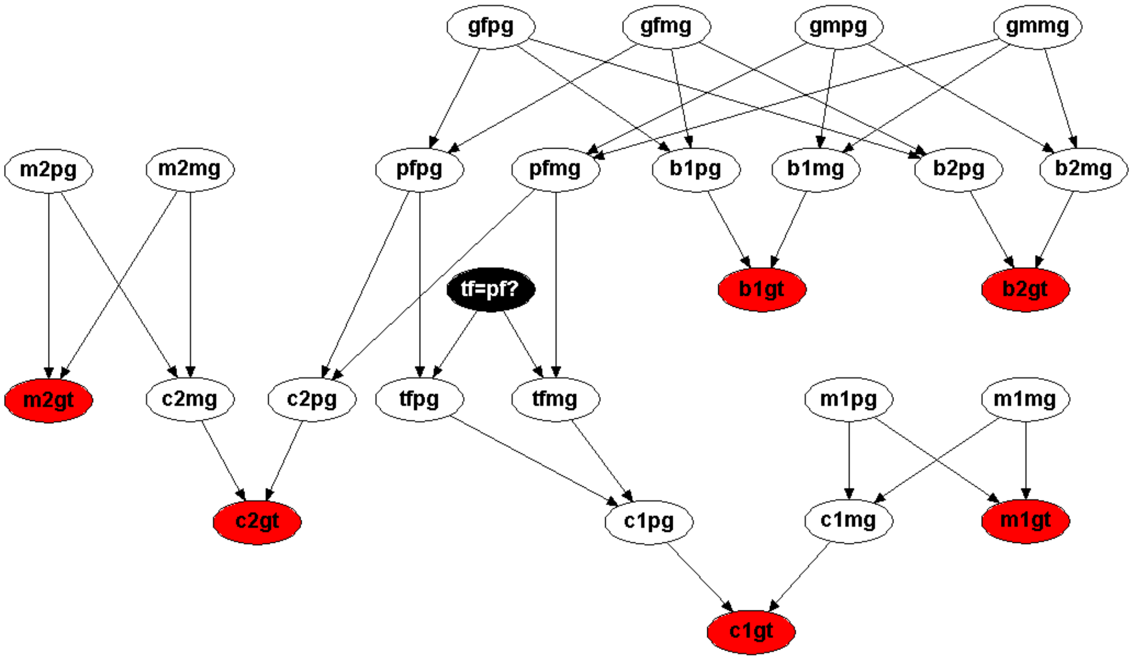


Figure 1. Probabilistic expert system representation of a complex paternity case

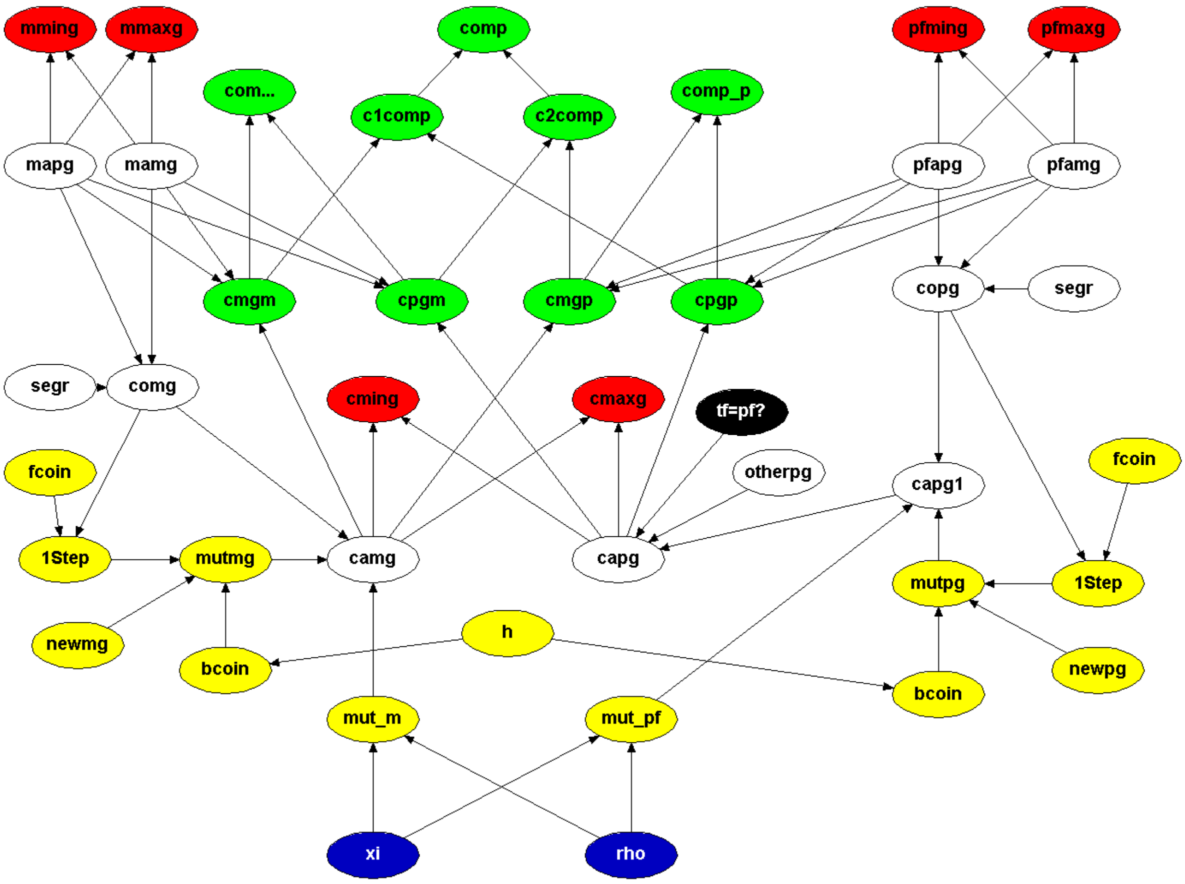


Figure 2. Non-paternity or mutation?

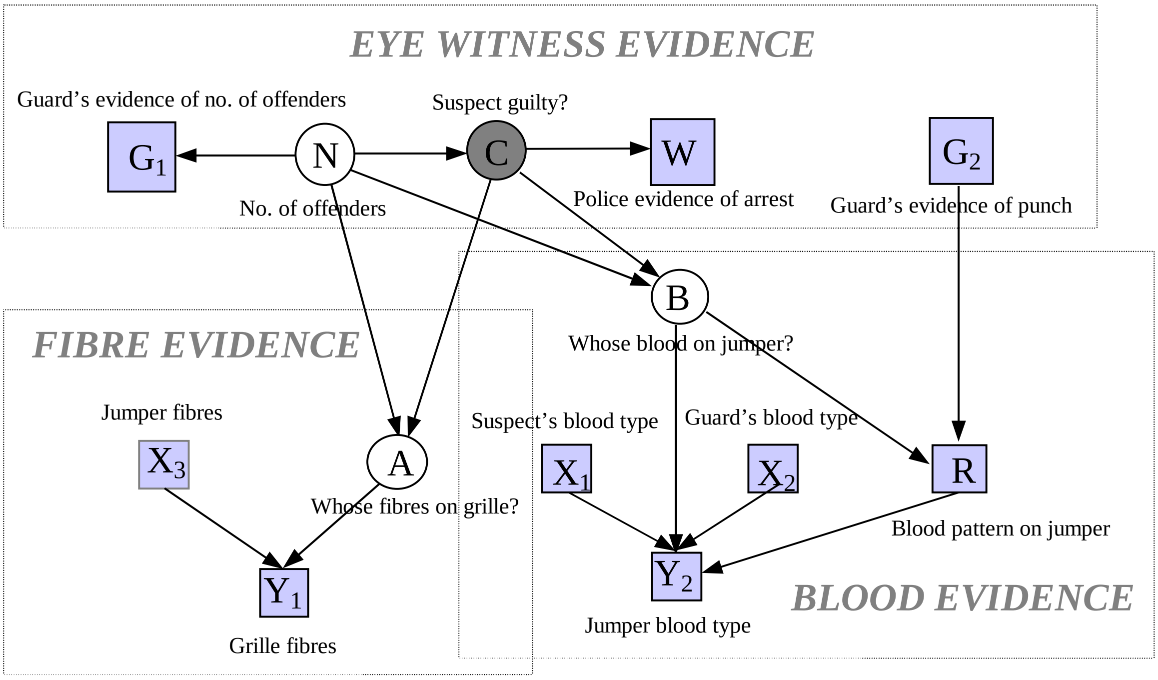
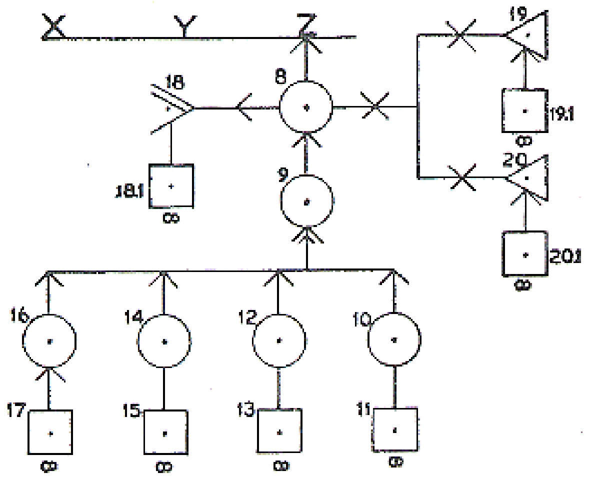


Figure 3. A case of mixed evidence: PES for robbery example

§ 33. Same: an Example Charted. We shall thus have charted the results of our reasoning upon the evidence affecting any single probandum. But this probandum will usually now in its turn (*ante,* § 8) become an evidentiary fact, towards another probandum in a catenate inference. The process of charting and valuation has then to be renewed for this new probandum; and so on until all the evidence has been charted, and the ultimate probanda in issue under the pleadings have been reached.

The following portion of a chart will illustrate (taken from the case of *Com.* v. *Umilian, post,* § 38):



Z is one of the ultimate probanda under the pleadings, viz. that the accused killed the deceased. Circle 8 is one of the evidentiary facts, viz., a revengeful murderous emotion. The arrowhead on the line from 8 to Z signifies provisional force given to the inference.

Figure 4. An original Wigmore chart

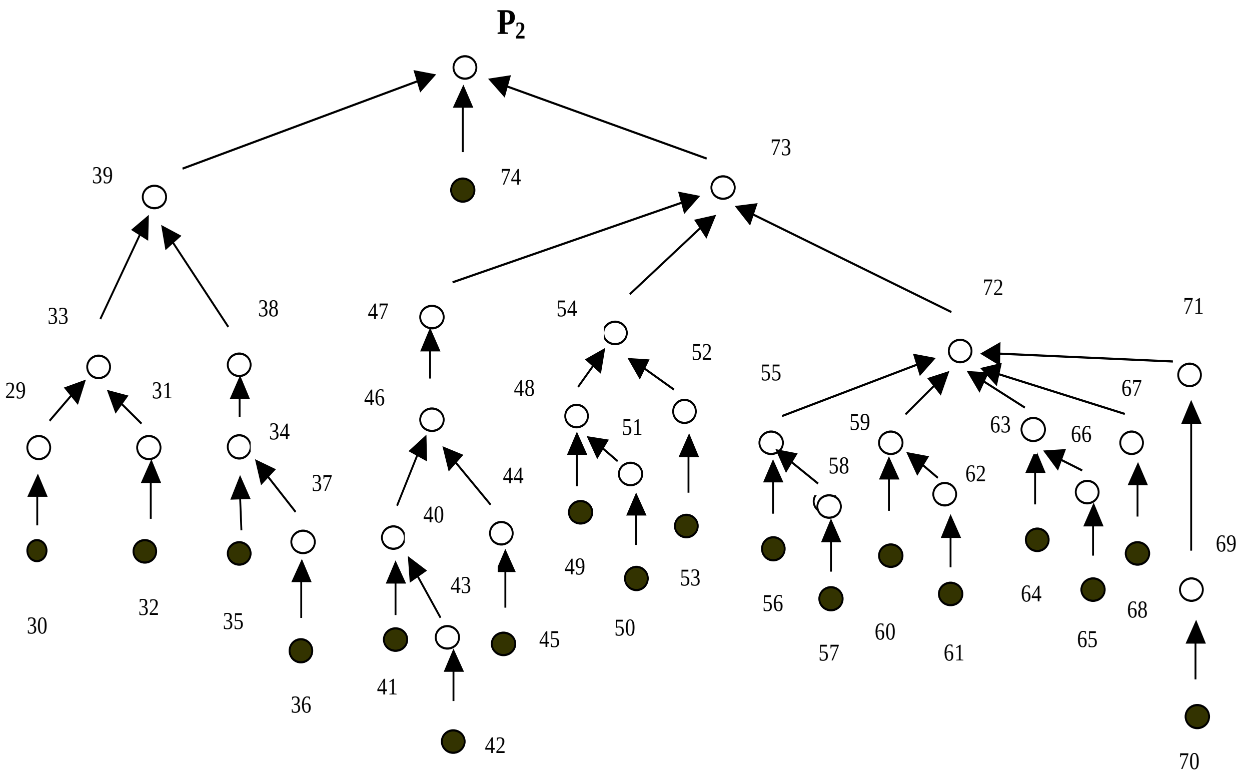


Figure 5. Wigmore chart for robbery

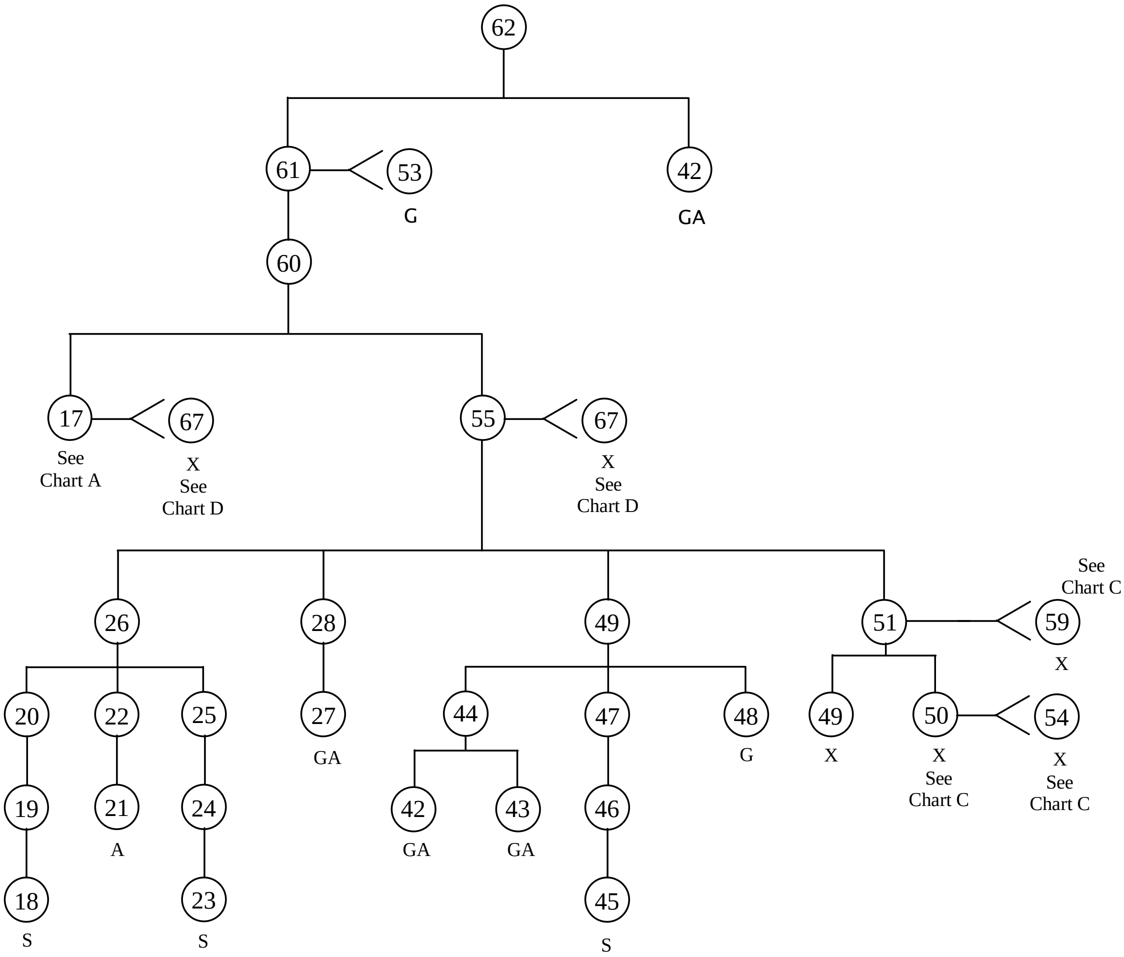


Figure 6. "The last wedge: When did the ability to read cuneiform script disappear?" Wigmore chart B, addressing penultimate probandum (62): ***Iamblichus knew and could read Akkadian cuneiform until the beginning of the third century A.D.***