Forensic Science International 331 (2022) 111174

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Use of Bayesian Networks for the investigation of the nature of biological material in casework

Lydie Samie[a,*], Christophe Champod[b], Séverine Delémont[a], Patrick Basset[a], Tacha Hicks[a,c], Vincent Castella[a]

[a] Forensic Genetics Unit, University Center of Legal Medicine, Lausanne - Geneva, Lausanne University Hospital and University of Lausanne, Ch. de la Vulliette 4, 1000 Lausanne 25, Switzerland
[b] Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice, University of Lausanne, Lausanne-Dorigny, Switzerland
[c] Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice and Fondation pour la formation continue UNIL-EPFL, University of Lausanne, Lausanne-Dorigny, Switzerland

ARTICLE INFO

Article history:
Received 5 October 2021
Received in revised form 14 December 2021
Accepted 1 January 2022
Available online 4 January 2022

Keywords:
Likelihood ratio
Body Fluid detection
OBTI
RSID Saliva
PSA
Christmas Tree staining

ABSTRACT

Chemical and staining methods, immunochromatography, spectroscopy, RNA expression or methylation patterns, do not allow to determine the nature of the biological material with certainty. However, to our knowledge, there are few forensic scientists that assess the value of such test results using a probabilistic approach. This is surprising as it would allow account for false positives and false negatives and avoid misleading conclusions.

In this paper, we developed three Bayesian Networks (BNs) to assess the presence of blood, saliva and sperm in the recovered material and combine potentially contradictory observations. The approach was successfully tested using 188 traces from proficiency tests. We have implemented an online user-friendly application (https://forensic-genetic.shinyapps.io/BodyFluidsApp/) that allows forensic scientists to assess the value of their results without having to build Bayesian Networks themselves. They can also input their own data, use the application to identify a potential lack of knowledge and report their conclusions regarding the presence of sperm, blood or/and saliva considering uncertainty.

1. Introduction

The characterization of the nature of biological fluids (e.g., blood, saliva or sperm) can be important for the investigation. Chemical and staining methods, immunochromatography, spectroscopy, RNA expression or methylation patterns are commonly used by forensic scientists (police and laboratories) to carry out that task. Usually, chemical and immunochromatographic tests are called "presumptive" and very specific tests such as staining methods in sperm morphological analysis are said to be "confirmatory". However, none of these tests is error free. False negatives as well as false positives, including misrecognition of type of cell, are known to occur. We define all these tests as presumptive in this paper. To our knowledge, few forensic scientists account for this when giving the results of their presumptive tests. For example, some forensic scientists report that their results indicate that the material is blood or simply that the result of the blood test is positive. In our laboratory we used to report that the results were better explained in the presence of blood. A layperson might understand that the material is undoubtedly blood. However, forensic scientists are aware that a positive test does not imply that the nature of the trace material has been established with 100% certainty. Forensic scientists may rely on other information such as the location of the recovered trace material, or other results (e.g., DNA quantification, quality of the DNA profile); but they generally do not disclose how they have taken into account the false positives or false negatives, and how they have combined all their results. This lacks transparency and leaves much room for misunderstanding: qualifying the results as providing an "indication" is akin to declare that the findings are "consistent with" one type of biological fluid. This type of conclusion has been shown to be misleading [1].

One of the reasons that could explain why results of presumptive tests are not formally assessed is that there are no specific tools to deal with this aspect. Bayesian Networks (BNs) have been shown to be very valuable for making such inferences (refer to [2] for a review of BNs). BNs allow to combine results that may be contradictory and

* Corresponding author.
E-mail address: Lydie.Samie-Foucart@chuv.ch (L. Samie).

https://doi.org/10.1016/j.forsciint.2022.111174
0379-0738/

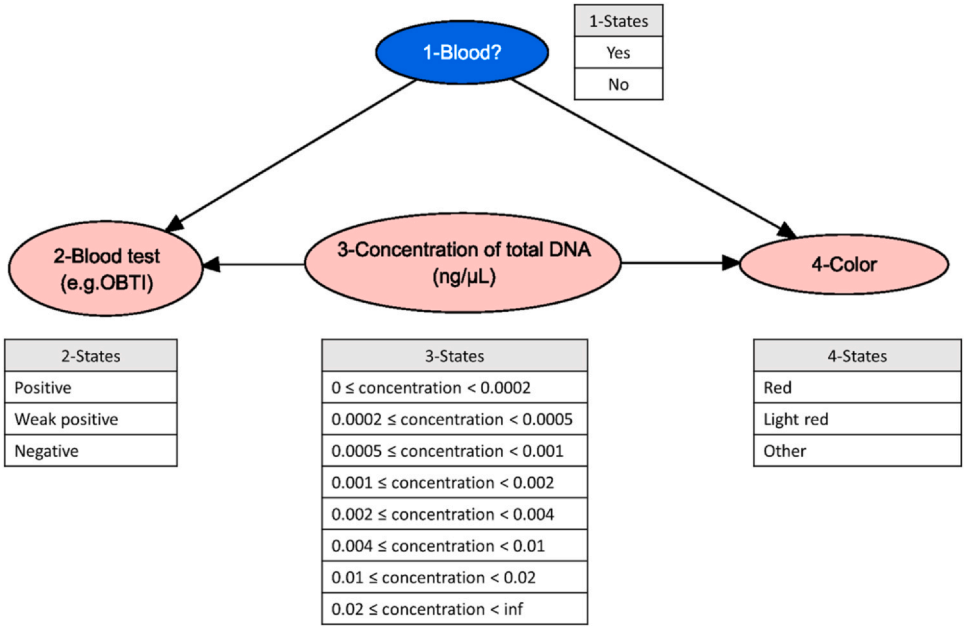

Fig. 1. BN used to investigate whether or not the trace contains blood.

difficult to evaluate without resorting to such a structured approach. Wolff et al. [3] show an example of BNs for saliva tests. Taylor et al. [4] have proposed an example of BNs for blood tests. Taylor [5] has presented an example on sperm recovered in sexual assault. However, for non-specialists, these BN would be difficult to use and/or adapt for tests other than those described in these papers. Furthermore, few forensic scientists use BNs. We believe that possible reasons for this could be: the lack of an easily accessible tool, the difficulty of integrating one's own data into the models and the lack of examples of how one can report such results.

The aim of this paper is as follows:

- To present three examples of BNs that allow to make inferences on the possible presence of respectively blood, saliva and sperm.
- To test these BNs and their parametrization against known ground truth cases.
- To propose a free online tool (i.e., Shiny application) for implementing the approach (https://forensic-genetic.shinyapps.io/BodyFluidsApp/). This user-friendly interface allows to easily integrate users' data within the Bayesian network that works behind the scene.
- To show examples of how one can report results of presumptive tests in a probabilistic way for investigative purposes.

In the Section 2, we define the scope of the article. In the Section 3, we define what we meant by false positives and false negatives results. In the Section 4, we describe the construction of the BNs, based on the protocols and the presumptive tests used in our laboratory (i.e., Hexagon OBTI test from Ruwag, for blood detection, RSID™ Saliva from Independent Forensics, for saliva detection, PSA from Seratec and Christmas Tree for sperm detection [7]). We then present how we have informed the conditional probability tables (CPTs) using domain knowledge derived from published data whenever possible. We conclude this section with the performance of the BNs by applying them to ground truth cases. In the Section 5, we describe and explain how this interface implementing the BNs works. The parametrization of the BNs can be easily adapted should the users want to keep the same BN construction, but use their own data pertaining to these tests. We conclude the section by showing, using sensitivity analysis, how this application helps assessing the impact paucity of data can have on the value of the results.

2. Scope of the proposed approach

In this paper, we focus only on the use of presumptive tests for investigative purposes when there is no person of interest or if the person of interest has not yet been questioned about the events. One might want to have elements regarding the nature of the fluid, for example to take the decision to analyze or not the material. As stated by Gill et al. [6]: "One can use likelihood ratios in both the investigative and the evaluative phase, the main difference is that in the evaluation phase, there will be a suspect/defendant. In this situation it will be necessary to account for the defense's view of events. The forensic scientist operates in 'investigator mode' in the initial stages of a case. A typical example is where a database search is carried out because there is no suspect associated with the crime-scene". Another typical example would be if a small brown fleck is found, whether this is blood and should be DNA profiled or not. Or, to investigate what could have happened to a possible victim of sexual abuse who has no recollection of the events (e.g., if a lot of spermatozoa have been found on internal vaginal swabs one could infer that there possibly was vaginal/penile intercourse). This information will guide the investigation and help ask the pertinent questions, once a person of interest is questioned. We will consider the following propositions:

- The trace contains the biological fluid of interest.
- The trace does not contain the biological fluid of interest (i.e., it contains another unknown material or no material).

The nature of the fluid may be useful as we have seen for the investigation, but also in the evaluative phase when a person of interest has been questioned about the events. As soon as a person is questioned, then evaluative reporting applies, which is not the topic of this paper but is discussed in Section 5.

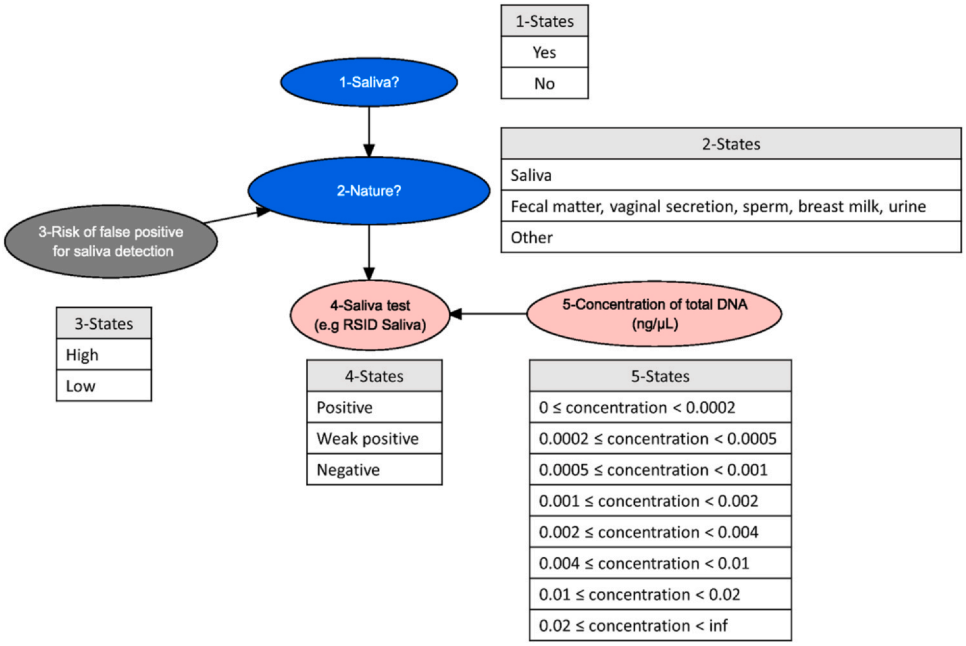

Fig. 2. BN to investigate whether or not the trace contains saliva.

3. Definition of false positives and false negatives

A result is considered as a false positive if the test is positive for a biological fluid different from the target fluid, even if the molecule targeted by the test is the correct one. For example, because sperm may contain low level of α-amylase [8], the RSID Saliva test, targeting the human α-amylase, can be positive in presence of sperm. Not detecting the fluid of interest when present is considered as a false negative.

4. Bayesian network

A BN has two main components: the graphical component that is the structure (variables and dependencies between each variables) and the quantitative component that is the conditional probability tables informed with data coming from experiments (published or not) and knowledge. That is what we referred to earlier as the parametrization of the BN. The BNs were developed using HUGIN researcher version 8.8 (www.hugin.com) [9].

4.1. Construction

For each of the studied biological fluids (blood, saliva and sperm), we have built a BN according to the laboratory protocol, as well as the sensitivity and specificity of each presumptive test. Despite the differences that exist in the structure of these 3 BNs, they have the same aim: providing investigative leads on whether or not the trace contains the biological fluid of interest. This is captured with a proposition node with two possible mutually exclusive and exhaustive states:

- State = Yes: the trace contains the biological fluid of interest.
- State = No: the trace does not contain the biological fluid of interest (i.e., it contains another unknown material or no material).

A node representing the total concentration of DNA (ng/µL) is used in the three BNs. Indeed, the outcome of each presumptive test result depends on the concentration of the target molecule. We have assumed that the measured concentration of DNA is correlated to the concentration of the target molecule. Even if some variability exist both within and between individuals, it seems reasonable to consider that the more biological fluid, the more target molecule and the more DNA.

For the immunochromatographic tests (i.e., Hexagon OBTI, RSID Saliva and PSA), we defined three possible outcomes [10]:

- "Positive" when a colored test line is observed.
- "Weak positive" when a weak colored test line is observed.
- "Negative" when no test line is observed.

The construction of each BN is presented in the following subsections.

4.2. Blood

Fig. 1 shows the BN developed for blood and the associated states for each variable.

The result of the presumptive test and the observation of the received material are respectively described in Node 2 (Blood test) and 4 (Color). We use the color of the swab as the observation rather than the color of the stain, as generally only swabs are transmitted to the laboratory. These observations depend on the nature of the trace (Node 1) and the concentration of DNA (Node 3).

4.3. Saliva

Fig. 2 shows the BN developed for saliva and the associated states for each variable.

The result of the presumptive test is represented by Node 4 (Saliva test). The RSID Saliva test outcome depends both on the nature of the fluid (Node 2) and on the DNA concentration (Node 5). Positive reactions for RSID Saliva have been observed with fecal matter, rat saliva [11,12], breast milk [12], urine, sperm, sweat [11,13] and vaginal secretions [11,13,14]. Node 3 allows to take into account false positives according to the type of item on which the sample was collected. For example, Node 3 is set to high if the item is underwear. A further example would be if the item is a bra and that the owner is breast feeding.

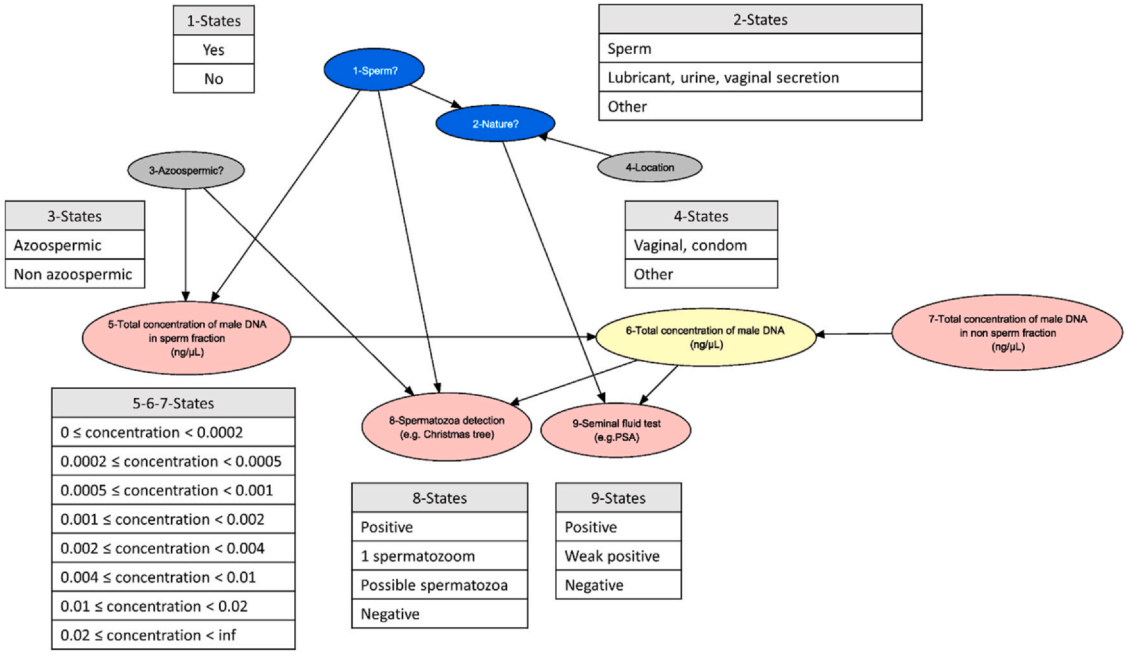

Fig. 3. BN to investigate whether or not the trace contains sperm.

4.4. Sperm

Fig. 3 shows the BN used for sperm and the associated states to each variable.

Nodes 8 (Spermatozoa detection) and 9 (Seminal fluid test) represent the results of the examination. We use Christmas Tree staining and PSA [10] to investigate respectively the presence of spermatozoa and seminal fluid.

For Christmas Tree, the result is defined as follows:

- "Positive" if more than one spermatozoon are observed.
- "1 spermatozoon" if only one spermatozoon is observed.
- "Possible spermatozoa" if the forensic scientist is uncertain that the cells are spermatozoa.
- "Negative" if no spermatozoon, actual or possible, is observed.

According to our protocol, differential DNA extraction is carried out with the Erase Sperm isolation kit (PTC Laboratories, Columbia, USA) when the presence of sperm is suspected. Following this differential extraction of DNA, two fractions are quantified: the so-called 'sperm fraction' and 'non-sperm fraction'. The concentration of male DNA (Node 6) is the sum of the concentration of male DNA in the 'sperm fraction' (Node 5) and in the 'non-sperm fraction' (Node 7). We used the entire specimen to carry out PSA, Christmas Tree and to determine the concentration of male DNA. Therefore, the results of Christmas Tree and PSA depend on the concentration of male DNA (Node 6). These observations also depend on the nature of the material (Node 2).

False positive for PSA have been observed with some contraceptive foams, female urine and breast milk [15], condom lubricants containing nonoxynol-9 [16], male urine [15,17,18] and vaginal secretions [14,15,19]. Node 4 is used to define whether the risk of false positive is high or low depending on the location.

With differential extraction of DNA, we expect the male DNA in the "sperm fraction" to come from spermatozoa. However, if a person is azoospermic, we expect to observe no or a small quantity of DNA in this fraction. Hence, the concentration of male DNA in Node 5 depends on the potential presence of sperm (Node 1) and whether or not the person is azoospermic (Node 3). The detection of spermatozoa (Node 8) obviously also depends on Node 3.

4.5. Conditional probability tables (CPTs) parametrization

In this sub-section, we present the CPTs for each node of the BNs.

4.5.1. Node 1 (Blood/Saliva/Sperm) of the 3 BNs

In general, forensic DNA scientists are given little information on the case circumstances during the investigation phase. By default, we have chosen to assign the same prior probability that the trace contains or not the biological fluid of interest (see Section 6). Note that this prior probability can be adjusted as required.

4.5.2. Node 2 (Nature) of saliva and sperm BNs

When considering the probability of a positive result given the alternative proposition (i.e. in the absence of the specific biological fluid), we typically consider the possibility of false positives. This probability depends on where the sample was collected.

The probabilities assigned for Node 2 are given in Tables 1 and 2 (BNs for saliva and sperm respectively). These probabilities were assigned according to our expectations and are justified as follows:

- If there is saliva or sperm (state: Yes), obviously the probability of saliva or sperm is 1.
- If there is no saliva (state: No) and the location (e.g., a vaginal swab) suggests a high risk of false positives, we have assigned a probability of 0.99 for the presence of material known to lead to false positives (e.g., vaginal secretions) and a probability of 0.01

Table 1
Conditional probability table of Node 2 (Nature of "other" than saliva) given nodes 1 and 3.

Node 1-Saliva?	Yes		No	
Node 3-Risk?	High	Low	High	Low
Saliva	1	1	0	0
Fecal matter, vaginal secretions, sperm, breast milk, urine	0	0	0.99	0.01
Other	0	0	0.01	0.99

Table 2
Conditional probability table of Node 2 (Nature of "other" than sperm) given nodes 1 and 4.

Node 1-Sperm?	Yes		No	
Node 4-Localization?	Vaginal, condom	Another	Vaginal, condom	Another
Sperm	1	1	0	0
Lubricant, urine, vaginal secretion	0	0	0.99	0.2
Other	0	0	0.01	0.8

for unknown material. Conversely, if the location (e.g., glass) suggests a low risk of false positives, we have assigned these probabilities as 0.01 and 0.99 respectively.

- If there is no sperm (state: No) and the location of the trace (e.g., vaginal swab or condom) suggests a high risk of false positives, we have assigned a probability of 0.99 for the presence of material known to lead to false positives (e.g., sexual lubricant, urine or vaginal secretion). Conversely, if the location (e.g., jacket) suggests a low risk of false positives, we have assigned a probability of 0.2 for the presence of sexual lubricant, urine or vaginal secretion and a probability of 0.8 for the presence of an unknown material.

For blood, it is known that a positive OBTI test can be triggered by urine, saliva, sperm, vaginal secretions and the high pH of the recovered material [20]. In our BN, this possibility was handled directly in the node referring to the OBTI test (Node 2).

4.5.3. Node 3 (Risk of false positive) of saliva BN and Node 4 (location) of sperm BN

When no information is available, the CPT for the Nodes 3 (Risk of false positive) of the BN for saliva and 4 (Location) of the BN for sperm are based on uniform probabilities. Once information is available, these nodes should be instantiated (i.e. the state is chosen by the user).

4.5.4. Node 3 (Azoospermic) of sperm BN

In this paper, sperm is defined as a seminal fluid with or without spermatozoa. The result of the detection of spermatozoa depends on the biological nature of the trace but also on whether the person at its source is azoospermic or not. This dependency is captured by Node 3. In investigative mode, we do not have a person of interest so we focus on the probability that a person in the population of interest would be azoospermic. This could be adapted if some information concerning the donor is available. The percentage of azoospermic men varies depending on the age of the person [21]. The probability of being vasectomized depends on the country and increases with age [22]. For our parametrization, we have assigned a default probability of 0.1 to be azoospermic based on Swiss data [23,24]. If required this probability can be adapted.

4.5.5. Concentration nodes

The alleged activities that led to the deposition of the biological material are usually not known at the investigation stage. Therefore, we are not in a position to know what are the expected DNA concentrations according to the type of material. For example, we expect to recover more semen (and therefore more DNA) from a panty following sexual intercourse with ejaculation rather than when being washed with another garment where sperm would be present. The lack of knowledge regarding the alleged activities [25] precludes inferring the expected concentration of DNA.

With the BNs proposed in this paper, the expected DNA concentration related to the nature of the trace is unknown. However, some considerations need to be taken into account. For example, the maximum DNA concentration that can be associated with the fluid cannot be higher than the total concentration of DNA obtained. Besides, the DNA profile also gives information on how to fill the CPTs for the DNA concentration. For example, if a mixed DNA profile is obtained, it is not possible to infer who is(are) the contributor(s) to the biological fluid and to instantiate the DNA concentration associated to these contributor(s). The impact of the DNA profile result on the DNA concentration nodes (Node 3 for Blood BN, Node 5 for Saliva BN, Nodes 5 and 7 for sperm BN) are described in Table 3.

These considerations have not been taken into account in the BN built in Hugin. The CPTs of the DNA concentration nodes need to be informed manually in each case. However, this is automatically done in the Shiny application, described in the corresponding section, when informing the results about the DNA profile and the total concentration of DNA.

There are several important aspects regarding the CPTs for the concentration of total male DNA in the 'sperm fraction' (Node 5). First, we expect that the DNA in the "sperm fraction" comes from the spermatozoa since a differential DNA extraction is performed for specimens potentially containing sperm. However, if the person at the source of the sperm is azoospermic, we expect to observe no or only a few cells in the "sperm fraction". As described earlier, the concentration of male DNA in the "sperm fraction" (Node 5) depends on whether the person is azoospermic or not (Node 3). The CPT for node 5 was informed using data from an internal non-published experiment and considering that an unobserved outcome is not impossible (Table 4).

4.5.6. Nodes pertaining to the observations

For the investigation of blood, the observations are instantiated in Node 2 (Blood test) and Node 4 (Color) respectively. For saliva, this is done in Node 4 (Saliva test) and for sperm in Nodes 8 (Spermatozoa detection) and 9 (Seminal fluid test), respectively. To

Table 3
Definition of each type of DNA profile that could be observed with the repercussions associated to this observation.

Observed DNA profile	Definition	Repercussions for the conditional probability tables of concentration nodes
NA	The DNA analysis was not performed.	When the observed DNA concentration is in interval x, the probabilities are uniformly distributed between 0 and x. Probabilities of zero are assigned for intervals above x.
None	0–5 alleles on the electropherogram (EPG).	Instantiation of the observed DNA concentration.
Not interpretable	DNA replicates are too variable or below the analytical threshold.	When the observed DNA concentration is in interval x, the probabilities are uniformly distributed between 0 and x. Probabilities of zero are assigned for intervals above x.
Single source	DNA profile of one person or a mixture with a major DNA component from one person and a minor component with low peak heights (i.e. < about 10% of the major peaks)	Instantiation of the observed DNA concentration.
Mixture	Any result different from the 4 described above.	When the observed DNA concentration is in interval x, the probabilities are uniformly distributed between 0 and x. Probabilities of zero are assigned for intervals above x.

Table 4
Conditional probability table used for node 5 (Concentration of total male DNA in ng/µL in the sperm fraction) given nodes 1 and 3.

Node 1-Sperm?	Yes		No	
Node 3-Azoospermic	Azoospermic	Non Azoospermic	Azoospermic	Non Azoospermic
[0–0.0002]	0.5	0.125	0.53	0.53
[0.0002–0.0005]	0.4	0.125	0.3	0.3
[0.0005–0.001]	0.05	0.125	0.06	0.06
[0.001–0.002]	0.01	0.125	0.05	0.05
[0.002–0.004]	0.01	0.125	0.03	0.03
[0.004–0.01]	0.01	0.125	0.01	0.01
[0.01–0.02]	0.01	0.125	0.01	0.01
[0.02–inf	0.01	0.125	0.01	0.01

The total DNA concentration in the BN for sperm is the sum of the DNA concentration observed in both 'sperm' and 'non sperm' fractions.

inform these variables, we have adopted a full Bayesian strategy meaning that we have initially set prior counts for each state based on our prior knowledge. Then, we have updated these prior counts using observed counts from published experiments to get posterior counts. Finally, these posterior counts were used to assign probabilities for each states to inform the appropriate CPTs. An example of the process is described below to explain how posterior probabilities of getting a positive, weak positive or negative OBTI result were assigned when the trace contains blood with a DNA concentration higher than 0.02 ng/µL.

To assign prior counts (before performing our experiments), we have considered that even when blood is present in the trace, one can observe a weak positive or a negative test result. The occurrence of either event was set as 1/10. Observed counts are based on experiments following the protocol described in [10]. In [10], 24/24 experiments gave a positive result for the situation when blood is present with a DNA concentration higher than 0.02 ng/µL. The posterior counts are then determined by the sum of prior and observed counts and used to assign the posterior probabilities as illustrated in Table 5.

The same reasoning was used for each state of the different variables about observations in the three BNs (blood, saliva and sperm). Supplementary data allowing to inform the BN of blood were obtained following protocols described in [10] and are presented in Appendix A. The full data allowing to inform the three BNs are presented in Appendix B. Prior and observed counts, as well as the probabilities associated to each state of these variables can be found in the "data" tab of the Shiny application described in the next section. This application also allows the users to easily integrate their own data (prior and observed counts) in the BNs.

4.6. Ground truth experiments: putting our BNs to the test

We compared the output of our BNs to ground truth experiments using data from 20 proficiency tests, organized by GEDNAP,[1] the SSML,[2] and the ISFG[3] French speaking working Group. A total of 68 traces were tested for blood, 58 for saliva and 61 for semen; in some cases mixtures of body fluids were present.

Table 6 describes, for each biological fluid investigated, the observations and the results that were obtained using the BNs. More detailed results are available in Appendix C. For transparency (if one is interested in reproducing the calculations), we have reported the LR and the posterior probability displayed by the Bayesian Network. However, in our reports, we propose to round the LRs (See Discussion, subsection 1).

The comparison of the results obtained using the BNs with ground truth results shows:

Table 5
Example of how posterior probabilities of getting a positive, weak positive or negative OBTI result were assigned if the trace contains blood with a DNA concentration higher than 0.02 ng/µL.

States [0.02–inf	Prior counts	Observed counts	Posterior counts	Probability
Positive	8	24	32	0.94
Weak positive	1	0	1	0.03
Negative	1	0	1	0.03
Total	10	24	34	1

- When the material is not human blood, the conclusion agrees with ground truth: even for traces containing animal blood or orange juice which are known to be problematic and even for traces which give a weak positive OBTI result. When the material is blood, the results obtained also agree with ground truth even with blood dilutions of 1/320. There is only one trace out of 68, where observations (DNA mixture profile, concentration higher than 0.02 ng/µL, white swab and negative OBTI test result) do not support ground truth. Results are 3 times more likely if the trace does not contain blood, when in fact it was blood diluted to 1/1000. It should be noted that even without the use of BN, because of the negative OBTI test, we would have concluded that the observations are best explained in the absence of human blood.
- For saliva, all probabilistic results agree with the nature of the 58 traces.
- For sperm, the probabilistic results also support ground truth, even for traces containing sperm without spermatozoa. There is one specimen out of 61, where observations (Unknown location, positive PSA, no spermatozoa detected, no male DNA profile in the "sperm fraction" with concentration of 0.0001 ng/µL, single male DNA profile in the "non-sperm fraction" with a concentration higher than 0.02 ng/µL) are equally probable if the trace contains or not sperm, when in fact we expected sperm from a vasectomized male. The specimen was a vaginal swab recovered 6 h after an intercourse with a vasectomized partner. The effect of knowing that the person of interest is vasectomized is shown in Table 7. This illustrates the importance of performing a new evaluation if the circumstances of the case change or if new information is made available and show that the use of BNs concur with our expectations.

5. Shiny application

For implementing these BNs in casework, it is helpful to have a user-friendly interface. In the next section, we describe this interface that takes the form of a Shiny application.

There are software applications that allow to construct and run BNs. We can cite for example the commercial application Hugin (www.hugin.com) or the open-source software GeNie (https://www.bayesfusion.com/genie/). However, they require specific technical skills that are not always available to the forensic scientists. The proposed Shiny application aims at providing a user-friendly

[1] GEDNAP: German DNA profiling (https://www.gednap.org/).
[2] SSML: Swiss Society of Legal Medicine (https://www.sgrm.ch/fr/ssml-home/).
[3] ISFG: International Society for Forensic Genetics (https://www.isfg.org/).

Table 6
Observations and associated probabilistic results (LR and posterior probability) obtained for known traces from proficiency tests in which the presence of blood, saliva or sperm was investigated. We assigned prior odds of 1:1. Observations that did not support the ground truth are outlined in grey.

Nature of the bodyfluid	Results of the tests	Number of traces	LR	BN result P(H|E,I)	Outcome
Human blood	OBTI +/Light red or Other colour or N/A	49	9-134	0.90-0.99	Expected
	OBTI -/ Other colour or N/A	1	0.28	0.22	Unexpected
No human blood	OBTI -/Red	4	0.27-0.90	0.21-0.47	Expected
	OBTI -/ Other colour or N/A	12	0.01-0.28	0.01-0.22	Expected
	OBTI weak +	2	0.27 and 0.35	0.21 and 0.26	Expected
Saliva	RSID +	14	11-16	0.92-0.94	Expected
No saliva	RSID -	44	0.03-1	0.03-0.50	Expected
Sperm	PSA +	20	1-Infinite	0.50-1	Expected
	PSA +/CT-	1	0.62	0.38	Unexpected
No Sperm	PSA-/CT- or N/A	40	0.05-0.2	0.05-0.18	Expected

interface that calls upon the required BNs and allows to update the network either by adding new data to parametrize the CPTs or conduct propagation based on chosen instantiated states. The application also allows to export the results of an assessment in a detailed report that can be added to case notes. The main functionalities of this Shiny application are described below. The web application is deployed on shinyapps.io: https://forensic-genetic.shinyapps.io/BodyFluidsApp/.

The development was carried out in R version 4.0.2 (2020-06-22) [26] coupled with RStudio Version 1.4.1103 [27] using the following key packages: For the user interface: shiny [28], shinyjs [29] and rintrojs [30]. For the BNs engine, we relied upon the graphical model engines from gRbase [31], gRain [32] and bnlearn [33].

The three BN presented in this paper can be downloaded from the web application.

5.1. Relations between the application and the BN

The BN structure and the conditional probability tables presented in this paper were integrated in R using the bnlearn package. Results of computations were successfully validated against the original BNs developed in Hugin (version researcher 8.8).

In the interface, the user can select different observations: e.g., the result of the presumptive test, the color of the sample, the location, and the risk of false positive. When the observations are selected, they are instantiated in the BN (i.e., a probability of 1 is assigned to the states corresponding to the selected observations).

Table 7
Effect of the probability of being azoospermic on the LR and on the probability that the trace contains sperm knowing the observations (Unknown location, positive PSA, no spermatozoa detected, no male DNA profile in the "sperm fraction" with concentration of 0.0001 ng/µL, single male DNA profile in the "non-sperm fraction" with a concentration higher than 0.02 ng/µL) considering a prior probability of 0.5.

Probability of being azoospermic	LR	P (Sperm = Observations,I)
0.1	0.6	0.38
0.5	2.8	0.73
1	5.6	0.84

However, in casework, some observations may not be known by the user. For example:

- The color of the swab may not have been recorded by the laboratory or cannot be recorded (for example, if the sampled area is dirty potentially masking the red color of the blood).
- For sperm detection, results may be available only for microscopy or for the immunochromatographic test.
- The location where the sample was collected may be unknown.

In these cases, "N/A" (not applicable) is selected from the drop-down menu meaning that no state of the variable is instantiated.

Once the observations are set, the application computes the posterior probabilities for the node of interest.

5.2. Integration of other background data

This application allows forensic scientists to import their own data and automatically update the conditional probability tables for each biological fluid. This allows other forensic scientists to use the application with the same BN construction, but with their own data associated with these tests, also in case the DNA extract volume is different from 50 µL.

If a laboratory applies different presumptive tests, the Shiny application can still be used once the appropriate data have been imported. However, the presumptive tests must be consistent with the BN presented in this paper (e.g., do the results of the test depend on another variable?). If not, the BN can be adapted externally and/or further tests can be added. If the users choose another model and modify the BN proposed in this paper, then they cannot use the Shiny application.

5.3. Managing the paucity of data

The Shiny application presents the result of the evaluation of the selected observations numerically and graphically. It is also possible to investigate the impact of the data informing the BN in order to highlight a potential lack of knowledge using sensitivity analysis. As

in [34], the sensitivity of the posterior probability that the trace contains the biological fluid of interest to the data underlying the observations can be explored. At each simulation, the counts from a Dirichlet distribution are re-sampled, for one or several nodes simultaneously simulated, using software R [35] and the freely available R libraries, gRain [32] and BNlearn [33]. The counts that are resampled, are made up of the prior counts and the observed counts. A total of 100 simulations are performed. A Likelihood Ratio (LR) is obtained for each simulation and the posterior probability is calculated from this LR and the selected prior probabilities in the Shiny application. The posterior probabilities obtained for each set of observations after the simulations, can be visualized as a boxplot in the Shiny application. For our laboratory, we have arbitrary defined that the variation of the posterior probability is high when the ratio interquartile to the median is higher than 0.2. A high variation means that there is potentially a lack of data for a set of observations. In this situation, the decision to acquire more data is based on a cost/benefit balance: for example, is this set of observations often obtained and is it worth investing in the acquisition of additional data?

Using the Shiny application, one can explore what would be the effect of carrying out more experiments (e.g., if 20 times more experiments for each set of observations were made, what would be the expected results?). These results are simulated using the same method and allow to pre-assess whether more experiments are needed or not.

These simulations are obviously only useful to highlight a potential lack of data if these are relevant in the case. Indeed, if poor data are used, the results of the simulations and the probabilistic results obtained from the BN is not meaningful.

5.4. Implementation in casework

The results obtained using the BNs were compared to ground truth experiments using proficiency tests. The results were judged very satisfactory. As the probabilistic results obtained with the Shiny application are the ones obtained from the BN, we are of the opinion that the Shiny application can be used in casework.

To facilitate implementation in casework, the Shiny application offers the possibility to generate a report containing all relevant information.

The Shiny application displays a message in the presence of unexpected results. This alerts the users they should investigate if there has been an input or a laboratory error, or if the presence of material from an animal is possible.

6. Discussion

Forensic scientists need to account for the possibility of false positives and false negatives, when reporting their findings regarding the nature of a biological material. To do so, assigning the value of the results within a Bayesian framework is recommended. It allows to convey the meaning of a positive or negative test in a balanced and logical way. One should not view presumptive tests as factual and leave the interpretation to a layperson. For example, if an OBTI test is positive, just giving the raw result leads the reader of the report to conclude that there is blood. But, depending on the color of the swab, the concentration of the DNA and the quality of the DNA profile, the probability that the trace contains blood may "only" be 0.85 and not “1” (Table 6). Because, presumptive tests do not lead to factual results, they need to be assessed probabilistically and the conclusion not presented as a fact. It is not sufficient to say that the results support or indicate that the material is blood, one needs to quantify the level of support. None of these tests, whether presented as “presumptive” or “confirmative”, is error free. Thus, categorical decisions regarding the nature of a body fluid should not be made by forensic scientists, except in the presence of a large quantity of sperm heads. Even with this exception, forensic scientists should keep in mind that it could be animal sperm.

To assess any result, we first need to define what the results are. We then need structured data to assign our probabilities. To determine the nature of the trace material, all observations should be used. BNs are a useful tool allowing to make inferences about the nature of the material based on structured data and on the combination of multiple observations. It is not uncommon to have conflicting results (i.e., a PSA positive and a Christmas Tree negative); in this case it is more difficult to assess the observations without the use of BNs. The Shiny application offers a user-friendly interface to operate the three BNs presented in this paper, making it possible to use this tool without having specific knowledge in the construction of BNs.

Below, we discuss the question on how one can report this probabilistic result, as well as the advantages and limitations of both BNs and the Shiny application.

6.1. Reporting the value of the results when the investigative issue regards the nature of the fluid

Some forensic scientists report that their results indicate that the material is blood or simply that the result of the blood test is positive. In our laboratory we used to report that the results were better explained in the presence of blood. We are of the opinion that this would be understood as a decision on the nature of the biological fluid. This is not the responsibility of forensic scientists, as they may be unaware of the costs associated with such a decision [36]. Forensic scientists know that false positives exist and that having a positive test does not necessarily imply that the target biological fluid is present. Forensic scientists should assess their results obtained in logical, transparent and robust way, using all the observations and available information. This paper shows how BNs and our Shiny application can help doing so. In casework it is rare to test one trace for the presence of several fluids, however should this be the case, each BN can be used separately.

In this section, we discuss the possibility of reporting a posterior probability, which is the probability that the trace contains the biological fluid of interest given the observations (i.e., results of presumptive tests, color of the swab, DNA concentration, DNA profile). However, to assign a posterior probability one needs to assign a prior probability. Maskell and Jackson [37] proposed to use a Bayesian approach to evaluate the results of presumptive drug test by assessing the posterior probability of a drug being present.

At the investigative stage, defining if the trace may contain blood (or any other biological material) or not, prior to any tests, based on only the circumstances, could be considered as the remit of the forensic DNA scientist. Indeed, the forensic scientists are given the necessary information by the investigators. Then, if new information is provided for example by the defense, one needs to assess the results (this time including the DNA profile of the person of interest) considering activity level propositions [38].

If the posterior probability is given, then three questions arise:

- Should our LR also be given? Since LRs are required to compute posterior probabilities, we are of the opinion that they should accompany prior probabilities.
- Should the posterior probabilities of both propositions be given? Indeed, indicating that the probability that the trace contains blood is 20% may be perceived differently than indicating that probability that the trace does not contain blood is 80%. If only one posterior probability is to be given, we recommend presenting the posterior probability referring to the proposition that is supported by the observations. For example if an LR larger than 1 is obtained, the posterior probability that the trace contains the biological fluid should be given. On the contrary, if an LR less than

1 is obtained, the posterior probability that the trace does not contain blood should be given.

- Which prior probability should we assign? In casework, we generally do not have any information on the case. But, since the police request a specific test, we assume that it is at least as likely that the trace contains the biological fluid of interest as if it did not. In [37], the authors claims that in the complete absence of any information, then an uninformative prior probability of 0.5, reflecting maximum uncertainty, could be adopted. Note that a prior probability of 0.5 is also used in the context of paternity cases with the Essen-Möller approach [39]. One assumes that the prior probability that the alleged father is the true biological father is the same as the prior probability of an unknown man. It might nevertheless be important to adapt this prior to take into account other information such as the result of a presumptive test performed before the trace was sent to the DNA laboratory. This can be done in the Shiny application which allows to cover prior probabilities from 0 to 1.

In our reports, we round our LRs to one significant figure (e.g. 4.5 - > 4, 134 - > 100) and present the posterior probability obtained based on the rounded LR. We do not report LRs smaller than 1 but reverse the propositions, and then round the LR (e.g. 0.7 - > 1.43 - > 1, 0.07 - > 14.28 - > 10). The probabilities are given as percentages truncated to the lower integer (e.g. 0.696 - > 69%).

In the box entitled "Meaning of the results" of the Shiny application, we propose examples of what could be indicated in a statement (i.e., our prior probability, LR and both posterior probabilities). The following situations are covered, taking as an example the BN for blood:

- LR > 1
  If an LR of 200 for example is obtained, we would report: "We have assigned a likelihood ratio of the order of 200. This means that our observations are in the order of 200 times more likely if the trace contains human blood than if it does not. In order to assign the probability that the trace contains human blood, one needs first to assign the probability that it contains blood, before performing the analyzes. This probability is known as the prior probability. Based on our likelihood ratio and a prior probability of 50%, the posterior probability that the trace contains human blood given our observations is 99%.

- LR = 1
  If an LR of 1 is obtained, we would report: "We have assigned a likelihood ratio in the order of 1. This means that our analytical results do not allow to discriminate the proposition that the trace contains human blood from the alternative proposition that it does not. As such they are uninformative. In order to assign the probability that the trace contains human blood, one needs first to assign the probability that it contains blood, before performing the analyzes. This probability is known as the prior probability. Based on our likelihood ratio and a prior probability of 50%, the posterior probabilities that the trace contains detectable human blood and that the trace does not, knowing our observations, remains unchanged, that is 50%."

- LR < 1
  We have assigned a likelihood ratio in the order of 0.5. As LRs smaller than one are difficult to grasp, we reversed the propositions. The LR is reversed as well. If we do so in this case, our LR is two. Said otherwise, our analytical results are twice more probable if the trace **does not contain human blood, rather than if it does. In order to assign the probability that the trace does not contain human blood, one needs first to assign the probability that it does not contain blood, before performing the analyzes. This probability is known as the prior probability. Based on a likelihood ratio of 2 and a prior probability of 50%, the posterior probability that the trace does not contain human blood given our observations is 66%.**

6.2. BNs/Shiny application presented in the literature

Wolff et al. [3], Taylor et al. [4] and Taylor [5] present BNs focused on a given biological fluid. However, for non-specialists, they can be difficult to use and adapt to other trace material. In this paper, we have built 3 BNs, one for each of the most common biological fluids. We have created one Shiny application to easily use all three without dealing with the implementation of the data and the instantiation of the states. This application facilitates the use and implementation of Bayesian networks.

On the other hand, any test or information involving a modification of the BN (e.g., addition of variables or modification of states) cannot be taken into account in the Shiny application. In this situation, forensic scientists can adapt the BNs by adding the necessary variables (e.g., any other examination).

6.3. Putting our BNs to the test

In this paper we have studied the performance of our BNs by using them to assess results where ground truth was known (traces from proficiency tests). Regarding the investigation of blood, there is 1/68 trace where we have concluded that the observations are more likely if the material was not human blood, when in fact it was blood with a dilution of 1/1000. Please note that, without the use of BN, because of the negative OBTI test, we would have also concluded that the observations were best explained in the absence of human blood.

Regarding the investigation of sperm, there was one case with a vaginal swab recovered 6 h after an intercourse involving a vasectomized partner.

In that case, our LR was of the order of 1 considering a probability of 0.1 that the person was vasectomized. However, knowing that the person of interest is vasectomized, a probability of 1 should be used instead. Considering this, our LR would be 5 and the probability that the trace contains sperm knowing our observations 84%.

The results of this study show that the use of BN produces results within our expectations. This validation allows to highlight the advantages of BNs. For instance, for one trace of a proficiency test, we observed no DNA profile, a white color swab and a weak positive OBTI test. With our previous protocol, we would have concluded that we could not determine whether the trace contained blood or not. Using all the information and the BN, our LR is 0.5 and the probability that the trace does not contain human blood knowing our observations is 66%. This example highlights the fact that BNs allow to combine several observations better than intuitive reasoning. Reporting the value of the results within a Bayesian framework allows to make a better use of the information content of the results.

Sometimes, observations may be missing. In this case, they are indicated as "N/A" both in the Shiny application and in Table 6. For an unknown DNA concentration, as it is not possible to select "N/A" in the Shiny application, we selected a concentration higher than 0.02 instead. This choice is motivated by the fact that for these traces a DNA profile was obtained or was expected. For sperm, according to our protocol, we usually perform differential DNA extraction in casework. However, for proficiency test, it was not always required. When only the total concentration of DNA was known (and unknown for each "sperm" and "non-sperm fraction"), the Shiny application could not be used. In the BN however, the node "Total concentration of male DNA" can be instantiated.

It is important to note that the protocols used in our laboratory have been modified since some of the proficiency tests. But the tests

(OBTI, PSA, RSID Saliva and Christmas Tree) are the same. Despite these differences, the LRs obtained using the 3 BNs of this article support the ground truth proposition. This was expected since we used data from our current validation, as well as data from the literature obtained with different protocols. Should a laboratory change protocols, one could use data from previous protocol as prior counts and update these with the new data.

6.4. BNs/Shiny application covering most of the cases

Our aim was to build BNs that would cover most of casework situations to avoid the need to adjust the BN. Thus, there are a few cases we decided to ignore. We have previously indicated that, for some of these tests (OBTI and RSID), fluids from animals can result in false positives. We have made a deliberate choice not to add a variable indicating that the fluid may be human, animal or non-biological, since it is - to our knowledge - a relatively rare situation. However, especially for blood, if a red swab is obtained and the OBTI test is negative, the material can contains animal blood. Thus, a specific alert indicating "Animal blood?" is displayed. These alerts should prompt the user to ask for more information or to perform further analysis or/and verify the results.

6.5. Perspective: nature of the fluid when a person of interest has been questioned

The nature of the recovered material can be important both for the investigation and the evaluative phases. In both cases, one can use a Bayesian framework [40,41] to draw scientifically supported conclusions. This article regards the investigation phase, when we are asked about the presence of a specific biological fluid. At this point, the police is investigating what could have happened and who could be involved. Once a person of interest has been identified, s/he might give information about the activities that might imply the presence of the material (e.g., blood) for legitimate reasons then evaluative reporting applies.

In the hierarchy of propositions for evaluative reporting (not investigative), there is a level that is designated as source. This type of propositions focuses on the issue of whether a given person is the origin of a specific biological material. A typical example would be: “The blood is from Mr Smith” or “The blood is from an unknown person”. As one can see both propositions state that the material is blood. Thus there is no contention that the material is blood. When the hierarchy of propositions was proposed for the first time by Cook et al. [42], it was not possible to produce a DNA profile from small quantities of material. The relationship between the nature (blood, semen, saliva) and the DNA profile was straightforward. With the present sensitivity of DNA analysis, this is generally no more the case. There are three possible approaches if the nature of the fluid is in question and a person of interest is questioned about the events (i.e., in the evaluative stage).

A possible approach could be to process in two stages: (1) Assess the nature of the fluid, say blood, with propositions “The trace contains blood or not” (2) Consider the value of the DNA comparison with propositions “DNA is from the person of interest or from an unknown person”. So, said otherwise, we guide as to the nature of the fluid, and subsequently consider source level propositions. This is a possible solution, but according to the ENFSI guideline [25] for evaluative reporting, there are two conditions for source level propositions to be meaningful: first, the issue should be whether a given person (or object) is the source of the material. Second, there should be no risk for the court to misinterpret the findings in the context of the alleged activities. This would be typically the case only when the material is found in such a quantity that (i) there is no need to consider its presence for reasons other than the alleged activity, (ii) the nature of the material can be safely assumed or (iii) the nature of the fluid is not contested. In our opinion these cases are rare.

A second approach, still in the context of evaluative reporting, is to use a proposition that both considers the nature and the source of the material. This can be quite a difficult task. These propositions are sometimes called source propositions but are in fact nature propositions. An example could be as in Taylor et al. [4]: “The source of the stain on the suspect’s top is the blood of the victim” and “The source of the DNA on the suspect’s top is the victim’s saliva and the stain is not human blood.” In such a case, the question seems to be only the nature of the material. In Zoete et al. [43] they consider multiple propositions source-nature: “The suspect contributed semen” or “The suspect did not contribute semen but contributed another cellular material” or “The suspect did not contribute to the trace”. But, if found in small quantities, then the two conditions for nature or nature-source propositions are not met.

A third approach is to consider activity level propositions once we know what the person of interest says about the events and that the nature of the fluid (and thus the activities) is contested. This is in line with the ENFSI [25] and ISFG guidelines [6,38] for evaluative reporting as value will be added in the process. This is in our opinion the most meaningful as it allows to assess all results (absence and presence of a fluid), and to consider important factors such as transfer, persistence, contamination, prevalence or so-called background. This can only be done when a person of interest has been questioned, so that one can also assess the value of the results considering both his/her version of events and the disputed activity.

There are situations where the only question asked to the forensic scientist concerns the nature of the biological material. We are concerned that non-forensic scientists would think: there is semen, it ‘matches’ the person of interest, thus it is his semen which shows there was sexual intercourse. As already mentioned, this can be misleading, especially if the material is found in small quantity.

The importance of the presumptive tests for evaluative purposes will be the topic of another paper. Here, the results of the BNs relate to the nature of the trace material when tested in the absence of a person of interest (POI) for investigative purpose. For example, if a trace is taken from a t-shirt and observations are more likely if the trace does not contain blood, this approach does not take into account the fact that the t-shirt may have been washed. It is not possible, with these BNs, to give any indications on the presence of blood on the t-shirt at the time of the alleged facts. It is not possible either to help associate the DNA profile or the concentration of DNA with the biological fluid. To contribute to answer these questions (presence/absence of a biological fluid coming from a person of interest at the time of an alleged activity), if there is a risk of misinterpretation, one shall assess the results given activity level propositions and specific case information, especially from the person of interest [25]. We consider that – except when the quantity of material is very large - there is a risk of misinterpretation.

CRediT authorship contribution statement

Lydie Samie: Conceptualization, Methodology, Software, Validation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Christophe Champod**: Conceptualization, Methodology, Software, Validation, Writing – review & editing, Supervision. **Séverine Delemont:** Conceptualization, Validation, Writing – review & editing. **Patrick Basset**: Conceptualization, Validation, Writing – review & editing. **Tacha Hicks**: Conceptualization, Validation, Writing – original draft, Writing – review & editing. **Vincent Castella**: Conceptualization, Validation, Resources, Writing – review & editing, Supervision, Project administration.

Conflict of Interest

No conflict of interest.

Acknowledgements

We would like to express our gratitude to Prisca Blandin and Annalisa Grini for performing experiments. We also wish to acknowledge the organizers of the proficiency tests from GEDNAP, from SSML and the French speaking subgroup of the ISFG.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.forsciint.2022.111174.

References

[1] G. Jackson, S. Jones, G. Booth, C. Champod, I.W. Evett, The nature of forensic science opinion—a possible framework to guide thinking and practice in investigation and in court proceedings, Sci. Justice 46 (1) (2006) 33–44.
[2] A. Biedermann, F. Taroni, Bayesian networks for evaluating forensic DNA profiling evidence: a review and guide to literature, Forensic Sci. Int.: Genet. 6 (2012) 147–157.
[3] T.R.D. Wolff, A.J. Kal, C.E.H. Berger, B. Kokshoorn, A probabilistic approach to body fluid typing interpretation: an exploratory study on forensic saliva testing, Law Probab. Risk 14 (4) (2015) 323–339.
[4] D. Taylor, D. Abarno, C. Champod, T. Hicks, Evaluating forensic biology results given source level propositions, Forensic Sci. Int. Genet. 21 (2016) 54–67.
[5] D. Taylor, Probabilistically determining the cellular source of DNA derived from differential extractions in sexual assault scenarios, Forensic Sci. Int. Genet. 24 (2016) 124–135.
[6] A.P. Gill, T. Hicks, J. Butler, E. Connolly, L. Gusmão, B. Kokshoorn, N. Morling, R. van Oorschot, W. Parson, M. Prinz, P. Schneider, T. Sijen, D. Taylor, DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence – Guidelines highlighting the importance of propositions Part I: evaluation of DNA profiling comparisons given (sub-) source propositions, Forensic Sci. Int. Genet. 36 (2018) 189–202.
[7] J.P. Allery, N. Telmon, R. Mieusset, A. Blanc, D. Rougé, Cytological detection of spermatozoa: comparison of three staining methods, J. Forensic Sci. 46 (2) (2001) 349–351.
[8] A.E. Kipps, P.H. Whitehead, The significance of amylase in forensic investigations of body fluids, Forensic Sci. 6 (1975) 137–144.
[9] S.K. Anderson, K.G. Olesen, F.V. Jennings, F. Jensen, HUGIN – a shell for building Bayesian belief universes for expert systems, in: Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, 1989, 1080–1085.
[10] P. Basset, P. Blandin, A. Grini, S. Delémont, L. Samie, V. Castella, A simplified protocol for the detection of blood, saliva and semen from a single biological trace using immunochromatographic tests, A simplified protocol for the detection of blood, saliva and semen from a single biological trace using immunochromatographic tests (2022), https://doi.org/10.1007/s12024-021-00453-2 In press.
[11] D.G. Casey, J. Price, The sensitivity and specificity of the RSIDTM-saliva kit for the detection of human salivary amylase in the Forensic Science Laboratory, Dublin, Ireland, Forensic Sci. Int. 194 (2010) 67–71.
[12] J. Old, B. Schweers, P.W. Boonlayangoor, K. Reich, Developmental validation studies of RSID-saliva, RSIDTM-lateral flow immunochromatographic strip test for the forensic detection of saliva, RSIDTM-developmental validation Rev. C, Indep. Forensics (2010) 1–33.
[13] I. Carboni, S. Rapi, U. Ricci, Stability of humana-salivary amylase in aged forensic samples, Leg. Med. 16 (2014) 214–217.
[14] J.M. Kishbaugh, S. Cielski, A. Fotusky, S. Lighthart, K. Maguire, L. Quarino, J. Conte, Detection of prostate specific antigen and salivary amylase in vaginal swabs using SERATEC® immunochromatographic assays, Forensic Sci. Int. 304 (2019) 109899.
[15] S.J. Denison, E.M. Lopes, L. D'Costa, J.C. Newman, Positive prostate-specific antigen (PSA) results in semen-free samples, Can. Soc. Forensic Sci. J. 37 (4) (2004) 197–206.
[16] S.E. Bitner, False positives observed on the seratec® PSA SemiQuant cassette test with condom lubricants, J. Forensic Sci. 57 (2012) 1545–1548.
[17] M.N. Hochmeister, B. Budowle, O. Rudin, C. Gehrig, U. Borer, M. Thali, R. Dirnhofer, Evaluation of prostate-specific antigen (PSA) membrane test assays for the forensic identification of seminal fluid, J. Forensic Sci. 44 (1999) 1057–1060.
[18] I. Feine, R. Gafny, I. Pinkas, Combination of prostate-specific antigen detection and micro-Raman spectroscopy for confirmatory semen detection, Forensic Sci. Int. 270 (2017) 241–247.
[19] E. Kafarowski, K. Dann, J.R.H. Frappier, J.C. Newman, Examination of Semen-Free Vaginal Swabs for p30 Using the SERATEC® PSA Test Kit: A Further Evaluation of the Specificity of p30/PSA for Semen Identification, MAAFS, MAFS, SAFS,CSFS Joint Meeting September 19–24, 2004, Orlando, Florida.
[20] M.N. Hochmeister, B. Budowle, R. Sparkes, O. Rudin, C. Gehrig, M. Thali, L. Schmidt, A. Cordier, R. Dirnhofer, Validation studies of an immunochromatographic 1-step test for the forensic identification of human blood, J. Forensic Sci. 44 (3) (1999) 597–602.
[21] M. Cocuzza, C. Alvarenga, R. Pagani, The epidemiology and etiology of azoospermia, Clinics 68 (S1) (2013) 15–26.
[22] H.H. Kim, P.N. Schlegel, Vastectomy, AUA Update Ser. 27 (2008) 197–207.
[23] R. Rahban, L. Priskorn, A. Senn, E. Stettler, F. Galli, J. Vargas, M. Van den Bergh, A. Fusconi, R. Garlantezec, T.K. Jensen, L. Multigner, N.E. Skakkebæk, M. Germond, N. Jørgensen, S. Nef, †the NICER Working Group, Semen quality of young men in Switzerland, Andrology 7 (2019) 818–826.
[24] OFS, Statistiques de la santé 2019, 2019.
[25] ENFSI, ENFSI Guideline for Evaluative Reporting in Forensic Science: Strengthening the Evaluation of Forensic Results Across Europe (STEOFRAE), 2015.
[26] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2018. ⟨https://www.R-project.org/⟩.
[27] RStudio Team, RStudio: Integrated Development Environment for R, RStudio, Inc, Boston, MA, 2015. ⟨http://www.rstudio.com/⟩.
[28] W. Chang, J. Cheng, J. Allaire, Y. Xie, J. McPherson, shiny: Web Application Framework for R, 2019. Retrieved from ⟨https://CRAN.R-project.org/package=shiny⟩.
[29] D. Attali shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds, 2020. Retrieved from ⟨https://CRAN.R-project.org/package=shinyjs⟩.
[30] C. Ganz, rintrojs: a wrapper for the Intro.js library, J. Open Source Softw. 1 (2016) 63.
[31] C. Dethlefsen, S. Højsgaard, A common platform for graphical models in R: the gRbase package, J. Stat. Softw. 14 (17) (2005) 1–12.
[32] S. Højsgaard, Graphical independence networks with the gRain package for R, J. Stat. Softw. 46 (2012) 1–26.
[33] M. Scutari, Learning Bayesian networks with the bnlearn R package, J. Stat. Softw. 35 (2010) 1–22.
[34] D. Taylor, T. Hicks, C. Champod, Using sensitivity analyses in Bayesian Networks to highlight the impact of data paucity and direct future analyses: a contribution to the debate on measuring and reporting the precision of likelihood ratios, Sci. Justice 56 (2017) 402–410.
[35] M. Plummer, Bayesian Graphical Models Using MCMC, RJAGS, 2012.
[36] A. Biedermann, J. Vuille, Understanding the logic of forensic identification decisions (without numbers), Sui-generis (2018) 397–413.
[37] P.D. Maskell, G. Jackson, Presumptive drug testing–the importance of considering prior probabilities, Wires Forensic Sci. 2 (4) (2020) e1371.
[38] P. Gill, T. Hicks, J.M. Butler, E. Connolly, L. Gusmão, B. Kokshoorn, N. Morling, R.A.H. van Oorschot, W. Parson, M. Prinz, P.M. Schneider, T. Sijen, D. Taylor, DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence – Guidelines highlighting the importance of propositions. Part II: Evaluation of biological traces considering activity level propositions, Forensic Sci. Int. Genet. 44 (2020) 102186.
[39] E. Essen-Möller, Die Beweiskraft, der Ähnlichkeit im Vaterschaftsnachweis – theoretische Grundlagen, Mitt. Anthropol. Ges. 68 (1938) 9–53.
[40] C. Champod, I.W. Evett, A probabilistic approach to fingerprint evidence, J. Forensic Identif. 51 (2001) 101–122.
[41] F. Taroni, C. Aitken, P. Garbolino, A. Biedermann, Bayesian Networks and Probabilistic Inference in Forensic Science, John Wiley & Sons, Chichester, England, 2006, pp. 1–345.
[42] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, Sci. Justice 38 (4) (1998) 231–240.
[43] J. Zoete, W. Oosterman, B. Kokshoorn, M. Sjerps, Cell type determination and association with the DNA donor, Forensic Sci. Int. Genet. 25 (2016) 97–111.