

Are expert opinions telling us both sides of the story?

A. Philip Dawid Gianluca Baio
 Department of Statistical Science
 University College London
 Gower Street, London WC1E 6BT, UK
 {dawid, gianluca}@stats.ucl.ac.uk

1. Introduction

- Expert opinions are increasingly often used as evidence for statistical inference in many fields of research, especially in forensic science, where the jurors' decision might be guided by information on the compatibility of scientific evidence with a particular hypothesis of interest for the case at hand
- Typically, experts report on complex pieces of information, possibly related to a comparison between two (or more) different samples, and provide their own assessment of how this evidence might support a causal relation between the crime and the suspect(s)
- In terms of decision making, this problem can be optimally solved by means of the Bayesian approach, where an expert updates all his/her prior knowledge of the uncertain states of the world by means of the observations of this particular case and chooses the explanation that maximises their expected utility
- However, it often happens that expert opinions only provide partial information on their interpretation of the forensic material. The actual process by which an expert declares a "match" between the crime and the suspect sample (i.e. favours the hypothesis of guilt of the suspect) might not take into account properly how more likely this hypothesis is in comparison to the alternative one(s), explicitly violating logic and Bayesian decision making axioms.

"... Psychologists in the mid-1990s established a set of traits of the Munchausen mother that are broad enough to cast suspicion on many whose children are genuinely ill. The signs include attentive mothering, a reluctance to leave the sick child's side, familiarity with medical terms and, most devastating, the denial of accusations of abuse." - The Times, May 25, 2006

2. The Torgensen case

- On 7 December 1956, the fire brigade were attending to a minor fire in a basement in the harbour area of Oslo (Norway) when they discovered the body of a 16 year old girl. She had been strangled, raped and mutilated
- Later a man was apprehended, brought to trial and convicted, partly on the basis of the evidence provided by a skin injury of 6 individual marks on the victim's breast
- It was agreed that the marks were caused by biting with a human set of teeth. The dental expert convinced the jury that the bite mark exhibited characteristics that were consistent with characteristics exhibited by the suspect's set of teeth
- We are interested here in a simplified version of the whole model derived from the evidence available for the real case (Figure 1).
- In particular, we consider the assessment made by an expert as to whether the bite marks found on the victim's body might be compatible with some characteristics of the suspect teeth

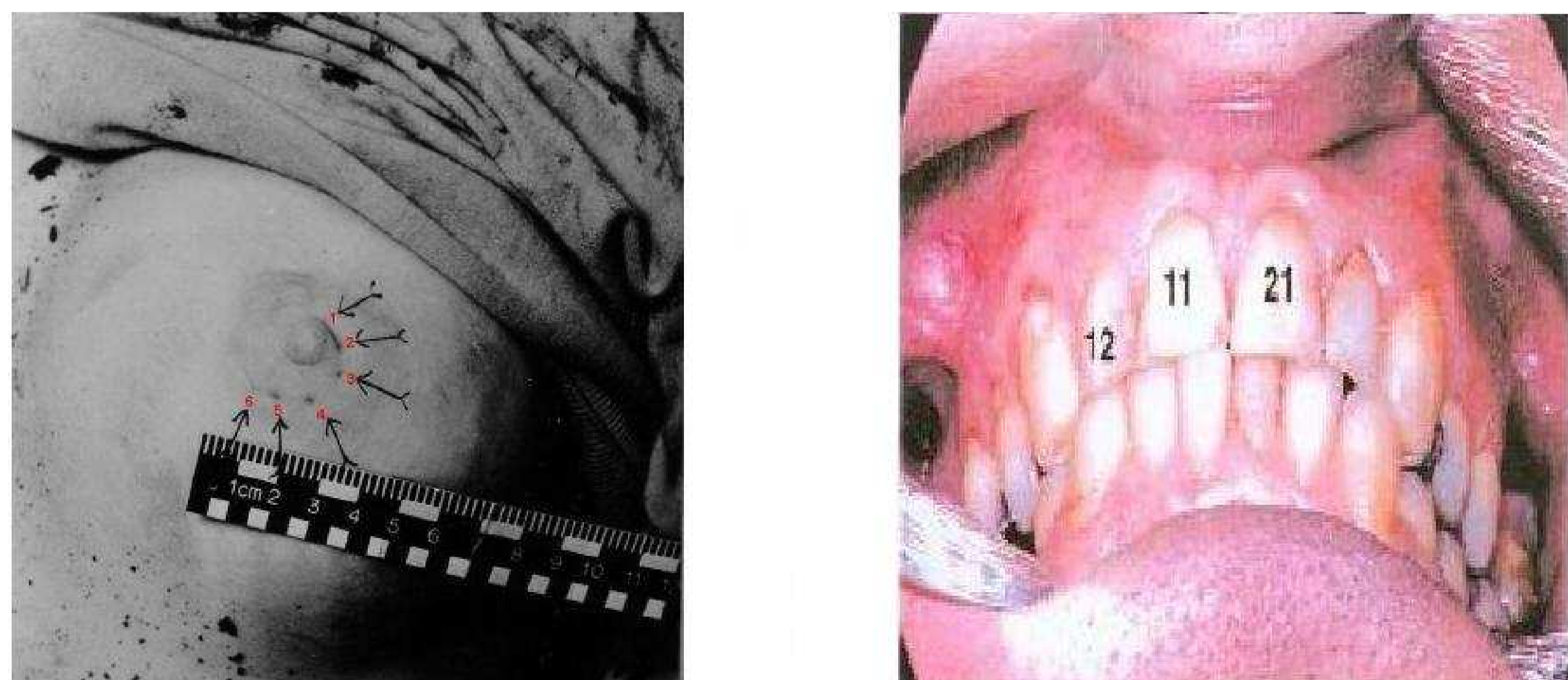


Figure 1: The data observed on April 1957: the skin marks on the victim's breast; and the suspect's teeth

3. Graphical representation of the model

- Standpoint: who are we?
 - We take the position of a rational assembler of the evidence
- Our main objective is to describe the process that leads to the expert's testimony

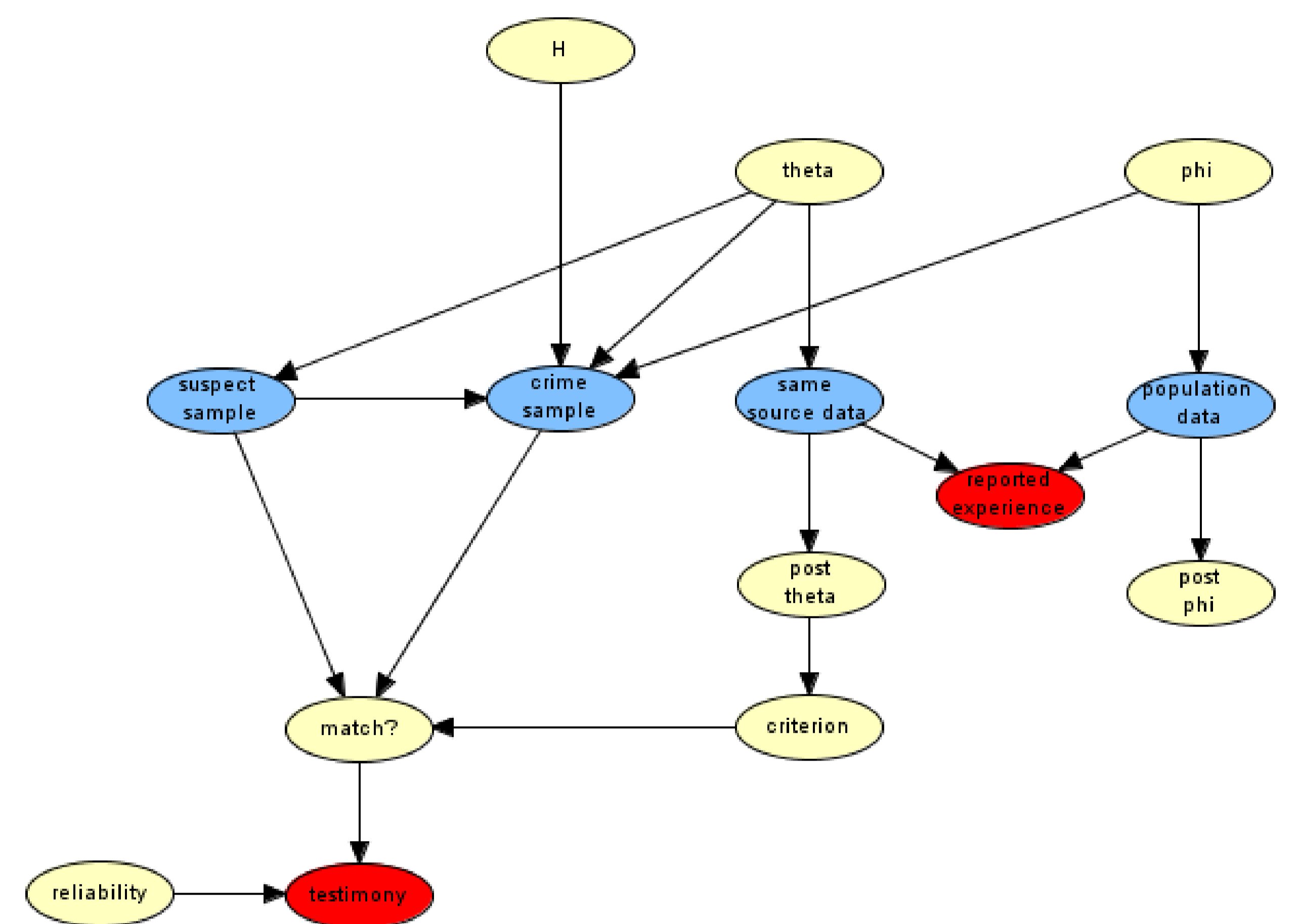


Figure 2: A possible representation of the model for the expert's opinion process, from the point of view of the "rational assembler of the evidence". Blue nodes represent evidence for the expert, whereas red nodes represent evidence for the jurors. Yellow nodes are unobservable variables

4. Description of the model

- H : is the working hypothesis, i.e. the teeth of the suspect are the cause of the skin mark #2 as found on the body of the victim;
- *crime sample*: the whole set of characteristics of the skin marks observed on the victim's breast;
- *suspect sample*: the whole set of characteristics of the suspect's teeth;
- ϕ : $\Pr(\text{crime sample} | \text{suspect sample}, H = \text{false})$ - Represents the true probability distribution of the crime sample, UNDER the assumption that the suspect did not cause it. On a different level, this parameter represents also the probability distribution of the crime sample, if picked at random from a reference population
- θ : $\Pr(\text{crime sample} | \text{suspect sample}, H = \text{true})$ - Represents the true probability distribution of the crime sample, conditionally on the suspect sample, UNDER the assumption that the suspect caused it;
- *same source data*: represents the data that the expert has analysed in the past, regarding cases similar to this one, where crime and suspect sample actually came from the same source;
- *population data*: represents the witness experience, in terms of knowledge about teeth characteristics in the population at large;
- *reported experience*: is what the expert declares in court as "evidence" of their capability. It could be described in terms of past participation in similar forensic cases, the number of publications in relevant journals, etc.;
- *post phi*: is the expert's posterior assessment of the probability distribution ϕ , given the witness' experience and the evidence. We assume that it is derived by the application of rational decision making (i.e. Bayesian Decision Theory);
- *post theta*: is the expert's posterior assessment of the probability distribution θ , given the witness' experience and the evidence. We assume that it is derived by the application of rational decision making (i.e. Bayesian Decision Theory);
- *criterion*: models the criterion that the witness uses to decide upon the match. A rational assessment of the evidence should be such that the criterion include the assessment of both *post theta* and *post phi*, not just one of the two!
- *match?*: logical node that describes whether the expert declares the match or not;
- *reliability*: assessment of reliability of the witness;
- *testimony*: the actual statement of the witness about the present case.

5. Further developments

- Compare the results of "face value" analysis with those obtained by a "rational" expert
- Provide a framework to help experts make good assessments from poor evidence
- Eventually improve investigation process highlighting possible deficiencies in expert judgements

Acknowledgments

We would like to thank Elisabet Lund for bringing this case to our attention and for helpful discussion and insights.

More info on the case: Document Collection to the Norwegian Appeals Committee of the Supreme Court, Volume I-V, 2001, C/O Professor Staale Eskeland, Department of Public and International Law, University of Oslo, St. Olavplass, PB 6706, 0130 Oslo, Norway.