LIKELIHOOD RATIOS

Transfer Evidence

Relevant Evidence

Rule 401 of the US Federal Rules of Evidence:

"Relevant evidence" means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.

Single Crime Scene Stain

Suppose a blood stain is found at a crime scene, and it must have come from the offender. A suspect is identified and provides a blood sample. The crime scene sample and the suspect have the same (DNA) "type."

The prosecution subsequently puts to the court the proposition (or hypothesis or explanation):

H_p: The suspect left the crime stain.

The symbol H_p is just to assist in the formal analysis. It need not be given in court.

Transfer Evidence Notation

G_S, G_C are the DNA types for suspect and crime sample.
$G_S = G_C$.
I is non-DNA evidence.

Before the DNA typing, probability of H_p is conditioned on I.

After the typing, probability of H_p is conditioned on G_S, G_C, I.

Updating Uncertainty

Method of updating uncertainty, or changing Pr(Hypothesis$_p$) to Pr(Hypothesis$_p$|Evidence) uses Bayes' theorem:

$$\text{Pr}(\text{Hypothesis}_p|\text{Evidence}) = \frac{\text{Pr}(\text{Evidence}|\text{Hypothesis}_p)\,\text{Pr}(\text{Hypothesis}_p)}{\text{Pr}(\text{Evidence})}$$

We can't evaluate Pr(Evidence) without additional information, and we don't know Pr(Hypothesis$_p$).

Can proceed by introducing alternative to Hypothesis$_p$.

First Principle of Evidence Interpretation

To evaluate the uncertainty of a proposition, it is necessary to consider at least one alternative proposition.

The simplest alternative explanation for a single stain is:

H_d: Some other person left the crime stain.

Evett IW, Weir BS. 1998. "Interpreting DNA Evidence."

Can be downloaded from:

`www.biostat.washington.edu/~bsweir/InterpretingDNAEvidence`

Updating Odds

From the odds form of Bayes' theorem:

$$\frac{\text{Pr}(\text{Hypothesis}_p|\text{Evidence})}{\text{Pr}(\text{Hypothesis}_d|\text{Evidence})} = \frac{\text{Pr}(\text{Evidence}|\text{Hypothesis}_p)}{\text{Pr}(\text{Evidence}|\text{Hypothesis}_d)} \times \frac{\text{Pr}(\text{Hypothesis}_p)}{\text{Pr}(\text{Hypothesis}_d)}$$

i.e. Posterior odds $=$ LR $\times$ Prior odds

where

$$\text{LR} = \frac{\text{Pr}(\text{Evidence}|\text{Hypothesis}_p)}{\text{Pr}(\text{Evidence}|\text{Hypothesis}_d)}$$

Questions for a Court to Consider

The trier of fact needs to address questions of the kind

- What is the probability that the prosecution proposition is true given the evidence,
  $\Pr(H_p|G_C, G_S, I)$?

- What is the probability that the defense proposition is true given the evidence,
  $\Pr(H_d|G_C, G_S, I)$?

Questions for Forensic Scientist to Consider

The forensic scientist must address different questions:

- What is the probability of the DNA evidence if the prosecution proposition is true,
  $\Pr(G_C, G_S | H_p, I)$?

- What is the probability of the DNA evidence if the defense proposition is true,
  $\Pr(G_C, G_S | H_d, I)$?

Important to articulate H_p, H_d. Also important not to confuse the difference between these two sets of questions.

Second Principle of Evidence Interpretation

Evidence interpretation is based on questions of the kind 'What is the probability of the evidence given the proposition.'

This question is answered for alternative explanations, and the ratio of the probabilities presented. It is not necessary to use the words "likelihood ratio". Use phrases such as:
'The probability that the crime scene DNA type is the same as the suspect's DNA type is one million times higher if the suspect left the crime sample than if someone else left the sample.'

Third Principle of Evidence Interpretation

Evidence interpretation is conditioned not only on the alternative propositions, but also on the framework of circumstances within which they are to be evaluated.

The circumstances may simply be the population to which the offender belongs so that probabilities can be calculated. Forensic scientists must be clear in court about the nature of the non-DNA evidence I, as it appeared to them when they made their assessment. If the court has a different view then the scientist must review the interpretation of the evidence.

Example

"In the analysis of the results I carried out I considered two alternatives: either that the blood samples originated from Pengelly or that the ... blood was from another individual. I find that the results I obtained were at least 12,450 times more likely to have occurred if the blood had originated from Pengelly than if it had originated from someone else."

Example

Question: "Can you express that in another way?"

Answer: "It could also be said that 1 in 12,450 people would have the same profile . . . and that Pengelly was included in that number . . . very strongly suggests the premise that the two blood stains examined came from Pengelly."

[Testimony of M. Lawton in R. v Pengelly 1 NZLR 545 (CA), quoted by
Robertson B, Vignaux GA, Berger CEH. 2016.*Interpreting Evidence (Second Edition)*. Wiley.

Likelihood Ratio

$$\text{LR} = \frac{\Pr(G_C, G_S|H_p, I)}{\Pr(G_C, G_S|H_d, I)}$$

Apply laws of probability to change this into

$$\text{LR} = \frac{\Pr(G_C|G_S, H_p, I)\,\Pr(G_S|H_p, I)}{\Pr(G_C|G_S, H_d, I)\,\Pr(G_S|H_d, I)}$$

Likelihood Ratio

Whether or not the suspect left the crime sample (i.e. whether or not H_p or H_d is true) provides no information about his genotype:

$$\Pr(G_S|H_p, I) = \Pr(G_S|H_d, I) \;=\; \Pr(G_S|I)$$

so that

$$\mathrm{LR} \;=\; \frac{\Pr(G_C|G_S, H_p, I)}{\Pr(G_C|G_S, H_d, I)}$$

This is the form that allows the consideration of relatives and/or population structure, as well as drop-out and drop-in.

Likelihood Ratio

$$\text{LR} = \frac{\Pr(G_C|G_S, H_p, I)}{\Pr(G_C|G_S, H_d, I)}$$

When $G_C = G_S$, and when they are for the same person (H_p is true):

$$\Pr(G_C|G_S, H_p, I) = 1$$

so the likelihood ratio becomes

$$\text{LR} = \frac{1}{\Pr(G_C|G_S, H_d, I)}$$

This is the reciprocal of the probability of the *match probability*, the probability of profile G_C, conditioned on having seen profile G_S in a different person (i.e. H_d) and on I.

Likelihood Ratio

$$\text{LR} = \frac{1}{\Pr(G_C|G_S, H_d, I)}$$

The next step depends on the circumstances I. If these say that knowledge of the suspect's type does not affect our uncertainty about the offender's type when they are different people (i.e. when H_d is true):

$$\Pr(G_C|G_S, H_d, I) = \Pr(G_C|H_d, I)$$

and then likelihood ratio becomes

$$\text{LR} = \frac{1}{\Pr(G_C|H_d, I)}$$

The LR is now the reciprocal of the *profile probability* of profile G_C.

Profile and Match Probabilities

Dropping mention of the other information I, the quantity $\Pr(G_C)$ is the probability that a person randomly chosen from a population will have profile type G_C. This profile probability usually very small and, although it is interesting, it is not the most relevant quantity.

Of relevance is the match probability, the probability of seeing the profile in a randomly chosen person after we have already seen that profile in a typed person (the suspect). The match probability is bigger than the profile probability. Having seen a profile once there is an increased chance we will see it again. This is the genetic essence of DNA evidence.

Likelihood Ratio

The estimated probability in the denominator of LR is determined on the basis of judgment, informed by I. Therefore the nature of I (as it appeared to the forensic scientist at the time of analysis) must be explained in court along with the value of LR. If the court has a different view of I, then the scientist will need to review the interpretation of the DNA evidence.

Random Samples

The circumstances I may define a population or racial group. The probability is estimated on the basis of a sample from that population.

When we talk about DNA types, by "selecting a person at random" we mean choosing a person in such a way as to be as uncertain as possible about their DNA type.

Convenience Samples

The problem with a formal approach is that of defining the population: if we mean the population of a town, do we mean *every* person in the town at the time the crime was committed? Do we mean some particular area of the town? One sex? Some age range?

It seems satisfactory instead to use a convenience sample, i.e. a set of people from whom it is easy to collect biological material in order to determine their DNA profiles. These people are not a random sample of people, but they have not been selected on the basis of their DNA profiles.

Meaning of Likelihood Ratios

There is a personal element to interpreting DNA evidence, and there is no “right” value for the LR. (There is a right answer to the question of whether the suspect left the crime stain, but that is not for the forensic scientist to decide.)

The denominator for LR is conditioned on the stain coming from an unknown person, and “unknown” may be hard to define. A relative? Someone in that town? Someone in the same ethnic group? (What is an ethnic group?)

Meaning of Frequencies

What is meant by "the frequency of the matching profile is 1 in 57 billion"?

It is an estimated probability, obtained by multiplying together the allele frequencies, and refers to an infinite random mating population. It has nothing to do with the size of the world's population.

The question is really whether we would see the profile in two people, given that we have already seen it in one person. This conditional probability may be very low, but has nothing to do with the size of the population.