To Test or Not to Test? A Question of Rational Decision Making in Forensic Biology

Simone Gittelson[1,2*] and Franco Taroni[1]

[1]School of Criminal Justice, University of Lausanne, le Batochime, Quartier UNIL-Sorge, Lausanne, 1015, VD, Switzerland.
[2]Department of Forensic Sciences, George Washington University, 2100 Foxhall Road NW, Washington, 20007, DC, United States.

*Corresponding author(s). E-mail(s): sgittelson@gwu.edu;
Contributing authors: franco.taroni@unil.ch;

Abstract

How can the forensic scientist rationally justify performing a sequence of tests and analyses in a particular case? When is it worth performing a test or analysis on an item? Currently, there is a large void in logical frameworks for making rational decisions in forensic science. The aim of this paper is to fill this void by presenting a step-by-step guide on how to apply Bayesian decision theory to routine decision problems encountered by forensic scientists on performing or not performing a particular laboratory test or analysis. A decision-theoretic framework, composed of actions, states of nature, and utilities, models this problem, and an influence diagram translates its notions into a probabilistic graphical network. Within this framework, the expected value of information (EVOI) for the submission of an item to a particular test or analysis addresses the above questions. The development of a classical case example on whether to perform presumptive tests for blood before submitting the item for a DNA analysis illustrates the use of this model for source level questions in forensic biology (i.e., questions that ask whether a crime stain consisting of a particular body fluid comes from a particular person). We show how to construct an influence diagram for this example, and how sensitivity analyses lead to an optimal analytical sequence. The key idea is to show that such a Bayesian decisional approach provides a coherent framework for justifying the optimal analytical sequence for a particular case in forensic science.

Keywords: Decision theory, Value of information, Utility, Bayesian decision network, Influence diagram, Forensic science

1 Introduction

In 1984, Kahneman and Tversky wrote the following words to underline the importance of decision-making:

> Making decisions is like speaking prose — people do it all the time, knowingly or unknowingly. [22, p. 341]

This is very true for decisions in forensic science. Decisions in forensic science are of the utmost importance, because the forensic scientist's work may determine whether the true perpetrator of a crime can be found and convicted. Hence, it is highly desirable for forensic scientists to make rational decisions which they can justify in a court of law.

However, there is currently a large void in logical frameworks for making decisions in forensic science. In practice, forensic scientists make decisions by blindly following laboratory protocols, which are not based on any logical framework. The scientific literature in forensic science is currently very limited with regard to models for rational decision-making. This literature currently contains a handful of publications on modeling a forensic scientist's conclusions as a decision: Krawczak and Schmidtke [28] support the minimax decision rule for identifying fathers in disputed paternity cases, Phillips et al. [35] advocate for the use of signal detection theory for forensic individualization conclusions, Biedermann et al. [2] introduce Bayesian decision theory for an identification or individualization conclusion, Biedermann et al. [3] apply Bayesian decision theory to the conclusion of whether an unknown proportion (e.g., of a consignment containing an illegal substance) is greater than a predefined threshold, Gittelson et al. [11] apply Bayesian decision theory to the conclusion of individualizing a person found through a database search, Gittelson et al. [14] apply Bayesian decision theory to the genotype designation of low-template DNA results, and Sironi et al. [42] apply Bayesian decision theory to declaring a person of unknown age a minor or an adult within the context of the law. A few publications attempt to create decision models for a forensic scientist's decision of performing a DNA analysis [43, 44, 50]. And two publications apply Bayesian decision theory to a forensic scientist's decision of performing a laboratory analysis in forensic science: Gittelson et al. [12] present a model for a forensic scientist's decision of whether to process a fingermark, and Gittelson et al. [15] present the results of applying Bayesian decision theory to a forensic scientist's decision of performing a single DNA analysis or two replicate DNA analyses on low-template DNA crime stains. Yet, these publications do not provide any explanation on how to apply Bayesian decision theory, how to construct an influence diagram, or how to perform sensitivity analyses.

The aim of this paper is to fill this void by presenting a step-by-step guide on how to apply Bayesian decision theory to decision problems encountered by forensic scientists on performing or not performing a particular laboratory test or analysis. We follow Raiffa's solution for the oil-wildcatter example [38], because this model is already widely published for medical decision problems on performing or not performing a test, decision problems which are structurally very similar to the forensic scientist's decision problems. Similar structures include, for example, a model for the decisions of testing and then treating an infant born to a mother infected with HIV [34], a model

for the decisions of performing a bacterial culture and then treating with antibiotics a patient with a sore throat [32], and a model for the decisions of performing diagnostic tests and then a thoracotomy on a patient with a lung-cancer tumor [32].

More specifically, this paper illustrates the application of Bayesian decision theory by presenting a decision-theoretic model for source level questions [8] in forensic biology. A source level question asks whether a crime stain consisting of a particular body fluid comes from a particular person. For example, in an assault where the perpetrator is injured and sheds blood, we are interested in whom the bloodstain(s) come(s) from.[1] Making an inference about whom a particular body fluid comes from requires both information about the type of body fluid the DNA comes from and information about the source of the DNA. To address both of these questions, the forensic scientist first performs one or several presumptive tests for the targeted body fluid, and second, submits this crime stain for a DNA analysis. So the forensic scientist performs a sequence of tests and analyses.

A probabilistic framework (e.g., a Bayesian network) that evaluates the combination of the results of multiple such tests and analyses has already been published elsewhere [49], and will not be treated here. What has not been published in previous studies is a model that takes the cost of each of the performed tests and analyses into account. Each presumptive test comes at a cost, and performing a DNA analysis is even more costly. So, the question we ask is, "How can the forensic scientist rationally justify performing a sequence of tests and analyses in a particular case?" That is, "When is it worth performing a presumptive test?" And, "When is it worth performing a DNA analysis?" The forensic problem under scrutiny here is a decisional extension of part of previously published work [49], and can be expressed in the following terms: "Do we need to perform a presumptive test for blood before submitting the sample for a DNA analysis?".

The aim of this paper is to place this decision problem into a decision-theoretic framework composed of actions, states of natures and utilities (these concepts are defined in Section 2.1.1), and to translate these notions into a probabilistic graphical model, that is, an influence diagram. This theory is then applied to a case example. This case study illustrates the use of such a model for making and justifying rational decisions in forensic science.

This paper is structured as follows. Section 2 introduces the reader to the notions of Bayesian decision theory, sensitivity analysis and influence diagrams, notions that will play a role in the decision analysis of the forensic problem approached in Section 3. Section 3 describes the case study and the decision-theoretic solution. Section 4 presents a general discussion, and Section 5 provides a synthesis and a conclusion of the paper.

[1] As opposed to whom the DNA comes from, because there could be background DNA unrelated to the assault on the surface of the bloodstain(s).

2 Methodology

The principal aim of this section is to provide a basis for forensic scientists to solve decision problems based on Bayesian decision theory. The aimed-for model is *normative* [29], meaning that it describes, or prescribes, how a rational decision maker would act, given the available information and his or her objectives and preference ranking. Note that this model may or may not represent a forensic scientist's choice made without a normative model on the basis of, for example, intuition or a pre-established laboratory protocol. It is important to highlight that this model does not intend to replace the forensic scientist in making decisions, but simply intends to provide a transparent tool, or framework – combining notions from probability theory, computational statistics, logic, and Bayesian decision theory – to reach a coherent[2] decision. Further, note that this model is *personal* in the sense that it reflects only one decision maker's valuation of the possible consequences and probability assignments. This makes the model flexible, and presents great opportunities for exploring different alternatives, and then justifying a chosen course of action in a particular case.

This section describes how Bayesian decision theory, sensitivity analyses and influence diagrams contribute to creating this normative model.

2.1 Bayesian decision theory

Bayesian decision theory focuses on choosing and justifying a rational course of action based on the inferences made in the presence of incomplete information [e.g., 39]. The difficulty is that there is uncertainty regarding the consequence of each possible action. The consequence is not only determined by the chosen action, but also by an unknown variable, or condition, referred to as the state of nature (or state of the world). Not knowing the true state of nature, the decision maker is only able to formulate a probability distribution over the space of possible states of nature. This is why the decision maker is uncertain about the resulting consequence (i.e., the combination of the chosen action with the true state of nature).

The decision-making process is not just based on probability distributions describing the unknown states of nature. Decision theory's major feature is that of combining the measure of uncertainty (probability distributions) with values describing the desirability of each of the possible consequences. By weighing the desirabilities of these consequences with the consequences' probabilities of occurring, the rational decision maker chooses the action with the maximum expected desirability. Formally, this is called the maximum expected utility.

[2] We use the word "coherent" following Lindley [29], p. 22:

> (...) it will not be possible to say that a decision is right but only that these decisions cohere, or not. It is the relationships between events or decisions that matter, not the individual events or decisions.

Hence, a normative framework provides constraints that ensure coherence in decision making. These constraints demand that the decision maker's degrees of belief in uncertain events, as well as his or her degrees of satisfaction with the choices' possible consequences, obey the laws of probability.

Here, we present the notions of utilities, expected utility maximization, and value of information. We begin by briefly describing the mathematical notation used throughout the rest of this paper.

2.1.1 Mathematical notation

To structure a decision problem, one has to define:

- an action space $\mathcal{A}$ consisting of an exhaustive[3] list of mutually exclusive[4] actions $a_1, a_2, \ldots, a_m$;
- a random variable Θ consisting of the possible states of nature, which are discrete in our example so that Θ has n possible states denoted $\Theta_1, \Theta_2, \ldots, \Theta_n$.

Further random variables will be denoted with capital roman letters according to the same scheme as Θ: the capital letter without subscripts denotes the variable's set of possible states, and the capital letter with a subscript one of its states. For boolean variables, the capital letter describing the random variable denotes the state "true" of this variable, and this capital letter preceded by the symbol "$\neg$" denotes its negation, that is, the state "false" of this variable. The results presented in this paper consider only discrete random variables so that each has a predefined number of exhaustive and mutually exclusive states. In some cases, the same capital letter will be used to denote multiple appearances of a particular variable in a problem, and we will distinguish between them by using superscripts.

Each variable is characterized by a probability distribution. This distribution represents the degree of belief the decision maker has in each of the states being true at a given point in time. These degrees of belief are described by *subjective* (or *personal* [31]) probabilities [10], which we denote using the notation $Pr(\cdot|\cdot)$.[5] This expression designates the conditional probability of the element(s) to the left of the vertical bar, given the element(s) to the right of the vertical bar. All of the decision maker's degrees of belief are conditional because they are conditioned on his or her knowledge. For example, if we denote by I the decision maker's knowledge at a given point in time, the decision maker's degree of belief in Θ_1 at this point in time is denoted by $Pr(\Theta_1|I)$.

The combination of $\mathcal{A}$ with Θ, i.e. $\mathcal{A} \times \Theta$, produces the space of the actions' possible consequences, which we will call outcomes: $O(a_i, \Theta_j)$, abbreviated by O_{ij}, denotes the outcome obtained when the decision maker chooses action a_i, $i \in \{1, 2, \ldots, m\}$, when state of nature Θ_j, $j \in \{1, 2, \ldots, n\}$, is true.

To quantify the desirabilities of the actions' possible outcomes, we will use utilities, denoted by the following expression:

$$u(a_i, \Theta_j) = u(O_{ij}) \text{ is the utility of outcome } O_{ij}.$$

The decision-theoretic analysis of a problem will compute the expected utility, of each of the possible actions. We will denote it by $\bar{u}(a_i|\cdot)$, $i = 1, 2, \ldots, m$, with the information available to the decision maker to the right of the conditioning bar. For

[3] The list is *exhaustive* when the decision maker inevitably chooses one of the actions in the list. Note that if it is possible for the decision maker to do nothing, then this possibility must be defined as one of the actions for the action space to be exhaustive.

[4] The actions are *mutually exclusive* if the decision maker can never choose more than one of them at one time.

[5] A discussion on the role of subjective probabilities and their relation with frequencies is introduced in [47].

example, $\bar{u}(a_1|I)$ is the decision maker's expected utility for performing action a_1 given the information contained in I.

2.1.2 Utilities

Utilities measure the desirability of each possible outcome, based on the decision maker's objectives and personal preferences. It is a subjective notion (in the sense of personal), meaning that decision makers with different objectives and preferences may have different preference orderings of the possible outcomes.

The practical application of the notion of utilities to decision problems was largely made possible by its concretisation through axioms [33, 40]. It was demonstrated that utilities are numerically measurable quantities, (i.e., numbers that may concretely be obtained through the comparision of gambles), and asserted that these numbers must obey the following axioms, where u_1, u_2 and u_3 are three utilities and p_1, p_2 and p_3 are three probabilities:

- It is possible for the decision maker to order the possible outcomes from best to worst (or to explicitly state her indifference between two, or several, of them).
- These preferences respect the property of transitivity: i.e., if preferences for three outcomes are defined by utilities u_1, u_2, and u_3, such that $u_1 > u_2$ and $u_2 > u_3$, then $u_1 > u_3$ must be true.
- If $u_2 < u_1$, then the outcome described by u_2 is less preferable than any gamble[6] between u_1 and u_2, and the outcome described by u_1 is more preferable than any such gamble.
- If $u_3 < u_2 < u_1$, then it is possible to define a gamble between u_1 and u_3 which is less preferable than u_2, and another gamble (still between u_1 and u_3) which is more preferable than u_2.
- The order in which utilities are combined is irrelevant: i.e., if p_1 and p_2 are two probabilities, then $p_1u_1 + p_2u_2 = p_2u_2 + p_1u_1$.
- Finally, the number of algebraic steps used in combining utilities is irrelevant: i.e., $p_1(p_2u_1 + (1-p_2)u_2) + (1-p_1)u_2 = p_3u_1 + (1-p_3)u_2$, with $p_3 = p_1p_2$.

These axioms provide decision theory with a means for quantifiying and unifying desirabilities of outcomes on a single scale. More specifically, the utility that should be assigned to a particular outcome is precisely defined through the comparison of two gambles. For this, consider a utility scale from 0 to 1, where a utility of 1 is assigned to the most desirable outcome (which we shall denote by O^+):

$$u(O^+) = 1 \ ,$$

and, analogously, 0 is assigned to the least desirable outcome (denoted by O^-):

$$u(O^-) = 0 \ .$$

[6] A gamble between two outcomes means that one of them will occur with a probability p and the other with a probability $(1-p)$.

For the remaining outcomes, the utilities will be somewhere between 0 and 1. The numerical value of any $u(O_{ij})$, $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$, is now determined by comparing the following two gambles:

gamble 1: obtain outcome O_{ij} for sure,
gamble 2: obtain the most desirable outcome O^+ with a probability of $Pr(O^+)$ and the least desirable outcome O^- with a probability of $Pr(O^-) = 1 - Pr(O^+)$.

Which gamble does the decision maker prefer? If $Pr(O^+)$ is very small, the decision maker will choose gamble 1. If $Pr(O^+)$ is very large, the decision maker will choose gamble 2. Thus, there must be a turning point somewhere between these two probabilities, that is, a single value for $Pr(O^+)$, for which the decision maker will be indifferent between gambles 1 and 2. This numerical value for the probability of $Pr(O^+)$ is equal to the decision maker's utility for outcome O_{ij}. In other words:

$$\begin{aligned} u(O_{ij}) &= u(O^+) \times Pr(O^+) + u(O^-) \times (1 - Pr(O^+)) \\ &= 1 \times Pr(O^+) + 0 \times (1 - Pr(O^+)) \\ &= Pr(O^+) \ . \end{aligned} \tag{1}$$

For example, a utility of $u(O_{ij}) = 0.7$ means that the decision maker is indifferent between obtaining outcome O_{ij} for sure and obtaining the best possible outcome with a probability of 0.7 and the worst possible outcome with a probability of 0.3.

2.1.3 Maximizing the expected utility

We consider - as expressed by Lindley [29] - that a rational decision maker wants to maximize the satisfaction she expects to obtain. This corresponds to maximizing her expected utility. For discrete states of nature, the most rational action is therefore:

$$\underset{i}{arg\ max}\ \bar{u}(a_i|\cdot) = \underset{i}{arg\ max} \sum_{j=1}^{n} u(O_{ij}) Pr(\Theta_j|\cdot).$$

This action that maximizes the expected utility is called the *Bayes action* [1]. Numerous examples apply this theory in forensic science [e.g., 2, 3, 11, 14, 42, 45].

2.1.4 Value of information

The probability distribution over the states of nature, Θ, is conditioned on the information the decision maker has. We denoted by I the information the decision maker has at a given point in time. Let this point in time be when the decision maker is first faced with the decision of choosing a given action in $\mathcal{A}$. The question is, should she make this decision with her current knowledge I, or should she acquire an additional piece of information before making the decision? Let us define a random variable E for this additional piece of information. E is a discrete random variable partitioned into q possible states denoted $E_1, E_2, \ldots, E_q$. A rational decision maker (i.e., a decision maker wanting to maximize the satisfaction she expects to obtain from choosing an action in $\mathcal{A}$) will acquire this additional piece of information if its expected value

is greater than the cost of acquiring it. This requires a quantification of the expected value of the additional piece of information.

Bayesian decision theory defines the *expected value of information* ($EVOI$) as the difference between the maximum expected utility *with* this information and the maximum expected utility *without* this information. The maximum expected utility without the additional piece information is:

$$max_i\ \bar{u}(a_i|I) = max_i\ \sum_{j=1}^{n} u(O_{ij})Pr(\Theta_j|I). \tag{2}$$

With the additional piece of information E, the maximum expected utility becomes the weighted average of the maximum expected utilities for each of the different possible realizations of this new piece of information (which is still unknown when the decision maker must decide to acquire or not acquire it):

$$\sum_{k=1}^{q} max_i\ \bar{u}(a_i|E_k, I)Pr(E_k|I) = \sum_{k=1}^{q} max_i\ \sum_{j=1}^{n} u(O_{ij})Pr(\Theta_j|E_k, I)Pr(E_k|I), \tag{3}$$

where $Pr(\Theta_j|E_k, I)$ is the updated probability of Θ_j upon learning E_k, which we call the posterior probability of Θ_j.

The mathematical relationship between the posterior probability $Pr(\Theta_j|E_k, I)$ and the prior probability $Pr(\Theta_j|I)$ is given by Bayes' theorem. The application of Bayes' theorem updates one's initial degree of belief in Θ_j so that this posterior probability assignment takes into account the new piece of information E_k:

$$Pr(\Theta_j|E_k, I) = \frac{Pr(E_k|\Theta_j, I)Pr(\Theta_j|I)}{Pr(E_k|I)}. \tag{4}$$

This updating process may be repeated as many times as necessary.

Inserting Eq. (4) into Eq. (3) produces:

$$\begin{aligned}\sum_{k=1}^{q} max_i\ \bar{u}(a_i|E_k, I)Pr(E_k|I) &= \sum_{k=1}^{q} max_i\ \sum_{j=1}^{n} u(O_{ij})\frac{Pr(E_k|\Theta_j, I)Pr(\Theta_j|I)}{Pr(E_k|I)}Pr(E_k|I) \\ &= \sum_{k=1}^{q} max_i\ \sum_{j=1}^{n} u(O_{ij})Pr(E_k|\Theta_j, I)Pr(\Theta_j|I). \end{aligned} \tag{5}$$

The $EVOI$ of E is the difference between Eq. (5) and Eq. (2), that is:

$$\sum_{k=1}^{q} max_i\ \bar{u}(a_i|E_k, I)Pr(E_k|I) - max_i\ \bar{u}(a_i|I)$$

$$= \sum_{k=1}^{q} max_i \sum_{j=1}^{n} u(O_{ij}) Pr(E_k|\Theta_j, I) Pr(\Theta_j|I) - max_i \sum_{j=1}^{n} u(O_{ij}) Pr(\Theta_j|I). \quad (6)$$

The $EVOI$ is always greater than or equal to zero, reflecting the informative value of additional information [e.g., 18, 37]. If this value is greater than the cost of obtaining E, the decision maker should acquire the additional information. If this value is smaller than the cost, the rational decision maker would choose not to acquire the information. Further information on using the $EVOI$ for solving forensic decision problems can be found in Gittelson et al. [12], Gittelson [13], Gittelson et al. [15].

2.2 Sensitivity analyses

According to Edwards [9]:

> Any decision [...] is, if made under conditions of uncertainty, equivalent to an assessment of a vector of probabilities. [9, p. 338]

Given a set of preferences, the decision between two actions comes down to comparing the posterior probability of a state of nature to a threshold probability calculated as a function of these preferences. Several examples illustrate this in the judicial context [e.g., 4, 23–25].

Consider the discrete states of nature Θ_j, $j \in \{1, 2, \ldots, n\}$. If we are interested, say, in the posterior probability of Θ_1, this refers to the probability of this state after observing a certain amount of evidence, say E_k: $Pr(\Theta_1|E_k)$. This probability is a function of the variables and parameters defined in the model for the relationship between Θ_1 and E_k. Thus, as Edwards [9] states, the decision is actually a function of the uncertain parameters in the model. Variations in these parameters may lead to different posterior probability distributions, which, in turn, may lead to different courses of action. To understand and justify the decision analysis's outcome, it is therefore imperative to study the decision model's behavior through sensitivity analyses. This information makes the decision maker aware of when the value of a parameter must be evaluated with high numerical precision, and when a simple order of magnitude suffices for a coherent decision analysis. Examples of such an approach can be found in forensic literature [e.g., 5, 11, 14, 16, 45].

2.3 Influence diagrams

2.3.1 Definition

Influence diagrams are normative expert systems combining probability and decision theory in a graphical model [e.g., 20, 26]. They consist of:

- *nodes* representing random variables in the form of circles (○), decisions in the form of squares (□), and the utilities in the form of diamonds (◇); and
- *arrows* denoting either the direct probabilistic relationships between these nodes (represented by unbroken arrows), or precedence links that indicate the order in which multiple decisions must be made (represented by dotted arrows) [27].

Each arrow points from a parent node to a child node, and together they form a directed acyclic graph.

Conditional probability tables respecting the direct probabilistic relationships between nodes are associated with each random variable.

An influence diagram is a translation of the elements in the decision problem into a graphical structure. The main advantage of such a graphical structure is its practical capacity of modeling a complex problem [13, 20, 41, 46]. Through its graphical representation it describes the assumed dependence relationships between the various elements of the problem. Underlying this representation, the laws of probability and decision theory rigorously govern the mathematical calculations. An example is developed in Section 3.2.

2.3.2 Structure for forensic decision problems

The basic structure of an influence diagram combines the three types of nodes as shown in Figure 1(a): utilities $u(O_{ij})$, $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$, are assigned to each outcome O_{ij} ($\diamond$), and each outcome O_{ij} depends on action $a_i \in \mathcal{A}$ ($\square$) and on the state of nature $\Theta_j \in \Theta$ ($\circ$).

An observation or test result updates the probability distribution over the possible states of nature. Figure 1(b) shows the most basic Bayesian network[7] for making an inference on Θ given evidence E. Combining this model with the influence diagram in Figure 1(a) produces the influence diagram in Figure 1(c). This model does the same as the influence diagram in Figure 1(a), except that the probability distribution over Θ may now be updated by observing evidence E (by instantiating the observed evidence in node E).

Finally, a decision $\mathcal{T}$ on whether or not to obtain the observation or test result (represented here by evidence E) can precede the main decision $\mathcal{A}$ at a cost of c. Figure 1(d) shows the generic influence diagram for the sequence of decisions *test decision* (denoted $\mathcal{T}$) $\rightarrow$ *terminal decision* (denoted $\mathcal{A}$) [27]. The decision to perform a particular test (node T) is a parent to the random variable E representing the possible observations or analytical results obtained if the test is performed. Node c, modeled as a child of this decision node accounts for the cost of performing this test. The dotted arrow between the action nodes T and A is a precedence link indicating that decision T precedes decision A. Such an influence diagram allows the decision maker to find the optimal course of action in a complex decision problem involving many intricately related random variables and a sequence of decisions.

[7] A Bayesian network (BN) is a graphical probability model containing only random variables as nodes [e.g., 20, 26].

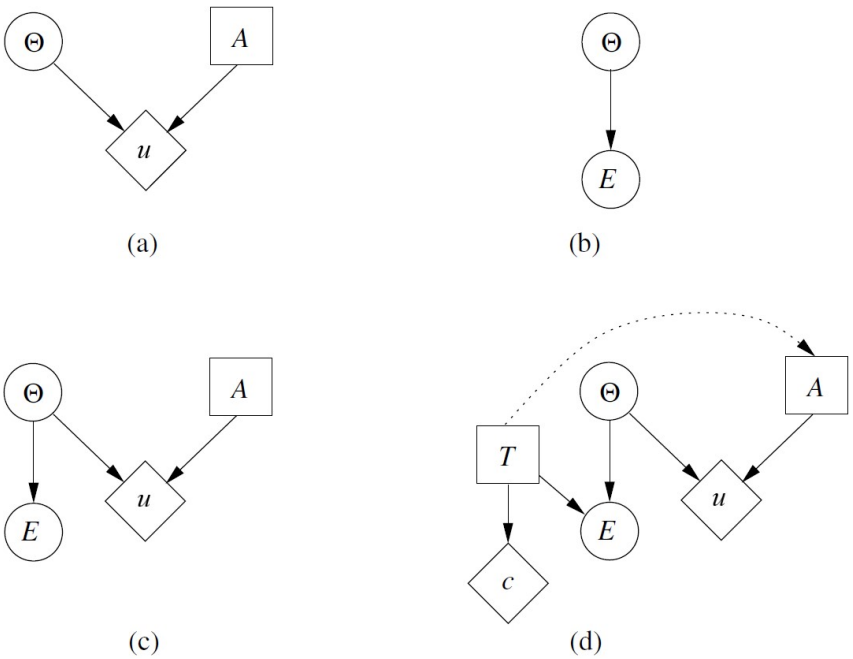


Fig. 1: Generic structures of graphical probability and decision models: (**a**) the general structure of an influence diagram combining actions (node A), utilities (node u), and states of nature (node Θ); (**b**) a Bayesian network for making inferences on the state of nature given a piece of evidence (node E); (**c**) the combination of models (a) and (b) produces an influence diagram for choosing an action in A given evidence E; (**d**) an extension of model (c) to include the preliminary test decision (node T) of performing or not performing the test which produces the evidence in E for a cost specified in node c. Model (d) is the generic model for the sequence of decisions *test decision* $\rightarrow$ *terminal decision* [27]. The unbroken arrows represent the network's functional relationships between the nodes (i.e., the dependence relationships that determine the model's evaluation process), and the dotted arrow stands for a precedence link indicating that decision T precedes decision A.

3 Results

In this section, we present a case study that applies the theory and model presented in Section 2. The forensic decision problem under scrutiny here is: "Should we perform a presumptive test for blood before submitting the sample for a DNA analysis?"

Consider the following scenario. There is an assault, and during this assault the perpetrator is injured and sheds blood. An investigator collects a stain on the crime scene, believing that it may contain blood coming from the offender. If this stain is human blood, the forensic scientist wants to obtain its DNA profile. However, if the stain is not human blood, a DNA analysis would not produce any relevant result, and

the cost of performing this analysis would be wasted.[8] The forensic scientist has the choice of directly performing a DNA analysis, or of first performing a presumptive test for hemoglobin. If she decides to first perform a presumptive test, then she must decide which type of presumptive test, for there are chemical tests that react with hemoglobin without distinguishing between animals and humans, and slightly more expensive immuno-chromatographic tests that react specifically with hemoglobin from primates. It is also possible to perform a sequence of both of these tests. The question is which presumptive test(s), if any, should the scientist perform? This decision deals with determining the sequence of tests to perform on a crime stain before deciding to submit it to a more costly analytical test, the DNA analysis. Hence, the scientist's terminal decision here consists of the action space $\mathcal{A} = \{a_1, a_2\}$, with:

a_1 - perform a DNA analysis,
a_2 - not perform a DNA analysis.

The possible states of nature in this decision problem are:

Θ_1 - the stain is human blood,
Θ_2 - the stain is not human blood,

These make the set of possible outcomes, $\mathcal{O} = \mathcal{A} \times \Theta$, consist of:

O_{11} - performing a DNA analysis and obtaining a profile because the stain is human blood,
O_{12} - performing a DNA analysis and not obtaining a profile because the stain is not human blood,
O_{21} - not performing a DNA analysis (and thus not obtaining a DNA profile), even though the stain is human blood.
O_{22} - not performing a DNA analysis (and thus not obtaining a DNA profile) when the stain is not human blood.

Outcomes O_{21} and O_{22} amount to the same physical result (i.e., not performing a DNA analysis and therefore not obtaining a DNA profile), so we will group them together and call this outcome O_{2-}.

There are different types of presumptive tests for hemoglobin (e.g. Luminol, leuchomalachite green, phenolphthalein, Hemastix, Hemident, Bluestar, Hexagon OBTI, Hemastix and Kastle-Meyer). Comments on performances of such tests can be found in Piva de Almeida et al. [36]. Other studies [e.g., 6, 17, 21] provide further information. Here, for the sake of illustration, we will use two that are used in European forensic genetics laboratories: the chemical Kastle-Meyer test and the immunochromatographic Hexagon OBTI test. The first targets the pseudoperoxidase activity of hemoglobin and will produce a positive result for both animal and human hemoglobin. The second uses antibodies specific for human hemoglobin and will produce a positive result for hemoglobin coming from a primate [19]. We will use the superscripts "KM" and "HO" to distinguish between these two tests, so that for the Kastle-Meyer test we have the test decision $\mathcal{T}^{\mathcal{KM}} = \{t_1^{KM}, t_2^{KM}\}$, with:

[8] Note that the forensic scientist here is interested specifically in obtaining the DNA profile of a bloodstain, and is not interested in the DNA profile of any other biological material of human origin that may be present on the surface on which the stain was recovered.

t_1^{KM} - perform a Kastle-Meyer test,
t_2^{KM} - not perform a Kastle-Meyer test,

and for the immunochromatographic Hexagon OBTI test, we have the test decision $\mathcal{T}^{\mathcal{HO}} = \{t_1^{HO}, t_2^{HO}\}$, with:

t_1^{HO} - perform a Hexagon OBTI test,
t_2^{HO} - not perform a Hexagon OBTI test.

We assume that each of these presumptive tests, if performed, will produce either a positive or a negative result [48]. Let us denote:

E_1^{KM} - positive Kastle-Meyer test result,
E_2^{KM} - negative Kastle-Meyer test result,
E_1^{HO} - positive Hexagon OBTI result,
E_2^{HO} - negative Hexagon OBTI result.

Finally, performing each of these tests comes at a cost, denoted c^{KM} for the cost of performing the Kastle-Meyer test and c^{HO} for the cost of performing the Hexagon OBTI test.

3.1 Utility function

Quantifying the $EVOI$ for each presumptive test requires assigning a utility function to the space of possible outcomes $\mathcal{O}$. In this decision problem, only the combination of a_1 (performing a DNA analysis) and Θ_1 (the stain is human blood) will lead to a DNA profile of the bloodstain. In all of the other combinations of $a_i \in \mathcal{A}$ and $\Theta_j \in \Theta$, the scientist will not obtain a DNA profile of the bloodstain, but will still have to pay for this analysis if a_1 was chosen. The utility function must therefore cover both the gain obtained from acquiring a DNA profile, and the cost produced by performing this analysis. On a monetary scale, this function is:

$u(O_{11}) =$ *gain from DNA profile - cost of DNA analysis*,
$u(O_{12}) =$ *cost of DNA analysis*,
$u(O_{2-}) = 0$.

For example, if the cost of one DNA analysis is 400 if it produces a DNA profile and 300 if it doesn't produce a DNA profile, we would obtain the following utility function:[9]

$u(O_{11}) =$ *gain from DNA profile - 400*,
$u(O_{12}) = -300$,
$u(O_{2-}) = 0$.

This leaves one unknown in the utility function: *gain from DNA profile*. How important is it in the case under investigation to obtain a DNA profile of this stain if this stain is indeed blood? Is this profile essential for finding and convicting the

[9] We purposely did not specify a monetary unit, as these costs will vary from lab to lab and from country to country, and these numbers are for illustrative purposes only.

offender? The answers to these questions will vary from one case to another. We assume in this study that outcome O_{11} is the most desirable, so that

$$u(O_{11}) > u(O_{2-}) \ ,$$

in other words, this means that in this case

$$\textit{gain from DNA profile} > 400 \ .$$

In Section 2.1.2, we saw how a utility function is defined by the probabilities of the most desirable outcome that make the decision maker indifferent between a gamble of obtaining this most desirable outcome with the chosen probability and the least desirable outcome with one minus this probability, and the gamble of obtaining the intermediate outcome for sure. Here, if the forensic scientist finds it difficult to put a numerical value on the *gain from DNA profile*, she may prefer to define the function on a scale from 0 (depicting the worst possible outcome) to 1 (for the best possible outcome), and then applying a linear transformation to obtain the equivalent function in monetary units. On a scale from 0 to 1, this utility function is given by:

$u(O_{11}) = 1$
$u(O_{12}) = 0$
$u(O_{2-}) = \kappa$,

where $0 < \kappa < 1$ is defined through the comparison of gambles explained in Section 2.1.2. Hence, the value of κ is equal to the probability that makes the forensic scientist indifferent between obtaining outcome O_{2-} (not performing a DNA analysis and therefore not obtaining a DNA profile) for sure, and obtaining outcome O_{11} (performing a DNA analysis and obtaining a DNA profile of the bloodstain) with a probability of κ and outcome O_{12} (performing a DNA analysis and not obtaining a DNA profile of the bloodstain) with a probability of $1 - \kappa$ (see the explanation in Section 2.1.2 for more details). The smaller κ, the more the forensic scientist wants to obtain a DNA profile, because the more she is willing to pay to perform the analysis when there is a small probability of obtaining a result. In other words, the smaller κ, the greater the gain obtained from a DNA profile. In terms of κ, the monetary gain acquired from a DNA profile in this example of the utility function is given by:

$$\textit{gain from DNA profile} = \frac{300}{\kappa}(1 - \kappa) + 400 \ .$$

Or, a given monetary gain from a DNA profile is equivalent to setting κ equal to:

$$\kappa = \frac{300}{\textit{gain from DNA profile} - 100} \ .$$

This utility function is plugged into Eq. (6) to evaluate the $EVOI$ for each of the presumptive tests.

3.2 Complete decision model – Influence diagram

The complete model consists of a sequence of two test decisions (the Kastle-Meyer test and the Hexagon OBTI test) and one terminal decision (performing a DNA analysis). The Kastle-Meyer's test result provides information about whether the crime stain is blood, and the Hexagon OBTI's test result provides information about whether the crime stain is human blood. We already have a variable for whether the stain is human blood: the decision problem's state of nature Θ. Yet we do not have a random variable for whether the stain is blood, so we add the boolean variable B with states:

B - the stain is blood,
$\neg B$ - the stain is not blood.

The state of nature Θ depends on whether the stain is blood (B) and on whether the stain is of human origin, so we must add the boolean variable G for whether the stain is of human origin. Its states are:

G - the stain is of human origin,
$\neg G$ - the stain is not of human origin.

We can now connect the random variables, decisions, utilities and costs to create the influence diagram for this decision problem. This influence diagram is shown in Figure 2. It is the influence diagram for the test decisions of both performing a Kastle-Meyer test (node T^{KM}) and performing a Hexagon OBTI test (node T^{HO}) on the crime stain before deciding whether to submit it for a DNA analysis.

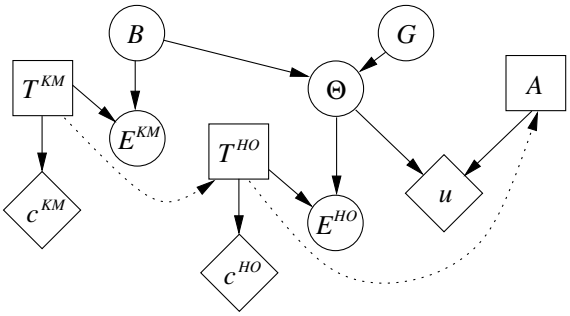


Fig. 2: An influence diagram for the sequence of decisions of performing a presumptive test for blood (node T) before deciding to submit an item for DNA typing (node A). The utility function in node u models the satisfaction obtained from decision A in function of whether the stain is human blood (node Θ). The prior probability distribution over this node depends on whether the stain is blood (node B) and on whether the stain is of human origin (node G). Here there are two presumptive tests possible: first, the scientist may perform a Kastle-Meyer test (node T^{KM}), specific for hemoglobin, producing result E^{KM} at a cost of c^{KM}; second, the scientist may perform a Hexagon OBTI test (node T^{HO}), specific for the hemoglobin of primates, producing result E^{HO} at a cost of c^{HO}.

If both tests are applied, the Kastle-Meyer test always precedes the Hexagon OBTI test, because the forensic scientist proceeds from the most general test to the most specific test. In this case (and in most cases), the most general test is also the least expensive test. Here, we fix the cost of the Kastle-Meyer test to be 0.60 monetary units, and the cost of the Hexagon OBTI test to be 5 monetary units. These costs are contained in nodes c^{KM} and c^{HO}, respectively. Nodes B and G in this influence diagram are the nodes for the boolean variables B (the stain is blood) and G (the stain is of human origin).

The conditional probability table for Θ is filled out as follows:

$$\Theta_1 = \begin{cases} 1 & \text{if } B \text{ and } G \text{ are both true ,} \\ 0 & \text{in all other cases .} \end{cases}$$

Nodes E^{KM} and E^{HO} model the test result of each of the presumptive tests, and have the following conditional probability tables:

		B	$\neg B$
E^{KM} :	E_1^{KM}	$1-\beta^{KM}$	α^{KM}
	E_2^{KM}	β^{KM}	$1-\alpha^{KM}$

		Θ_1	Θ_2
E^{HO} :	E_1^{HO}	$1-\beta^{HO}$	α^{HO}
	E_2^{HO}	β^{HO}	$1-\alpha^{HO}$

where α represents the probability of a false positive and β the probability of a false negative.[10] According to the literature, the Kastle-Meyer test may produce false positives when in contact with certain types of food (e.g., potatos, tomatos, red kidney bean, horseradish), bleach solutions and materials [e.g., 7, 51], and the Hexagon OBTI test may produce false positives when in contact with certain detergents and bleach solutions, in addition to blood from primates other than humans [e.g., 19]. A false negative may occur when the blood is highly diluted, that is, diluted more than 1:100,000 for the Kastle-Meyer test [51] and more than 1:1,000,000 for the Hexagon OBTI test [19]. As the false positive and false negative probabilities are highly dependent on the case circumstances,[11] it is not possible to put a single, general numerical value on each of these. Instead, we examine the impact of α and β, along with the impact of B, G, and κ in the sensitivity analyses in Section 3.3.

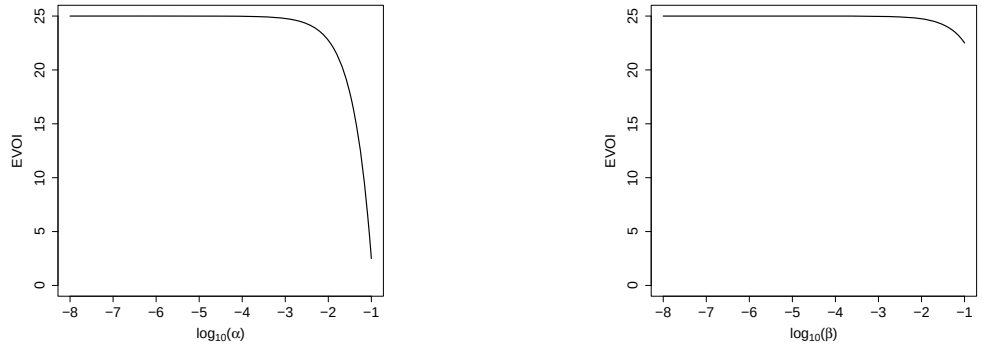
3.3 Sensitivity analysis

The $EVOI$ associated with each of the tests is a function of:

- the false positive and false negative probabilities, α and β,

[10] The literature also commonly uses the terms *sensitivity* and *specificity*, which are $1-\beta$ and $1-\alpha$, respectively.

[11] For example, was a gorilla present on the crime scene? Or, more realistically, what is the probability that the surface on which the trace was recovered was freshly cleaned with a bleach solution?



(a) (b)

Fig. 3: The $EVOI$ of a test (here for the Hexagon OBTI test) depends on (**a**) α (the test's false positive probability) and (**b**) β (the test's false negative probability). These graphs show that the $EVOI$ is only affected by α and β if these probabilities are very high. In other words, if $\alpha < 10^{-3}$, the $EVOI$ is independent of α's precise numerical value, and if $\beta < 10^{-2}$, the $EVOI$ is independent of β's precise numerical value.

- the prior probability distributions over B and G (which determine the probability distribution over Θ), and
- parameter κ in the forensic scientist's utility function.

3.3.1 False positive and false negative probabilities

Figure 3 presents the $EVOI$ in function of α and β.

According to Figure 3(a), the exact value of α is of no importance when $\alpha < 10^{-3}$. And according to Figure 3(b), the exact value of β is of no importance at all, except if β would take unrealistically high values of 10^{-1} or 10^{-2}. This means that unless there are particular case circumstances that are favorable to false positives (i.e., making $\alpha \geq 10^{-3}$ and/or $\beta \geq 10^{-2}$), assigning more precise numerical values for α and β does not change the $EVOI$ of the test. For the sensitivity analyses that follow, we will use α equal to 10^{-3} and 10^{-1} and β equal to 10^{-3} and 10^{-1} for both the Kastle-Meyer and the Hexagon OBTI tests.

3.3.2 Prior probability distributions over B and G and parameter κ

A presumptive test is useful when there is uncertainty surrounding the presence of the test's target molecule. For the test to be useful, this uncertainty (quantified here by the probabilities $Pr(B|I)$ and $Pr(\Theta_1|I)$) must fall into a particular range, that is, a range of probabilities where the test's result will have an impact on the choice of the terminal action. In other words, if the probability that the target molecule is present is already very high, a presumptive test will probably produce a positive result, and this will add very little information to what is already known. Conversely,

if the probability of the target molecule being present is extremely small, the test will probably produce a negative result, also adding very little information. And if the test result does happen to go in the opposite direction of what is expected (i.e., negative for a very high prior probability, or positive for a very low prior probability), then its value may still be too small to counterbalance prior probabilities that are very close to 0 or to 1, so that the scientist would not change her choice of $a_i \in \mathcal{A}$ upon learning the test result. Thus, a presumptive test is useful when the probability of the target molecule is in a range excluding values very close to 0 and values very close to 1.

So, when is it worth performing a Kastle-Meyer test, and when is it worth performing a Hexagon OBTI test? Figure 4 presents the $EVOI$ (solid black line) of a Kastle-Meyer test in function of $Pr(B|I)$ and given values for κ, α, β and $Pr(G|I)$. Figure 5 presents the $EVOI$ (solid black line) of a Hexagon OBTI test in function of $Pr(\Theta_1|I)$ and given values for κ, α and β. For κ, the values 0.1, 0.5 and 0.9 were chosen to cover the typical range of cases a forensic scientist encounters: (1) a serious case where it is worth investing in a DNA analysis even when there is as little as a 0.1 probability of obtaining a DNA profile ($\kappa = 0.1$), (2) a case of medium severity where it is worth investing in a DNA analysis when the odds of obtaining a DNA profile are 50:50 or higher ($\kappa = 0.5$), and (3) a case of low severity where it is only worth investing in a DNA analysis when there is a probability of 0.9 or greater of obtaining a DNA profile ($\kappa = 0.9$). These graphs also plot the maximum expected utility without the test (dotted red line) and the maximum expected utility with the test (dashed green line), whose difference produce the $EVOI$ (Eq. (6)).

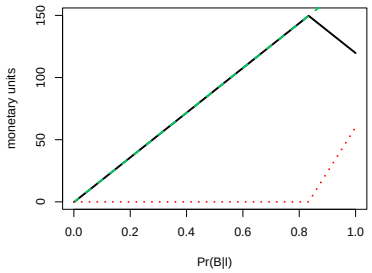
The maximum expected utility with the test is a linearly increasing function of $Pr(B|I)$ for the Kastle-Meyer test (Fig. 4) and a linearly increasing function of $Pr(\Theta_1|I)$ for the Hexagon OBTI test (Fig. 5), with a minimum value equal to $u(O_{2-})$ for a probability of 0 and a maximum value equal to $u(O_{11})$ for a probability of 1.[12] The maximum expected utility without the test is first equal to $u(O_{2-})$ when $Pr(B|I)$ takes values from 0 to $\frac{\kappa}{Pr(G|I)}$ for the Kastle-Meyer test (Fig. 4) and when $Pr(\Theta_1|I)$ takes values from 0 to κ for the Hexagon OBTI test (Fig. 5), and then after this threshold, its value increases linearly[13] to the maximum value of $u(O_{11})$ for a probability of 1. Since the $EVOI$ is the difference between these two functions, its maximum is at the thresholds $Pr(B|I) = \frac{\kappa}{Pr(G|I)}$ (Fig. 4) and $Pr(\Theta_1|I) = \kappa$ (Fig. 5), where it is equal to the maximum expected utility with the test.

More specifically, these sensitivity analyses produce the following results for the $EVOI$:

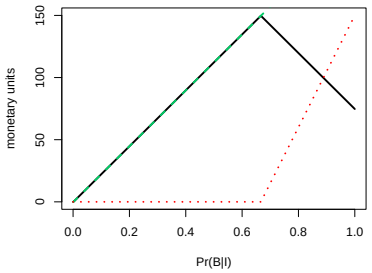
- **Kastle-Meyer test:** The Kastle-Meyer test has an $EVOI$ equal to 0 for $Pr(G|I) \leq \kappa$. When $Pr(G|I) \leq \kappa$, a_2 will always maximize the expected utility, even when the test's result is positive, because the low $Pr(G|I)$ prevents $Pr(\Theta_1|I)$ from exceeding the threshold necessary for a_1 to be the Bayes action. The information provided by a Kastle-Meyer test is therefore useless whenever $Pr(G|I) \leq \kappa$. However, when

[12] These extremes represent perfect information which attain the maximum utility values for performing a DNA analysis (i.e., $u(O_{11})$) and for not performing a DNA analysis (i.e., $u(O_{2-})$).

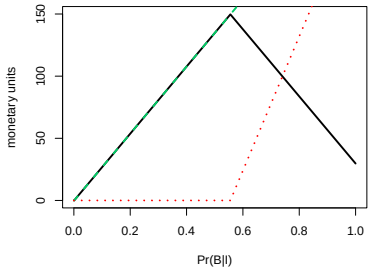
[13] Here, the maximum expected utility without the test is equal to $Pr(B|I) \times u(O_{11}) + Pr(\neg B|I) \times u(O_{12})$ for the Kastle-Meyer test (Fig. 4) and to $Pr(\Theta_1|I) \times u(O_{11}) + Pr(\Theta_2|I) \times u(O_{12})$ for the Hexagon OBTI test (Fig. 5).



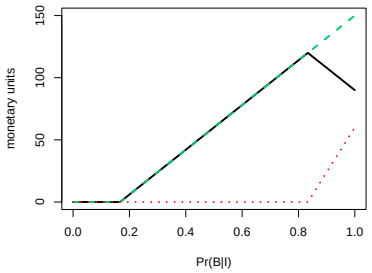
(a) $Pr(G|I) = 0.60$, $\alpha = \frac{1}{1000}$, $\beta = \frac{1}{1000}$



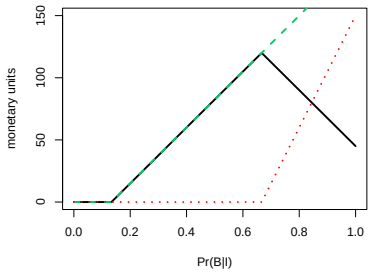
(b) $Pr(G|I) = 0.75$, $\alpha = \frac{1}{1000}$, $\beta = \frac{1}{1000}$



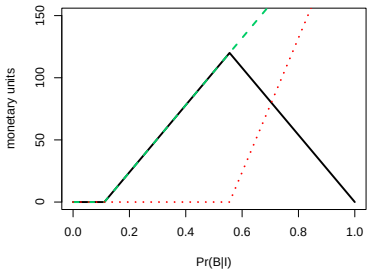
(c) $Pr(G|I) = 0.90$, $\alpha = \frac{1}{1000}$, $\beta = \frac{1}{1000}$



(d) $Pr(G|I) = 0.60$, $\alpha = \frac{1}{10}$, $\beta = \frac{1}{10}$

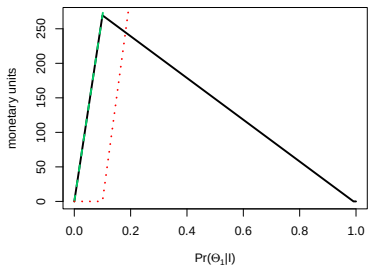


(e) $Pr(G|I) = 0.75$, $\alpha = \frac{1}{10}$, $\beta = \frac{1}{10}$

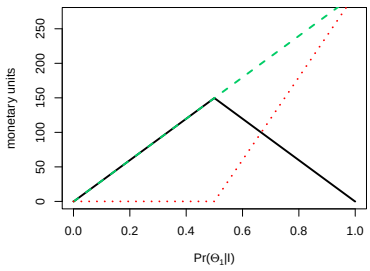


(f) $Pr(G|I) = 0.90$, $\alpha = \frac{1}{10}$, $\beta = \frac{1}{10}$

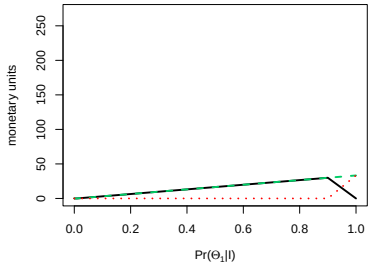
Fig. 4: The $EVOI$ of the Kastle-Meyer test in function of the probability of the crime stain being blood, $Pr(B|I)$, for $\kappa = 0.5$, for values of $Pr(G|I)$ of 0.60, 0.75 and 0.90, and for values of α and β of $\frac{1}{1000}$ and $\frac{1}{10}$. The solid black line depicts the $EVOI$, the dotted red line the maximum expected utility without performing the test, and the dashed green line the maximum expected utility with performing the test.



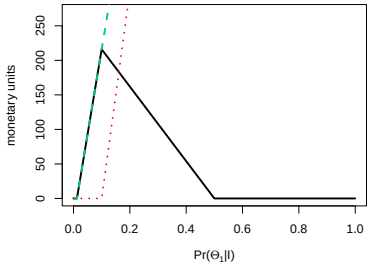
(a) $\kappa = 0.1$, $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$



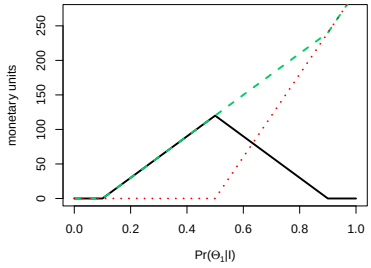
(b) $\kappa = 0.5$, $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$



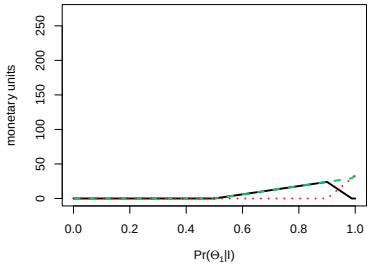
(c) $\kappa = 0.9$, $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$



(d) $\kappa = 0.1$, $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$



(e) $\kappa = 0.5$, $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$



(f) $\kappa = 0.9$, $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$

Fig. 5: The $EVOI$ of the Hexagon OBTI test in function of the probability of the crime stain being human blood, $Pr(\Theta_1|I)$, for values of κ of 0.1, 0.5, and 0.9, and for values of α and β of $\frac{1}{1000}$ and $\frac{1}{10}$. The solid black line depicts the $EVOI$, the dotted red line the maximum expected utility without performing the test, and the dashed green line the maximum expected utility with performing the test.

$Pr(G|I) > \kappa$, the $EVOI$ attains its maximum value for

$$Pr(B|I) = \frac{\kappa}{Pr(G|I)} \ .$$

Figure 4 shows this. Further, the greater $Pr(G|I)$, the greater is the maximum value of the $EVOI$. This is because the $EVOI$ is a function of the expected utility of a_1, which is a function of $Pr(\Theta_1|I)$, and a greater $Pr(G|I)$ produces a greater $Pr(\Theta_1|I)$. Higher false positive and false negative probabilities (i.e., $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$) produce slightly lower values for the $EVOI$ than lower false positive and false negative probabilities (i.e., $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$) because the test produces less information when the probabilities of a false positive or a false negative are higher. Higher false positive and false negative probabilities also produce $EVOI = 0$ for very low values of $Pr(B|I)$, because a positive test result in this case would not provide enough information to justify performing a DNA analysis. In our example with $c^{KM} = 0.60$ monetary units, it is worth performing the Kastle-Meyer test whenever

$$EVOI^{KM} \geq 0.60 \ .$$

For the instances plotted in Figure 4 where $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$, this is the case when:

for $Pr(G\|I) = 0.60$:	$0.011 \leq Pr(B\|I) \leq 0.997$,	i.e.,	$0.007 \leq Pr(\Theta_1\|I) \leq 0.598$
for $Pr(G\|I) = 0.75$:	$0.005 \leq Pr(B\|I) \leq 0.997$,	i.e.,	$0.004 \leq Pr(\Theta_1\|I) \leq 0.748$
for $Pr(G\|I) = 0.90$:	$0.003 \leq Pr(B\|I) \leq 0.997$,	i.e.,	$0.003 \leq Pr(\Theta_1\|I) \leq 0.897$

For the instances plotted in Figure 4 where $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$, this is the case when:

for $Pr(G\|I) = 0.60$:	$0.170 \leq Pr(B\|I) \leq 1$,	i.e.,	$0.102 \leq Pr(\Theta_1\|I) \leq 0.600$
for $Pr(G\|I) = 0.75$:	$0.136 \leq Pr(B\|I) \leq 1$,	i.e.,	$0.102 \leq Pr(\Theta_1\|I) \leq 0.750$
for $Pr(G\|I) = 0.90$:	$0.114 \leq Pr(B\|I) \leq 1$,	i.e.,	$0.103 \leq Pr(\Theta_1\|I) \leq 0.900$

- **Hexagon OBTI test:** The Hexagon OBTI test has a maximum $EVOI$ for $Pr(\Theta_1|I) = \kappa$. This is shown in Figure 5. In addition, Figure 6 plots the $EVOI$ in function of κ, showing how the $EVOI$ increases to this maximum value and then decreases. The smaller κ, the greater is the maximum value of the $EVOI$. This is because a smaller κ represents a greater gain from obtaining a DNA profile, making the information on the possible success of obtaining a DNA profile more valuable. Like for the Kastle-Meyer test, higher false positive and false negative probabilities produce slightly lower values for the $EVOI$, as well as $EVOI = 0$ for very low values of $Pr(\Theta_1|I)$. In our example with $c^{HO} = 5.00$ monetary units, the Hexagon OBTI test is worth performing whenever

$$EVOI^{HO} \geq 5.00 \ .$$

In the instances presented in Figure 5 where $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$, this is the case when:

for $\kappa = 0.1$:	$0.002 \leq Pr(\Theta_1\|I) \leq 0.983$
for $\kappa = 0.5$:	$0.017 \leq Pr(\Theta_1\|I) \leq 0.983$
for $\kappa = 0.9$:	$0.151 \leq Pr(\Theta_1\|I) \leq 0.983$

In the instances presented in Figure 5 where $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$, this is the case when:

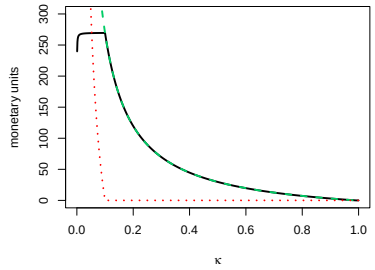
for $\kappa = 0.1$:	$0.015 \leq Pr(\Theta_1\|I) \leq 0.490$
for $\kappa = 0.5$:	$0.117 \leq Pr(\Theta_1\|I) \leq 0.883$
for $\kappa = 0.9$:	$0.584 \leq Pr(\Theta_1\|I) \leq 0.969$

3.3.3 Both presumptive tests in sequence

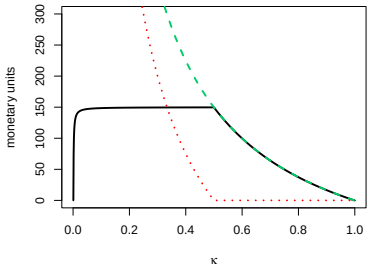
With two presumptive tests, there is also the possibility of performing both tests in sequence. So, when is it worth performing both tests? Given a positive result from a Kastle-Meyer test, is it worth performing the Hexagon OBTI test? Figure 7 plots the $EVOI^{HO}$ for a case where the Kastle-Meyer test has already produced a positive result. The graphs in this figure plot the $EVOI^{HO}$ and the maximum expected utility values with and without this additional test in function of the prior probability of the stain being human blood ($Pr(\Theta_1|I)$), that is, the probability of the stain being human blood before learning that the Kastle-Meyer test produced a positive result. These graphs show that the sequence of tests Kastle-Meyer $\rightarrow$ Hexagon OBTI is the most cost effective for small prior probabilities for Θ_1 when $\kappa = 0.1$ and $\kappa = 0.5$, and for larger prior probabilities for Θ_1 when $\kappa = 0.9$. This is because the forensic scientist is willing to risk performing a DNA analysis and not getting a DNA profile more for lower values of κ than for a high value of κ. Hence, the forensic scientist will risk performing a DNA analysis with less information for lower values of κ so that the value of additional information decreases and plateaus as the prior probability $Pr(\Theta_1|I)$ increases for these values of κ (Fig. 7(a),(b),(d) and (e)). For a large value of κ, the forensic scientist is more risk-adverse and will therefore require more information in order to perform a DNA analysis. Here, the graphs (Fig. 7(c) and (f)) show that the value of the Hexagon OBTI test is beneficial for larger prior probabilities for Θ_1, because for these prior probabilities the information of this second test will provide the required useful information for making the decision of performing or not performing a DNA analysis. These graphs indicate that the very small prior probabilities for Θ_1 are too small for the information of the two tests to suffice for a justification of performing a DNA analysis. When comparing the $EVOI$ values with the cost of the Hexagon OBTI test ($c^{HO} = 5.00$), we note that the cost of this test is low enough that it is cost effective to perform the Hexagon OBTI test for $0.001 \leq Pr(\Theta_1|I) \leq 1$ when $\alpha^{KM} = \alpha^{HO} = \frac{1}{1000}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{1000}$, and for $\kappa = 0.1$, $\kappa = 0.5$, and $\kappa = 0.9$.

Where $\alpha^{KM} = \alpha^{HO} = \frac{1}{10}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{10}$, this range of probabilities is:

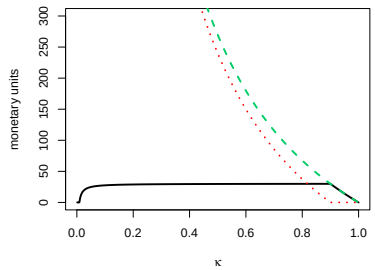
for $\kappa = 0.1$:	$0.002 \leq Pr(\Theta_1\|I) \leq 0.105$
for $\kappa = 0.5$:	$0.015 \leq Pr(\Theta_1\|I) \leq 0.769$
for $\kappa = 0.9$:	$0.153 \leq Pr(\Theta_1\|I) \leq 1$



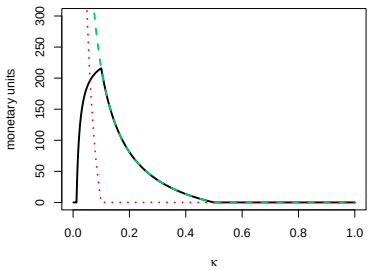
(a) $\Theta_1 = 0.1$, $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$



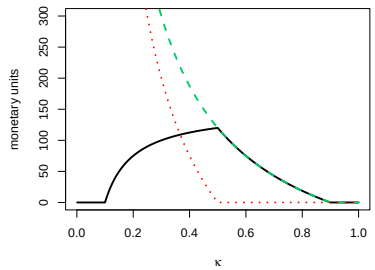
(b) $\Theta_1 = 0.5$, $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$



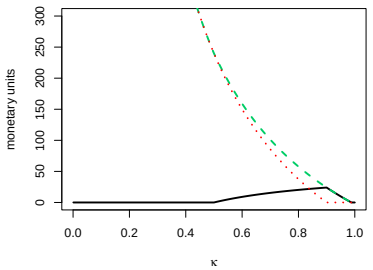
(c) $\Theta_1 = 0.9$, $\alpha = \frac{1}{1000}$ and $\beta = \frac{1}{1000}$



(d) $\Theta_1 = 0.1$, $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$

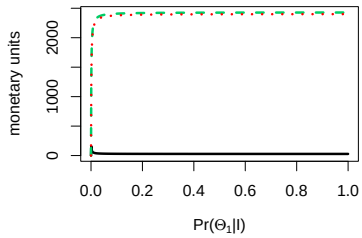


(e) $\Theta_1 = 0.5$, $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$

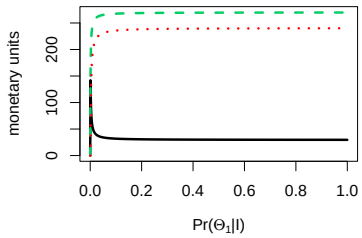


(f) $\Theta_1 = 0.9$, $\alpha = \frac{1}{10}$ and $\beta = \frac{1}{10}$

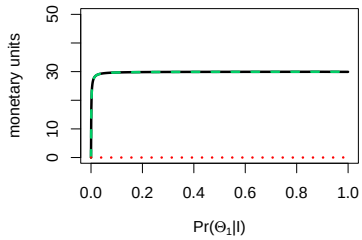
Fig. 6: The $EVOI$ of the Hexagon OBTI test in function of κ, the probability that makes the forensic scientist indifferent between obtaining outcome O_{2-} for sure and obtaining outcome O_{11} with a probability of κ and outcome O_{12} with a probability of $1 - \kappa$, plotted here for values of $Pr(\Theta_1|I)$ of 0.1, 0.5, and 0.9, and for values of α and β of $\frac{1}{1000}$ and $\frac{1}{10}$. The solid black line depicts the $EVOI$, the dotted red line the maximum expected utility without performing the test, and the dashed green line the maximum expected utility with performing the test.



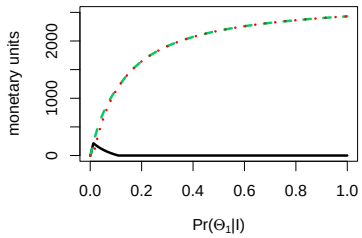
(a) $\kappa = 0.1$, $\alpha^{KM} = \alpha^{HO} = \frac{1}{1000}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{1000}$



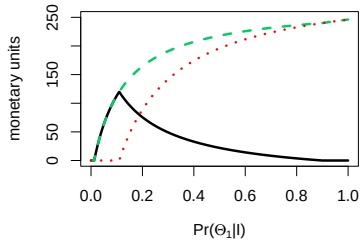
(b) $\kappa = 0.5$, $\alpha^{KM} = \alpha^{HO} = \frac{1}{1000}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{1000}$



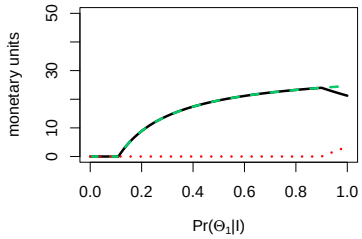
(c) $\kappa = 0.9$, $\alpha^{KM} = \alpha^{HO} = \frac{1}{1000}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{1000}$



(d) $\kappa = 0.1$, $\alpha^{KM} = \alpha^{HO} = \frac{1}{10}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{10}$



(e) $\kappa = 0.5$, $\alpha^{KM} = \alpha^{HO} = \frac{1}{10}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{10}$



(f) $\kappa = 0.9$, $\alpha^{KM} = \alpha^{HO} = \frac{1}{10}$ and $\beta^{KM} = \beta^{HO} = \frac{1}{10}$

Fig. 7: The *EVOI* of the Hexagon OBTI test after having obtained a positive result from the Kastle-Meyer test in function of the prior probability of the stain being human blood $Pr(\Theta_1|I)$, for values of κ of 0.1, 0.5, and 0.9, and for values of $\alpha^{KM} = \alpha^{HO}$ and $\beta^{KM} = \beta^{HO}$ of $\frac{1}{1000}$ and $\frac{1}{10}$. The solid black line depicts the *EVOI*, the dotted red line the maximum expected utility without performing the Hexagon OBTI test, and the dashed green line the maximum expected utility with performing the Hexagon OBTI test.

For higher probabilities of false positives and false negatives, the $EVOI$ values are slightly lower, and therefore the range of prior probabilities for which such a test would be cost effective is more restricted.

3.4 Decision strategy

The results of these sensitivity analyses lead to the following decision strategy for our example:

If $Pr(G|I) \leq \kappa$:

- if $Pr(\Theta_1|I)$ falls into the range defined by κ for the Hexagon OBTI test
  $\rightarrow$ **perform the Hexagon OBTI test**.
- if $Pr(\Theta_1|I)$ does not fall into the range defined by κ for the Hexagon OBTI test
  $\rightarrow$ **do not perform any presumptive test**.
  - if $Pr(\Theta_1|I)$ is close to 0 $\rightarrow$ **do not perform the DNA analysis**.
  - if $Pr(\Theta_1|I)$ is close to 1 $\rightarrow$ **perform the DNA analysis**.

If $Pr(G|I) > \kappa$:

- if $Pr(\Theta_1|I)$ falls into the range defined by κ for the Hexagon OBTI test
  $\rightarrow$ **perform the Hexagon OBTI test**.
- if $Pr(\Theta_1|I)$ falls into the range defined by κ and $Pr(G|I)$ for the Kastle-Meyer test
  $\rightarrow$ **perform the Kastle-Meyer test**.
- if $Pr(\Theta_1|I)$ falls into the range defined for the Hexagon OBTI test after a positive Kastle-Meyer result
  $\rightarrow$ **perform the Kastle-Meyer test**,
  and if Kastle-Meyer produces a positive result
  $\rightarrow$ **perform the Hexagon OBTI test**.
- if $Pr(\Theta_1|I)$ does not fall into any of the above ranges
  $\rightarrow$ **do not perform any presumptive test**.
  - if $Pr(\Theta_1|I)$ is close to 0 $\rightarrow$ **do not perform the DNA analysis**.
  - if $Pr(\Theta_1|I)$ is close to 1 $\rightarrow$ **perform the DNA analysis**.

Note that the overlap of the ranges defined by κ for the Hexagon OBTI test and by κ and $Pr(G|I)$ for the Kastle-Meyer test for $Pr(G|I) > \kappa$, make it possible that performing each of these tests is a cost-effective choice. In this case, a more in-depth analysis focusing specifically on the values of $Pr(G|I)$, $Pr(B|I)$ and κ in the case at hand is required to see which course of action maximizes the expected utility.

4 Discussion

The model presented in this study contains several test decisions for a single terminal decision. With this case study, we have shown how a decision-theoretic analysis of such a case allows the forensic scientist to come up with the optimal analytical sequence for a particular case.

There is a clear need to apply a coherent framework to make and justify decisions based on the available information. The major difficulty in making rational decisions is the inevitable presence of uncertainty. This uncertainty is due to incomplete knowledge of the event which ultimately determines how desirable the outcome of the decision is. In forensic casework, this uncertainty is reflected in questions such as, "What traces are present on an item of evidence under examination?" or "Where do these traces come from?". These questions place the forensic scientist in front of, not only a single decision on what test or analysis to perform, but, a whole sequence of such decisions. All of these decisions are interconnected, and together the chosen actions ultimately determine the value of the scientific evidence in a given case.

However, it is common that scientists do not bother with decision analysis because they are sure that they will make excellent choices without a formal analysis. Yet, how the quality of a decision is measured in the absence of decision analysis remains a mystery. A justification of an action, even if only implicit, should be supported:

> More often than not, the decisions you make in your personal or professional life can be made without a lot of fuss. Either your best choice is clear to you without much analysis, or the decision is not important enough to warrant any great amount of attention. Occasionally, however, you probably find yourself in a situation where you feel it is worth your time and effort to think systematically and hard about the different courses of action you might pursue. You might even be willing to push a few numbers around, if you thought it would help you make a better decision. [38, p. ix]

Forensic scientists have not always appreciated that they can make a contribution to decision making; that they can not only present the data informatively by using the coherent metric known as the Bayes factor or the likelihood ratio, but also explain how those data can be used to assist in choosing the most rational course of action [30]. The decisions made by forensic scientists are an integral part of the judicial process that may lead to the conviction or acquittal of an individual. It is therefore crucial that these decisions be made and justified on rational foundations, such as those provided by Bayesian decision theory. The model presented in this paper provides a foundational framework for making rational decisions in forensic science about obtaining additional information at a given cost, and its application provides a means for choosing and justifying a sequence of analytical tests.

5 Conclusion

Decision analysis plays an important role in forensic science. It is becoming more and more important for forensic scientists to justify their choices (e.g., a given analytical sequence of tests and analyses in the laboratory). Therefore, it is desirable for scientists to conform to a line of reasoning that can be explained and where the chosen course of action is justified. Here, a typical example of choosing a forensic analytical sequence has been introduced and developed by putting forward the role of Bayesian decision theory, the importance of sensitivity analyses, and the support offered by probabilistic graphical models called influence diagrams.

Acknowledgments. This research was funded by the Swiss National Science Foundation grant numbers 100014-135340 and 100011_204554/1. The authors would also like to thank two anonymous reviewers for their valuable comments.

Declarations

Funding. This research was funded by the Swiss National Science Foundation grant numbers 100014-135340 and 100011_204554/1.

Competing interests. The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Ethics approval. Not applicable

Consent to participate. Not applicable

Consent for publication. Not applicable

Availability of data and materials. Not applicable

Code availability. Not applicable

Authors' contributions. Conceptualization: Franco Taroni; Methodology: Simone Gittelson; Formal analysis and investigation: Simone Gittelson; Writing—original draft preparation: Simone Gittelson; Writing—review and editing: Simone Gittelson and Franco Taroni; Funding acquisition: Franco Taroni.

References

[1] Berger JO (2010) Statistical decision theory and Bayesian analysis, 2nd edn. Springer-Verlag, New York

[2] Biedermann A, Bozza S, Taroni F (2008) Decision theoretic properties of forensic identification: Underlying logic and argumentative implications. Forensic Sci Int 177:120–132

[3] Biedermann A, Bozza S, Garbolino P, Taroni F (2012) Decision-theoretic analysis of forensic sampling criteria using Bayesian decision networks. Forensic Sci Int 223:217–227

[4] Biedermann A, Bozza S, Taroni F, Vuille J (2020) Computational normative decision support structures of forensic interpretation in the legal process. SCRIPTed: A Journal of Law, Technology and Society 17:83–124

[5] Bozza S, Broseus J, Esseiva P, Taroni F(2014) Bayesian classification criterion for forensic multivariate data. Forensic Sci Int 244:295–301

[6] Butler J, Chaseling J, Wright K (2019) A comparison of four presumptive tests for the detection of blood on dark materials. J Forensic Sci 64:1838–1843

[7] Castello A, Francès F, Verdu F (2017) Bloodstains on leather: examination of false negatives in presumptive test and human hemoglobin test. J Forensic Sci 62:1308–13133

[8] Cook R, Evett IW, Jackson G, Jones PJ, Lambert JA (1998) A hierarchy of propositions: Deciding which level to address in casework. Sci Just 38:231–240

[9] Edwards W (1988) Summing up: The society of Bayesian trial lawyers. In Tillers P, Green ED (eds) Probability and inference in the law of evidence: The uses and limits of Bayesianism. Kluwer Academic Publishers, Dordrecht, pp. 337–342

[10] de Finetti B (1939) The theory of probability in its relationships with the analysis. In Monari P, Cocchi D (eds) (1993) Probabilità e induzione. Editrice Clueb, Bologna, pp. 365–374

[11] Gittelson S, Biedermann A, Bozza S, Taroni F (2012) The database search problem: A question of rational decision making. Forensic Sci Int 222:186–199

[12] Gittelson S, Bozza S, Biedermann A, Taroni F (2013) Decision-theoretic reflections on processing a fingermark. Forensic Sci Int 226:e42–e47

[13] Gittelson S (2013) Evolving from inferences to decisions in forensic science. Dissertation, University of Lausanne

[14] Gittelson S, Biedermann A, Bozza S, Taroni F (2014) Decision analysis for the genotype designation in low-template-DNA profiles. Forensic Sci Int Genet 9:118–133

[15] Gittelson S, Steffen CR, Coble MD (2016) Low-template DNA: A single DNA analysis or two replicates? Forensic Sci Int 264:139–145

[16] Gittelson S, Steffen CR, Coble MD (2016) Expected net gain data of low-template DNA analyses. Data in Brief 8:375–386

[17] Gomes C, Lopez-Matayoshi C, Palomo-Diez S, López-Parra AM, Cuesta-Alvaro P, Baeza-Richer C, Gibaja JF, Arroyo-Pardo E (2017) Presumptive tests: A substitute for benzidine in blood samples recognition. Forensic Sci Int Genet 6:e546–e548

[18] De Groot MH (1970) Optimal statistical decisions. John Wiley & Sons, Hoboken

[19] Hochmeister MN, Budowle B, Sparkes R, Rudin O, Gehrig C, Thali M, Schmidt L, Cordier A, Dirnhofer R (1999) Validation studies of an immunochromatographic 1-step test for the forensic identification of human blood. J Forensic Sci 44:597–602

[20] Jensen FV, Nielsen TD (2007) Bayesian networks and decision graphs, 2nd edn. Springer, New York

[21] Johnston E, Ames CE, Dagnall KE, Foster J, Daniel BE (2008) Comparison of presumptive blood test kits including Hexagon OBTI. J Forensic Sci 53:687–689

[22] Kahneman D, Tversky A (1984) Choices, values and frames. American Psychologist 39: 341–350

[23] Kaplan J (1968) Decision theory and the factfinding process. Stanford Law Review 20:1065–1092

[24] Kaye DH (1988) Introduction: What is Bayesianism? In Tillers P, Green ED (eds) Probability and inference in the law of evidence: The uses and limits of Bayesianism. Kluwer Academic Publishers, Dordrecht, pp. 1–19

[25] Kaye DH (1999) Clarifying the burden of persuasion: What Bayesian decision rules do and do not do. Int J Evid Proof 3:1–29

[26] Kjaerulff UB, Madsen AL (2008) Bayesian networks and influence diagrams: A guide to construction and analysis. Springer, New York

[27] Korb KB, Nicholson AE (2011) Bayesian artificial intelligence, 2nd edn. CRC Press, Boca Raton

[28] Krawczak M, Schmidtke JO (1992) The decision theory of paternity disputes: Optimization considerations applied to multilocus DNA fingerprinting. J Forensic Sci 37:1525–1533

[29] Lindley DV (1985) Making decisions, 2nd edn. John Wiley & Sons, Chichester

[30] Lindley DV (2000) The philosophy of statistics. The Statistician 49:293–337

[31] Lindley DV (2006) Understanding uncertainty. John Wiley & Sons, Hoboken

[32] Nease RF, Owens DK (1997) Use of influence diagrams to structure medical decisions. Medical Decision Making 17:263–276

[33] von Neumann J, Morgenstern O (1947) The theory of games and economic behavior, 2nd edn. Princeton University Press, Princeton

[34] Owens DK, Shachter RD, Nease RF (1997) Representation and analysis of medical decision problems with influence diagrams. Medical Decision Making 17:241–262

[35] Phillips VL, Saks MJ, Peterson JL (2001) The application of signal detection theory to decision-making in forensic science. J Forensic Sci 46:294–308

[36] Piva de Almeida J, Glesse N, Bonorino C (2011) Effect of presumptive tests reagents on human blood confirmatory tests and DNA analysis using real time polymerase chain reaction. Forensic Sci Int 206:58–61

[37] Raiffa H, Schlaifer R (1961) Applied statistical decision theory. The M.I.T. Press, Cambridge

[38] Raiffa H (1968) Decision analysis – Introductory lectures on choices under uncertainty. Addison-Wesley, Reading

[39] Savage LJ (1951) Theory of statistical decision. Journal of the American Statistical Association 46:55–67

[40] Savage LJ (1972) The foundations of statistics, 2nd edn. Dover, New York

[41] Shachter RD (1986) Evaluating influence diagrams. Operations Research 34:871–882

[42] Sironi E, Gittelson S, Bozza S, Taroni F (2021) Minor or adult? Introducing decision analysis in forensic age estimation. Science & Justice 61:47–60

[43] Taroni F, Bozza S, Aitken CGG (2005) Decision analysis in forensic science. J Forensic Sci 50:894–905

[44] Taroni F, Bozza S, Bernard M, Champod C (2007) Value of DNA tests: A decision perspective. J Forensic Sci 52:31–39

[45] Taroni F, Bozza S, Biedermann A, Garbolino G, Aitken CGG (2010) Data analysis in forensic science: A Bayesian decision perspective. John Wiley & Sons, Chichester

[46] Taroni F, Biedermann A, Bozza S, Garbolino P, Aitken CGG (2014) Bayesian networks for probabilistic inference and decision analysis in forensic science. John Wiley & Sons, Chichester

[47] Taroni F, Garbolino P, Biedermann A, Aitken CGG, Bozza S (2018) Reconciliation of subjective probabilities and frequencies in forensic science. Law Prob Risk 17:243–262

[48] Taroni, F, Garbolino P, Bozza S (2020) Coherently updating degrees of beliefs: Radical probabilism, the generalization of Bayes' theorem and their consequences on evidence evaluation. Law Prob Risk 19: 293–316

[49] Taylor D, Abarno D, Hicks T, Champod C (2016) Evaluating forensic biology results given source level propositions. Forensic Sci Int Genet 21:54–67

[50] Tillmar AO, Mostad P (2014) Choosing supplementary markers in forensic casework. Forensic Sci Int Genet 13:128–133

[51] Tobe SS, Watson N, Daéid NN (2007) Evaluation of six presumptive tests for blood, their specificity, sensitivity, and effect on high molecular-weight DNA. J Forensic Sci 52:102–109