A generalised Bayes' factor formula for evidence evaluation under activity level propositions: variations around a fibres scenario

Citation for published version:
Taroni, F, Garbolino, P & Aitken, C 2021, 'A generalised Bayes' factor formula for evidence evaluation under activity level propositions: variations around a fibres scenario', *Forensic Science International*, vol. 110750.
https://doi.org/10.1016/j.forsciint.2021.110750

Digital Object Identifier (DOI):
10.1016/j.forsciint.2021.110750

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Forensic Science International

Forensic Science International 322 (2021) 110750

Contents lists available at ScienceDirect

Forensic Science International

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

A generalised Bayes' factor formula for evidence evaluation under activity level propositions: Variations around a fibres scenario

Franco Taroni[a,*], Paolo Garbolino[b], Colin Aitken[c]

[a] School of Criminal Justice, University of Lausanne, Lausanne, Switzerland
[b] Department of Architecture and Arts, IUAV University, Venice, Italy
[c] School of Mathematics and Maxwell Institute, University of Edinburgh, Edinburgh, United Kingdom

A R T I C L E I N F O

Article history:
Received 4 October 2020
Received in revised form 7 February 2021
Accepted 8 March 2021
Available online 15 March 2021

Keywords:
Bayes' factor
Scientific evidence evaluation
Activity level proposition

A B S T R A C T

Generalised Bayes' factors and associated Bayesian networks are developed for the transfer of extrinsic evidence at the activity level, developments that extend previous work on activity level evaluation. A strategy for the assessment of extrinsic evidence is developed in stages with progressive increases in complexity. The final development is illustrated with an example involving fibres from clothing. This provides a list of factors involved in the consideration of a transfer case with activity level propositions and their roles in the determination of evidential value.

1. Introduction

For the assessment of scientific evidence, the forensic scientist should consider (at least two) propositions proposed by the prosecution and the defence, respectively, to illustrate their description of the facts under examination. Propositions are formalised representations of the framework of circumstances and depend on case information and the allegations of each of the parties, parties which have different key issues. The key issues are formally defined by the ENFSI Guideline for Evaluative Reporting in Forensic Science [6] as follows:

> The key issue(s) represent those aspects of a case on which a Court, under the law of the case, seeks to reach a judgement. The key issue(s) provide the general framework within which requests to forensic practitioners and propositions (for evaluative reporting) are formally defined. (p. 21)

A classification (a so-called *hierarchy*) of these propositions into three main categories or levels has been proposed by Cook et al. [4], notably the source level (level I), the activity level (level II), and the offence level (level III) propositions. Generally, the lower the level (with offence level being the highest level) at which the evidence is assessed, the more limited will be the importance of the results in the context of the case as a whole discussed in court. For ease of simplicity, note that even if the value of the evidence is such as to add considerable support to the proposition that the evidence comes from the person of interest, this does not help determine whether the recovered material had been transferred during the criminal action or for some innocent reason. Consequently, there is often dissatisfaction if the scientist's evaluation is restricted to level I propositions. Comments on this aspect can be found in [3,14,10].

The ENSFI guideline [6] emphasised this aspect, supporting evaluations under propositions at activity level. The guideline specifies:

> Activity level propositions should be used when expert knowledge is required to consider factors such as transfer mechanisms, persistence and background levels of the material which could have an impact on the understanding of scientific findings relative to the alleged activities. This is particularly important for trace materials such as microtraces (fibres, glass, gunshot residues, other particles) and small quantities of DNA, drugs or explosives (p. 11)

In summary, the available information, the context of the case and the key issue(s) influence the choice of propositions. Propositions should be amenable to a reasoned assignment of credibility by a judicial body and be able to be used for rational inference as emphasised by the European Network of Forensic Science Institutes [6].

This paper focuses on evaluation using Bayes' factors and Bayesian networks (Bayes' nets, BNs for short) for the transfer of so-called *extrinsic* evidence at activity level. A distinction needs to be

* Corresponding author.
E-mail address: Franco.Taroni@unil.ch (F. Taroni).

https://doi.org/10.1016/j.forsciint.2021.110750
0379-0738/

drawn between what is called *intrinsic* evidence and what is called *extrinsic* evidence. Intrinsic evidence is evidence that is immutably associated with a particular source. An example is the DNA evidence of a person. Extrinsic evidence is evidence that is not immutably associated with a particular person. An example is that of fibres from a woollen pullover. The pullover may be associated in general with a particular person but, without further evidence, it cannot be associated with the person at the time of interest to a criminal investigation.

This paper extends previous work on activity level evaluation, mainly in the field of DNA evidence (intrinsic evidence), by Hicks et al. [11] and Taylor et al. [17–20]. A strategy for the assessment of extrinsic evidence is developed in stages with progressive increases in complexity from a basic scenario. A particular scenario is not presented but there is an incremental development of a complete probabilistic model. The development concludes with an explanation of the potential for Bayes' Nets to deal with the most complex of cases.

The paper also extends work done by the current authors on transfer evidence, activity level propositions and Bayes' nets. The formal development generalises previous work to consider various extensions culminating in (27). This equation incorporates terms that allow for

- the role of various intermediate propositions such as transfer from a person other than the person of interest (PoI),
- legitimate transfer from a known source,
- secondary transfer from the person of interest (PoI) via a third party,
- transfer by a third party,
- uncertainty associated with the origin of the source and
- a potential false positive association.

A Bayes net, Fig. 1, provides a graphical representation of (27). Cross-transfer evidence is not considered. Bayes' nets are used to help provide results for a generalised model.

Aitken et al. [1] considered intrinsic evidence (DNA), cross-transfer evidence and potential false associations. Note that the node for the transfer of background material in the BN in [1] is not included in the BN here (Fig. 1) because the current formulation renders it unnecessary. The probabilities for the transfer of background material are considered in the conditional probability table of the node representing evidential material. Taroni et al. [13] considered extrinsic evidence with uncertainty about the true source but did not consider evidence of cross-transfer or potential false associations. Taroni et al. [16] considered extrinsic evidence with cross-transfer but with no uncertainty about the true source, no consideration of false positive associations or the possibility of secondary transfer. Of the factors considered in these papers, only cross-transfer is not considered here.

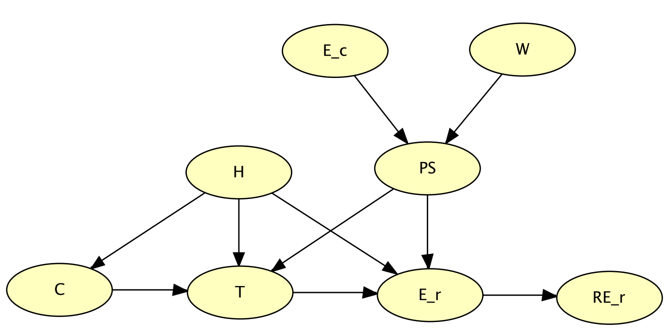

Fig. 1. Bayesian network for reported evidence applicable for transfer from PoI to the victim or from victim to PoI. Node descriptors are given in Table 1 with definitions that depend on a direction of transfer from the PoI to the victim.

All the other factors are brought together in one model for extrinsic evidence.

The ideas are illustrated with an example involving fibres from clothing. The activity is that of a hit on a person, known hereafter as the victim. The hit may have been accidental and so not an offence. Thus the person who hit the victim is referred to as the hitter, not the offender. The person who is the subject of the activity level propositions is known as the person of interest, who may or may not be the hitter. This illustration provides a list of the factors involved in the consideration of a transfer case with activity level propositions and illustrates their roles in the determination of evidential value.

The paper is structured as follows. Section 2.1 develops a Bayes' factor for direct transfer evidence between the hitter and the victim. The transfer is considered separately in each direction. Cross-transfer evidence is beyond the scope of the paper. Section 2.2 develops a Bayes' factor to allow for a contact between the hitter and the victim in which there is no transfer of trace evidence. Section 3 discusses the possibility of an absence of the transfer of evidential material when there was contact between the hitter and the victim. Section 4 extends the formulation of the Bayes' factor to consider, first, transfer material left on the victim by a third party, not the PoI, and, second, a secondary transfer from the PoI via a third person. Generalisations of the formulae for the Bayes' factors developed in the previous sections to allow for uncertainty about the source of the transfer material and the possibility of error in reporting the evidence are given in Section 5. Section 6 extends such generalisation through a representation with Bayes' nets. Some illustrative numerical examples of application of the generalised Bayes' factor are presented in Section 7. A summary of the paper is given in Section 8.

2. Direct transfer between victim and hitter

2.1. Material assumed to be transferred from the hitter to the victim

An expression for the Bayes' factor in cases invoking propositions that put forward actions committed by the PoI or by another person was originally developed in a seminal paper by Evett [7]. For the sake of illustration and clarification of notation, consider the following scenario involving an assault where characteristics describing recovered material (say, textile fibres), E_r, found on the victim, are similar to characteristics of control material coming from the PoI, E_c. Define the evidence E as (E_r, E_c) and more precisely as a single group of recovered material. The size of the group is assumed unchanged throughout and is thus not included in the notation. Consider activity propositions.

H_p: The PoI hit the victim;
H_d: Some person other than the PoI hit the victim.

For the activity to be considered as an offence a factor such as intent to cause harm has to be considered. The action of hitting may or may not have been illegal. The Bayes' factor can be expressed as

$$V = \frac{\Pr(E|H_p, I)}{\Pr(E|H_d, I)},$$

where I characterises the background information that will be omitted in what follows for ease of notation. The letter V has been chosen for the notation to emphasise that the Bayes' factor quantifies the *value* of the evidence.

There are two explanations for the presence of the evidence of the fibres.

- The recovered group of fibres was transferred, has persisted and has been successfully recovered from the victim. In this situation, the group of fibres were not present on the victim before the action of the hit, known hereafter as the activity. Denote this explanation T.

- The recovered group of fibres was not transferred in the activity. In this situation, the recovered fibres are unconnected with the action under investigation: the fibres were on the victim before the activity. Denote this explanation $\bar{T}$.

These two explanations are association propositions [8].

Assuming that all the fibres that have been transferred are from one source, the Bayes' factor leads to:

$$V = \frac{\Pr(E|T, H_p,)\Pr(T|H_p) + \Pr(E|\overline{T}, H_p)\Pr(\overline{T}|H_p)}{\Pr(E|T, H_d,)\Pr(T|H_d) + \Pr(E|\overline{T}, H_d)\Pr(\overline{T}|H_d)}. \quad (1)$$

In the numerator of (1), $\Pr(E|T, H_p)$ represents the probability of observing a group of corresponding fibres on the victim, given that the group was transferred during the activity, has persisted and was recovered successfully, and that the PoI hit the victim. If these conditions are true, this implies the group of fibres was not there before the activity. The probability of this event is denoted b_0, the probability of the presence by chance of no groups of fibres[1].

$\Pr(T|H_p)$ represents the probability that a group of corresponding fibres was transferred, has persisted and was recovered successfully from the victim, given that the PoI hit them. This probability is denoted t_p.

$\Pr(E|\overline{T}, H_p)$ is the probability that a group of corresponding fibres are recovered from the victim, given that the PoI hit them and that there was no transfer of fibres during the activity, and hence no persistence or recovery. If the group of fibres was not transferred in the activity, it was present on the victim beforehand. Let $b_1 \times \gamma$ represent the probability b_1 of the chance occurrence of a group of fibres on the victim linked to the relevant population proportion γ for the corresponding characteristics. The development of the generalised Bayes' factor assumes the discovery of one, and only one, group of fibres. This group is either from a background population or not. Thus $b_0 + b_1 = 1$.

$\Pr(\overline{T}|H_p)$ represents the probability that no group of fibres was transferred, persisted or recovered successfully from the PoI's clothing to the victim. This probability is denoted $1 - t_p$.

Consider the terms in the denominator of (1). $\Pr(E|T, H_d)$ represents the probability of finding a group of corresponding fibres given that the PoI did not hit the victim. The victim had this group of fibres before the activity and the event of the shared characteristics is one of chance. This probability is $b_0 \times \gamma$. In the numerator the probability γ is replaced by 1 since the source of the fibres is assumed by the conditioning. Hence the corresponding term is just b_0.

$\Pr(T|H_d)$ represents the probability that a group of corresponding fibres was transferred, persisted and recovered successfully from the victim given that the PoI did not hit the victim. Denote this probability by t_d.

$\Pr(E|\overline{T}, H_d)$ is the probability that a group of corresponding fibres is observed on the victim given that the PoI did not hit the victim and that this group of fibres was not transferred, persisted or recovered successfully during the activity. If the group of fibres was not transferred, it was present on the victim before the activity, with associated probability $b_1 \times \gamma$.

$\Pr(\bar{T}|H_d)$ represents the probability there was no transfer from the hitter to the victim, persistence or recovery of a group of fibres. This probability is denoted $1 - t_d$.

Therefore, (1) becomes

$$V = \frac{b_0 t_p + b_1\gamma(1 - t_p)}{b_0\gamma t_d + b_1\gamma(1 - t_d)}. \quad (2)$$

[1] Definitions of all the probabilistic notation are listed in Table 8 in the Appendix A.

2.2. Material assumed to be transferred from the victim to the hitter

Consider a transfer of fibres in the other direction from that of Section 2.1. A group of external fibres has been recovered on clothing of the PoI (e.g. a pullover). The defence deny that their client (the PoI) hit the victim, and hence the recovered fibres are not related to the activity. These fibres are on the PoI's pullover by chance alone from another activity unconnected to the activity of H_p. This is an important point for the development of the Bayes' factor. The numerator of the Bayes' factor is the same as in (2) but there is a change in the denominator. There is no reason to develop $\Pr(E|H_d)$ using the association propositions T and $\bar{T}$ because the fibres are not considered to be the result of transfer, persistence and recovery following an alleged activity. The Bayes' factor thus reduces to:

$$V = \frac{b_0 t_p + b_1\gamma(1 - t_p)}{b_1\gamma}. \quad (3)$$

Consideration of the previous scenarios may change with consideration of the contextual information and the strategy of the defence. For example, a third party may have deposited the fibres on the victim or there may have been a secondary transfer from the PoI to the victim via a third party. These situations are considered in Section 4 and the appropriate Bayes' factor is introduced. A generalised formula can be deduced and adapted for the various scenarios of interest (see Section 5).

3. Consideration of the possible absence of the transfer of evidential material between the hitter and the victim

When considering activity propositions the scientist should pay attention to the logical consequence of the activities. If a person hit a victim, this person (and possibly their clothes) had a physical contact with the victim.

Reconsider the scenario involving the activity of one person hitting another, known as the victim, and transfer from that person to the victim. The characteristics of the recovered material (fibres) found on the victim (E_r) are similar to the characteristics of the control material found on the PoI (E_c). The evidence E, with $E = (E_r, E_c)$, may be thought of as a single group of fibres.

The main (activity level) propositions of interest are: .

H_p: The PoI hit the victim;
H_d: Some person other than the PoI hit the victim.

The presence of the evidence of the fibres on the victim may be explained by consideration of the following association propositions:

T: There was a transfer from the PoI to the victim;
$\bar{T}$: There was not a transfer from the PoI to the victim.

An extension using intermediate association propositions can be suggested by taking into account the logical contact caused by the action. If the PoI hit the victim, they have a physical contact with the victim. Such a contact may involve a transfer of evidential material. Define

C: The victim's clothing has been in contact with that of the PoI;
$\bar{C}$: The victim's clothing has not been in contact with that of the PoI.

Consider the Bayes' factor $V = \Pr(E|H_p)/\Pr(E|H_d)$. As previously noted in Section 2.1, the numerator of the Bayes' factor is obtained by extending the conversation to propositions T and $\overline{T}$.

$$\Pr(E|H_p) = \Pr(E|T, H_p)\Pr(T|H_p) + \Pr(E|\overline{T}, H_p)\Pr(\overline{T}|H_p), \quad (4)$$

where $\Pr(E|T, H_p) = b_0$. The conditional probability $\Pr(T|H_p)$ needs to take into consideration the uncertainty about propositions C and $\bar{C}$, leading to

$$\Pr(T|H_p) = \Pr(T|C, H_p)\Pr(C|H_p) + \Pr(T|\bar{C}, H_p)\Pr(\bar{C}|H_p). \quad (5)$$

Consideration needs to be given to the four conditional probabilities of (5).

$\Pr(T|C, H_p)$ represents the probability that a transfer to the victim occurred, given contact and H_p. Denote this probability by the letter t_1.

$\Pr(C|H_p)$ is the probability that there has been contact between the clothes of the victim and the clothes of the PoI, given that the PoI hit the victim. This probability is c. This probability assignment depends largely on the circumstances of the case, in particular the information on how the activity occurred. It is possible that c may not be equal to 1; the PoI may have hit the victim without their clothes coming into contact.

$\Pr(T|\bar{C}, H_p) = 0$; there is no possibility for transfer when there was no contact.

$\Pr(\bar{C}|H_p)$ is the complement of $\Pr(C|H_p)$ and equals $(1 - c)$.

Consider now the third conditional probability of the right-hand-side of (4), $\Pr(E|\bar{T}, H_p)$. It equals $b_1\gamma$, the probability of transfer other than to the victim from the PoI. This is the chance occurrence of a single group of fibres on the victim's clothing linked to the relevant population proportion γ for the observed characteristics of the clothing of the PoI.

Finally, $\Pr(\bar{T}|H_p)$ is also obtained in an extension to C, that is

$$\Pr(\bar{T}|H_p) = \Pr(\bar{T}|C, H_p)\Pr(C|H_p) + \Pr(\bar{T}|\bar{C}, H_p)\Pr(\bar{C}|H_p), \quad (6)$$

where $\Pr(\bar{T}|C, H_p) = (1 - t_1)$, $\Pr(C|H_p) = c$, $\Pr(\bar{T}|\bar{C}, H_p) = 1$ and $\Pr(\bar{C}|H_p) = (1 - c)$.

The numerator of the Bayes' factor becomes

$$\Pr(E|H_p) = b_0 t_1 c + b_1\gamma[(1 - t_1)c + (1 - c)]. \quad (7)$$

Consider now the denominator $\Pr(E|H_d)$ of the Bayes' factor. It is also obtained by taking into account the uncertainty about propositions T and $\bar{T}$ by writing

$$\Pr(E|H_d) = \Pr(E|T, H_d)\Pr(T|H_d) + \Pr(E|\bar{T}, H_d)\Pr(\bar{T}|H_d), \quad (8)$$

where $\Pr(E|T, H_d) = b_0$ and the conditional probability $\Pr(T|H_d)$ is obtained by an extension to C and $\bar{C}$:

$$\Pr(T|H_d) = \Pr(T|C, H_d)\Pr(C|H_d) + \Pr(T|\bar{C}, H_d)\Pr(\bar{C}|H_d). \quad (9)$$

$\Pr(T|C, H_d) = f$ represents the transfer probability of fibres from the PoI to the victim, given that someone other than the PoI hit the victim. Note that the nature of the activity and the position where the group of fibres was found may lead to different probability assignments for t_1 and f, under H_p and H_d, respectively.

$\Pr(C|H_d) = d$ is the probability associated with the event that the clothing of the PoI could have been in contact with the clothing of the victim for reasons other than the activity.

Note also that $\Pr(T|\bar{C}, H_d) = 0$. As under proposition H_p, there is no possibility for the transfer of fibres when there was no contact.

The denominator of the Bayes' factor becomes

$$\Pr(E|H_d) = b_0 f d + b_1\gamma[(1 - f)d + (1 - d)]. \quad (10)$$

The Bayes' factor is

$$V = \frac{b_0 t_1 c + b_1\gamma[(1 - t_1)c + (1 - c)]}{b_0 f d + b_1\gamma[(1 - f)d + (1 - d)]}. \quad (11)$$

From this development, it can be seen that, if $c = 1$ (meaning that it is assumed that the clothing of the victim and the clothing of the PoI have been in contact, given that the PoI hit the victim) and either $d = 0$ (meaning that it is assumed that the clothing of the PoI and the clothing of the victim have not been in contact for reasons other than the activity) or $f = 0$, recognising the circumstance that though there was a contact of clothing, $d \neq 0$, there was not a transfer of fibres, then the Bayes' factor becomes analogous to (3):

$$V = \frac{b_0 t_1 + b_1\gamma t_0}{b_1\gamma},$$

where t_1 denotes the probability that one group of fibres has been transferred, persisted and successfully recovered from clothing from the victim and $t_0 = 1 - t_1$.

Other scenarios of interest were presented in [2,9].

Consider an alternative scenario involving recovered fibres on the seat of a car that belongs to a man who is suspected of abducting a woman and attempting to rape her. There is a single group of foreign red woollen fibres that have been collected from the passenger seat of the car. The victim was wearing a red woollen pullover. According to the PoI, no one other than his wife ever sits on the passenger seat. In addition, the car seats had been vacuumed recently but before the alleged rape is thought to have taken place. The PoI denies that the victim has ever been in contact with the car. In such a case, an issue of concern is that the victim sat on the passenger seat of the PoI's car (H_p) with its converse, the victim has never sat on the passenger seat of the PoI's car (H_d). Proposition C refers to the event that victim has been in contact with the seat and alternatively, $\bar{C}$, the victim has never been in contact with the seat. Here, it appears reasonable to assume that $\Pr(C|H_p) = c = 1$ and $\Pr(C|H_d) = d = 0$. The numerator of the Bayes' factor is then $b_0 t_1 + (1 - t_1)b_1\gamma$. Given H_d, the transfer probability $\Pr(T|C, H_d) = f = 0$ and the denominator of the Bayes' factor is $b_1\gamma$. The resulting Bayes' factor is:

$$V = \frac{b_0 t_1 + b_1\gamma(1 - t_1)}{b_1\gamma}. \quad (12)$$

Further considerations of the possibilities of transfer are those of secondary transfer and third party transfer.

4. Secondary transfer from the hitter and third party transfer to the victim

Consider again the scenario of Section 2.1 where textile fibres are recovered from the clothing of a victim. The characteristics of the recovered fibres are similar to those of control material from a PoI. The propositions of interest are H_p, the PoI hit the victim, and H_d, some person other than the PoI hit the victim.

Recall (4) and consider the numerator first.

$$\Pr(E|H_p) = \Pr(E|T, H_p)\Pr(T|H_p) + \Pr(E|\bar{T}, H_p)\Pr(\bar{T}|H_p).$$

As before, $\Pr(E|T, H_p) = b_0$ and $\Pr(E|\bar{T}, H_p) = b_1\gamma$. The probabilities $\Pr(T|H_p)$ and $\Pr(\bar{T}|H_p)$ are extended with consideration of the possibilities for the mechanism of transfer. There are three possibilities for the occurrence of a transfer: a transfer can occur if (a) there was contact of clothing of the victim and the PoI, either legitimately or as part of the activity (hit), (b) a third party committed the activity, and (c) a secondary transfer (involving fibres similar to those of the PoI's) occurred. Consider the following three (exhaustive) associate intermediate propositions:

C: the PoI has been in contact with the victim;
T_A: a third party has been in contact with the victim;
T_S: the PoI has been in contact with a third party who transferred the PoI's fibres to the victim.

Under proposition H_p (the PoI hit the victim), only proposition C is of interest and

$$\Pr(T|H_p) = \Pr(T|C, H_p)\Pr(C|H_p) = t_1, \quad (13)$$

where $\Pr(C|H_p) = 1$ and $\Pr(T|C, H_p) = t_1$.

$$\Pr(\bar{T}|H_p) = \Pr(\bar{T}|C, H_p)\Pr(C|H_p), \quad (14)$$

with $\Pr(C|H_p) = 1$ and $\Pr(\bar{T}|C, H_p) = (1 - t_1)$.

Under the alternative proposition H_d (some person other than the PoI hit the victim), the denominator is

$$\Pr(E|H_d) = \Pr(E|T, H_d)\Pr(T|H_d) + \Pr(E|\bar{T}, H_d)\Pr(\bar{T}|H_d),$$

where $\Pr(E|T, H_d) = b_0\gamma$ because the correspondence of the characteristics of the fibres is one of chance given that the PoI is not the hitter, and $\Pr(E|\bar{T}, H_d) = b_1\gamma$.

$\Pr(T|H_d)$ and $\Pr(\bar{T}|H_d)$ are extended considering relevant intermediate associate propositions under H_d, say propositions T_A and T_S. So

$$\Pr(T|H_d) = \Pr(T|T_A, H_d)\Pr(T_A|H_d) + \Pr(T|T_S, H_d)\Pr(T_S|H_d), \tag{15}$$

where $\Pr(T_A|H_d)$ and $\Pr(T_S|H_d)$ refer to the probability that a third party exists and has transferred textile fibres, and to the probability that a secondary transfer has occurred from the PoI via the third party given that the PoI is not the hitter. Note that $\Pr(T_A|H_d) + \Pr(T_S|H_d) = 1$.

The Bayes' factor becomes

$$V = \frac{b_0 t_1 c + b_1\gamma(1 - t_1)}{b_0\gamma[t_1' t_a + t''_1 t_s] + b_1\gamma[(1 - t'_1)t'_a + (1 - t''_1)t'_s]}, \tag{16}$$

where t'_a is the probability of a contact between clothing of the victim and clothing of a third party given the PoI did not hit the victim, t'_1 is the probability of a transfer of a group of fibres from a third party to the victim, given there was contact between clothing of the victim and the clothing of the third party, whether or not the PoI hit the victim, t'_s is the probability of the occurrence of a secondary contact from the clothing of the PoI, via a third party, to the clothing of the victim given the PoI did not hit the victim and t''_1 is the probability of a transfer of a group of fibres to the victim from the PoI, given there was secondary contact from clothing of the PoI, via the clothing of a third party, to clothing of the victim, whether or not the PoI hit the victim.

Consider the following situation. Given that $c = 1$ and $t'_s = 0$, meaning that a secondary transfer is considered impossible (and so $t'_a = 1$), the value of the evidence becomes:

$$V = \frac{b_0 t_1 + b_1\gamma(1 - t_1)}{b_0\gamma t'_1 + b_1\gamma(1 - t'_1)},$$

as in (2).

On the other hand, if $t'_a = 0$, meaning that a third party cannot be considered as relevant (and so $t'_s = 1$), then the Bayes' factor is

$$V = \frac{b_0 t_1 + b_1\gamma(1 - t_1)}{b_0\gamma t''_1 + b_1\gamma(1 - t''_1)}.$$

Consider the situation where the victim's pullover is known to be new and hence has never been in contact with other articles of clothing so b_0 can be taken equal to 1, and hence $b_1 = 0$, then the Bayes' factor becomes

$$V = \frac{t_1}{\gamma t''_1},$$

If $t_1 = t''_1$, the Bayes' factor reduces to the classical source level expression $1/\gamma$.

Consider the situation where the possibility of third party involvement and the possibility of a secondary transfer of fibres are deemed equally likely, so that $t'_a = t'_s = 0.5$. The Bayes' factor then becomes

$$V = \frac{b_0 t_1 + b_1\gamma(1 - t_1)}{0.5\{b_0\gamma[t'_1 + t''_1] + b_1\gamma[(1 - t'_1) + (1 - t''_1)]\}}.$$

In the extreme situation characterised by $b_0 = 1$ and $b_1 = 0$, the Bayes' factor is

$$V = \frac{t_1}{0.5\{\gamma[t'_1 + t''_1]\}},$$

where the denominator is a weighted value between two types of transfer mechanisms.

This development of the Bayes' factor takes into account the list of potential mechanisms for a transfer of the recovered material E_r that corresponds to control material E_c which, together, characterise the evidence E. Three mechanisms have been considered: (a) direct transfer from the PoI to the victim, either legitimately or as the result of a hit, (b) a third party hitter, and (c) a secondary transfer from the PoI via a third party. The numerator and the denominator of the Bayes' factor should take into account the relevant mechanisms under propositions H_p and H_d.

The extended expression of the Bayes' factor is

$$V = \frac{b_0[t_1 c + t'_1 t_a + t''_1 t_s] + b_1\gamma[(1 - t_1)c + (1 - t'_1)t_a + (1 - t''_1)t_s]}{b_0\gamma[fd + t'_1 t'_a + t''_1 t'_s] + b_1\gamma[(1 - f)d + (1 - t'_1)t'_a + (1 - t''_1)t'_s]}, \tag{17}$$

where the three terms in each of the square brackets [...] indicate the probabilistic contributions of the three mechanisms (a), (b) and (c). Eq. (17) provides a general expression for the Bayes' factor for a one-way activity in which the question of issue is whether a PoI hit, or did not hit, a victim and the evidence under consideration is that of the possible transfer of trace evidence from the PoI to the victim.

The probability $b_0\gamma$ in the denominator can be reduced to b_0 if the origin of the recovered fibres is not disputed. In general, $t_a = t_s = 0$ as it is assumed (H_p) that the PoI is the source of the fibres on the victim. As a consequence, the probabilities of secondary transfer (t_s) and of a third party source (t_a) are deemed sufficiently small to be treated as zero. The terms are included in the expression for V in (17) for completeness.

5. A further generalisation for the Bayes' factor

As mentioned in Section 1, evidence of fibres differs from that of DNA in the sense that they are not intrinsic to a given individual; they are *extrinsic*. A given individual has, as far as most of the common typing techniques in forensic science are concerned, one and only one DNA profile (leaving aside biological anomalies and other special cases) and it cannot be deliberately modified; the profile is said to be *intrinsic* to the individual. Most people, however, almost certainly, have more than one pullover. The characteristics of the fibres are not individual to a particular pullover; many pullovers have fibres with similar characteristics. Thus, with items such as pullovers, it is necessary to make assumptions regarding the relationship between a particular pullover and a particular PoI such as when the clothes were worn. The same line of reasoning can be adopted in scenarios involving shoe prints, for example.

The problem of interest is that of uncertainty about the item itself actually worn by the PoI in the event that they committed the action of interest. It may not be known if the item worn by the PoI during the alleged facts (say, event *PS* for PoI's source) is in fact the item available (and analysed) as a known source (control material E_c). So, there is uncertainty associated with the origin of the source and this aspect is not considered in the most general formula (see (17)). The fact that the suspect wore (did not wear) the known source E_c at the time of the activity under investigation has to be taken into account. In all the examples presented earlier in this paper, it was tacitly assumed that there was no uncertainty about the assumed known source. This assumption can be relaxed. Such an extension was presented in [13] and it should be integrated in a generalised formula.

Consider a further step of realistic extension. Note that observations made by the scientist on E_r are subject to uncertainty and it is important to capture and represent this uncertainty explicitly in any probabilistic model adopted. So, consider as evidence the *reported* observation made by the forensic scientist (call this event RE_r) and extend further the conditional probability of interest $\Pr(RE_r|E_c, H_p)$ by considering the *true*, but unknown, characterisation of the recovered material, E_r. The numerator of the Bayes' factor becomes:

$$\Pr(RE_r|E_r, E_c, H_p)\Pr(E_r|E_c, H_p) + \Pr(RE_r|\bar{E}_r, E_c, H_p)\Pr(\bar{E}_r|E_c, H_p).$$

This development has been called 'cascaded inference' [12] and it has been used in [21] to consider the effect of false positives. Note that RE_r is conditionally independent of the hypotheses H and of the control evidence E_c, given E_r, so that the previous expression becomes $\Pr(RE_r|E_r)\Pr(E_r|E_c, H_p) + \Pr(RE_r|\bar{E}_r)\Pr(\bar{E}_r|E_c, H_p)$. A similar expression is derived for the denominator. The Bayes' factor becomes

$$V = \frac{\Pr(RE_r|E_r)\Pr(E_r|E_c, H_p) + \Pr(RE_r|\bar{E}_r)\Pr(\bar{E}_r|E_c, H_p)}{\Pr(RE_r|E_r)\Pr(E_r|E_c, H_d) + \Pr(RE_r|\bar{E}_r)\Pr(\bar{E}_r|E_c, H_d)}. \tag{18}$$

Eq. (18) may be used for the development of a generalised expression for the value of the evidence. Consider the numerator first. The probability $\Pr(RE_r|E_r)$ refers to a value that relates on false negatives of the laboratory inspection process. Denote this probability a. A (unrealistic) value of 1 assumes an inspection process which has no false negatives.

$\Pr(E_r|E_c, H_p)$ refers to the probability the fibres are of type of interest, say x, if the PoI hit the victim and the control characteristics are of type x. To be able to quantify such a probability, one needs to condition on the control material reputedly worn by the PoI during the hitting (PS). The material worn by the PoI at the time of the hitting is unknown; only the control material, E_c, worn by the PoI when he was arrested is known. $\Pr(E_r|E_c, H_p)$ can be extended by taking into account these uncertainties: $\Pr(E_r|PS, E_c, H_p)\Pr(PS|E_c, H_p) + \Pr(E_r|\overline{PS}, E_c, H_p)\Pr(\overline{PS}|E_c, H_p)$ and the numerator of the Bayes' factor becomes:

$$\begin{aligned} &a[\Pr(E_r|PS, E_c, H_p)\Pr(PS|E_c, H_p) + \Pr(E_r|\overline{PS}, E_c, H_p)\Pr(\overline{PS}|E_c, H_p)] \\ &+ \Pr(RE_r|\bar{E}_r, H_p)[\Pr(\bar{E}_r|PS, E_c, H_p)\Pr(PS|E_c, H_p) \\ &+ \Pr(\bar{E}_r|\overline{PS}, E_c, H_p)\Pr(\overline{PS}|E_c, H_p)]. \end{aligned} \tag{19}$$

It can be assumed that E_r is conditionally independent from E_c given PS or $\overline{PS}$. In fact, if one knows that the suspect wore or did not wear the control material (PS and $\overline{PS}$) having characteristics x during the hitting, then the fact that the control material (E_c) seized at the time of arrest has characteristics x is of no longer of interest; it is only the material the PoI wore during the criminal action that is of interest. Therefore, $\Pr(E_r|PS, E_c, H_p) = \Pr(E_r|PS, H_p)$ and $\Pr(E_r|\overline{PS}, E_c, H_p) = \Pr(E_r|\overline{PS}, H_p)$.

The probability the PoI wore the seized pullover of characteristics x, $\Pr(PS|E_c, H_p)$, should take into account that the article of clothing used as the source of control fibres is the one worn by the criminal at the crime, so the possibility the PoI wore the pullover during the commission of the crime (call this event W).

Also, PS is independent of whether or not the PoI is the hitter (H_p) or not (H_d) Thus, $\Pr(PS|E_c, H_p)$ may be written as $\Pr(PS|E_c, W)\Pr(W) + \Pr(PS|E_c, \bar{W})\Pr(\bar{W})$.

- $\Pr(PS|E_c, W)$ equals 1 because if the PoI wore the pullover (W), and that the pullover has a given characteristic of interest (E_c), then it is certain that such a characteristic will be observed for the PoI's source (PS).
- The probability $\Pr(W) = w$ represents the probability the PoI wore the pullover at the time of the relevant activity. Note that such a probability cannot generally be considered as equal to 1 as for intrinsic evidence; it depends on the circumstances of the case.
- $\Pr(PS|E_c, \bar{W}) = \gamma'$. This probability is considered to be different from the general population proportion of the fibres characteristic described by the characteristic x. The reason is that there may be information that indicates that the collection of textile habits of the PoI is not representative of the general population proportions of the various fibres types.

For the special case $\Pr(W) = w = 1$, $\Pr(PS|E_c, W)\Pr(W) + \Pr(PS|E_c, \bar{W})\Pr(\bar{W}) = 1$. For simplicity of notation, denote $\Pr(PS|E_c, W)\Pr(W) + \Pr(PS|E_c, \bar{W})\Pr(\bar{W}) = \rho$, so that the numerator of the Bayes' factor (19) can be re-written as:

$$\begin{aligned} &a[\Pr(E_r|PS, H_p)\rho + \Pr(E_r|\overline{PS}, H_p)(1-\rho)] \\ &+\Pr(RE_r|\bar{E}_r, H_p)[\Pr(\bar{E}_r|PS, H_p)\rho + \Pr(\bar{E}_r|\overline{PS}, H_p)(1-\rho)]. \end{aligned} \tag{20}$$

Consider the term $\Pr(E_r|PS, H_p)$ in (20). The probability of E_r with characteristics, say x, given that the PoI hit the victim and that the PoI's source has characteristics x depends on the fact that fibres have (have not) been transferred, persisted and have been recovered (events T and $\bar{T}$) as introduced in Section 2.1. Therefore,

$$\begin{aligned} \Pr(E_r|PS, H_p) &= \Pr(E_r|PS, T, H_p)\Pr(T|PS, H_p) \\ &+\Pr(E_r|PS, \bar{T}, H_p)\Pr(\bar{T}|PS, H_p). \end{aligned} \tag{21}$$

In an analogous way, $\Pr(E_r|\overline{PS}, H_p)$ becomes

$$\begin{aligned} \Pr(E_r|\overline{PS}, H_p) &= \Pr(E_r|\overline{PS}, T, H_p)\Pr(T|\overline{PS}, H_p) \\ &+\Pr(E_r|\overline{PS}, \bar{T}, H_p)\Pr(\bar{T}|\overline{PS}, H_p), \end{aligned} \tag{22}$$

$\Pr(\bar{E}_r|PS, H_p)$ becomes

$$\begin{aligned} \Pr(\bar{E}_r|PS, H_p) &= \Pr(\bar{E}_r|PS, T, H_p)\Pr(T|PS, H_p) \\ &+\Pr(\bar{E}_r|PS, \bar{T}, H_p)\Pr(\bar{T}|PS, H_p), \end{aligned} \tag{23}$$

and $\Pr(\bar{E}_r|\overline{PS}, H_p)$ becomes

$$\begin{aligned} \Pr(\bar{E}_r|\overline{PS}, H_p) &= \Pr(\bar{E}_r|\overline{PS}, T, H_p)\Pr(T|\overline{PS}, H_p) \\ &+\Pr(\bar{E}_r|\overline{PS}, \bar{T}, H_p)\Pr(\bar{T}|\overline{PS}, H_p). \end{aligned} \tag{24}$$

Given that (a) $\Pr(E_r|PS, T, H_p) = b_0$, (b) $\Pr(E_r|PS, \bar{T}, H_p) = b_1\gamma$, (c) $\Pr(E_r|\overline{PS}, T, H_p) = 0$ because of the incompatibility between the characteristics of E_r and E_c and (d) $\Pr(\bar{E}_r|PS, T, H_p) = 0$, the numerator of the Bayes' factor reduces to:

$$\begin{aligned} &a\{[b_o\Pr(T|PS, H_p) + b_1\gamma\Pr(\bar{T}|PS, H_p)]\rho + [b_1\gamma\Pr(\bar{T}|PS, H_p)(1-\rho)]\} \\ &\quad+\Pr(RE_r|\bar{E}_r, H_p)[(1-b_1\gamma)\Pr(\bar{T}|PS, H_p)]\rho + [Pr(T|\overline{PS}, H_p) \\ &\qquad+(1-b_1\gamma)\Pr(\bar{T}|\overline{PS}, H_p)](1-\rho). \end{aligned} \tag{25}$$

Following the same line of extension reasoning, $\Pr(T|PS, H_p)$ becomes

$$\Pr(T|C, H_p)\Pr(C|H_p) + \Pr(T|T_A, H_p)\Pr(T_A|H_p) + \Pr(T|T_S, H_p)\Pr(T_S|H_p).$$

Call this extension z. $\Pr(\bar{T}|PS, H_p)$ becomes

$$\Pr(\bar{T}|C, H_p)\Pr(C|H_p) + \Pr(\bar{T}|T_A, H_p)\Pr(T_A|H_p) + \Pr(\bar{T}|T_S, H_p)\Pr(T_S|H_p).$$

Call this extension $\bar{z}$.

The numerator of the Bayes' factor is:

$$\begin{aligned} &a\{[b_oz + b_1\gamma\bar{z}]\rho + [b_1\gamma\bar{z}(1-\rho)]\}+ \\ &\Pr(RE_r|\bar{E}_r, H_p)[(1-b_1\gamma)\bar{z}]\rho + [z + (1-b_1\gamma)\bar{z}](1-\rho), \end{aligned} \tag{26}$$

where $\Pr(RE_r|\bar{E}_r, H_p)$ represents the probability to declare a false positive result: the probability to declare $E_r = x$ given that $E_r \neq x$. Denote this probability e.

Finally, given

- $z = t_1 c + t_1^{'} t_a + t_1^{''} t_s$,
- $\bar{z} = (1 - t_1)c + (1 - t_1^{'})t_a + (1 - t_1^{''})t_s$,
- $\rho = w + [\gamma^{'}(1 - w)]$

and assuming that

- the PoI wore the control evidence at the relevant activity time ($w = 1$), so that $\rho = 1$ and $(1 - \rho) = 0$,
- the probability a equals 1, so that there is no false negative, and
- the probability e equals 0, so that there is no false associations (false positives), $e = 0$,

the numerator of the Bayes' factor, Eq. (26), reduces to

$$b_0[t_1 c + t'_1 t_a + t''_1 t_s] + b_1\gamma[(1 - t_1)c + (1 - t'_1)t_a + (1 - t''_1)t_s]$$

as in Eq. (17).

The denominator of the Bayes' factor is developed in an analogous way, with $z' = fd + t'_1 t'_a + t''_1 t'_s$, $\bar{z}' = (1 - f)d + (1 - t'_1)t'_a + (1 - t''_1)t'_s$ and $d + t'_a + t'_s = 1$. The final general expression for the Bayes' factor is therefore

$$V = \frac{a\{[b_o z + b_1\gamma\bar{z}]\rho + [b_1\gamma\bar{z}(1 - \rho)]\} + e\{[(1 - b_1\gamma)\bar{z}]\rho + [z + (1 - b_1\gamma)\bar{z}](1 - \rho)\}}{a\{[b_o\gamma z' + b_1\gamma\bar{z}']\rho + [b_1\gamma\bar{z}'(1 - \rho)]\} + e\{[(1 - b_1\gamma)\bar{z}']\rho + [z' + (1 - b_1\gamma)\bar{z}'](1 - \rho)\}}. \quad (27)$$

6. Representation with a Bayesian network

A Bayes' net is a graphical model whose elements are *nodes*, *arrows* (also called arcs) between nodes, and *probability assignments*. A finite set of nodes together with a set of arrows (i.e. directed link) between nodes forms a mathematical structure called a *directed graph*. If there is an arrow pointing from node, say Y to node, say X, it is said that Y is a *parent* of X and X is a *child* of Y. A node with no parents is called a *root* node. A sequence of consecutive arrows connecting two nodes X and Y, independently from the direction of the arrows, is known as a *path* between X and Y. Nodes represent random variables where the random variable may be either discrete or continuous. In what follows, all the nodes that will be used to describe the scenario of interest will be discrete. The discrete nodes take a finite set of mutually exclusive *states*.

Arrows represent direct relationships amongst variables. For each variable X with parents $Y_1, Y_2, \ldots, Y_n$, there is an associated conditional probability table $\Pr(X|Y_1, Y_2, \ldots, Y_n, I)$, where I denotes, as usual, background information, all the relevant knowledge which does not appear explicitly as a node in the graph. If X is a root node, then its table reduces to probabilities $\Pr(X)$, unconditional on other nodes in the graph, where I has been omitted for ease of notation.

Let X be a discrete variable with n mutually exclusive and exhaustive states $x_1, \ldots, x_n$ where a state is a possible value associated with a random variable. If X is a root node, then the (un)conditional probability table $\Pr(X)$ will be an $n-$table (a table with n entries) containing the probability distribution $\{\Pr(X = x_i), i = 1, \ldots, n\}$, with $\sum_{i=1}^{n} \Pr(X = x_i) = 1$. When the context is sufficiently clear that there will be no confusion in so doing, the subscript i is omitted from the notation and one writes $\Pr(X = x)$ and $\sum \Pr(X = x_i) = 1$.

Table 1
Node descriptors for transfer of fibres from the PoI to the victim.

Node	Definition and states
C	The victim has been in contact with the PoI (C), a third party has been in contact with the victim (T_A), the victim has been in contact with a third party (T_S) who transferred the PoI's fibres to the victim
H	The PoI hit the victim (H_p) or did not hit the victim (H_d)
E_c	Control fibres from PoI $\{x, \bar{x}\}$
E_r	Recovered fibres from victim $\{x, \bar{x}\}$
	E_c and E_r have two possible states $\{x, \bar{x}\}$ which are characteristics of fibres from control source (PoI) and recovered source (victim)
RE_r	Reported observation of E_r
PS	has two states, A and $\bar{A}$, depending on whether or not the person of interest did or did not wear an article of clothing at the time of the crime which had characteristics of the fibres found on the victim (clothing fibres of PoI similar/dissimilar to those found on victim)
T	Transfer (T) or not ($\bar{T}$) of recovered group of fibres transferred to the victim (maybe or maybe not transferred from the PoI), persisted and recovered from the victim
W	The PoI wore (W) or did not wear ($\bar{W}$), the known source of the fibres (E_c) at the time of the activity under investigation

Let Y be a variable with m states y (using the abbreviated notation introduced above). If Y is a parent of X, then the conditional probability table $\Pr(X|Y)$ will be an $n \times m$ table containing all the probability assignments $\Pr(X = x|Y = y)$. Notice that the notation for the states of the variables as used here below may vary with the domain of application. It may happen, as in the model depicted in Fig. 1, that the name of a variable, say E_r, also characterises a state of that variable (e.g., the observed characteristics of the recovered fibres and their converse, all other possible characteristics these fibres may have). On other occasions, the states of a variable may be described differently from the name of the variable. A more formal description and definitions of BNs and their components can be found in [5,15].

The illustrative example here for activity propositions is transfer of trace evidence in the form of fibres from a PoI to a victim. A formulaic generalisation of the Bayes' factor for the evaluation of trace evidence with activity propositions is provided in Section 5 and (27). A template for the corresponding Bayes' net is illustrated in Fig. 1 with the definitions of the node descriptors given in Table 1. The nodes in the network list the factors that should be considered in the evaluation of evidence. The links in the network list the conditional probabilities that need to be assigned in order to obtain a numerical value for the evidence.

Consider Fig. 1. The connections between nodes H, T and E_r characterise a model which describes the results expressed by Eqs. (2) and (3). Extension of the model using node C, as defined in Table 1, gives the results of (11) and (12).

Nodes E_c, W and PS take account of the possibility that there is uncertainty about the known source of the fibres found on the victim. Node RE_r represents the uncertainty inherent in the observations E_r made by the scientist. The separation of the node RE_r from the node H by the node E_r represents the assumption that RE_r is conditionally independent of H, given E_r.

The reasoning associated with the Bayes' network of Fig. 1 is as follows. There are three root nodes. The first, H, represents the activity level propositions. For the fibres scenario described in this paper, these are that the person of interest hit, or did not hit, the victim. The second root node, E_c, represents the characteristics of the fibres of an article of clothing (pullover) associated with the PoI

Table 2
Conditional probabilities for the node *PS* of Fig. 1, with parent nodes E_c and W. Node *PS* has two states, A and $\bar{A}$, depending on whether or not the person of interest did or did not wear an article of clothing at the time of the crime which had characteristics of the fibres found on the victim (clothing fibres of PoI similar/dissimilar to those found on victim).

	Node W: At the time of the crime, person of interest			
	wore pullover (W)		did not wear pullover ($\bar{W}$)	
	Node E_c: Known source - pullover of person of interest - was composed of fibres similar (x) or not ($\bar{x}$) to those found on the victim			
	x	$\bar{x}$	x	$\bar{x}$
A	1	0	γ'	γ'
$\bar{A}$	0	1	$(1-\gamma')$	$(1-\gamma')$

which are (x) or are not ($\bar{x}$) similar, in some sense, to fibres found on the victim at the time of the crime. The third root node, W, concerns whether (W) or not ($\bar{W}$) the PoI wore the article in question at the time of the crime. The second and third root nodes are parent nodes for a node *PS* that is not directly observed. Node *PS* has two states, A and $\bar{A}$, depending on whether or not the person of interest did or did not wear an article of clothing at the time of the crime which had characteristics of the fibres found on the victim (clothing fibres of PoI similar/dissimilar to those found on victim). The states of *PS* and the associated conditional probabilities are given in Table 2.

Consider node *PS*. If state W is true, then the characteristics of the control source E_c correspond to that of the PoI's source (state A): $\Pr(A|E_c = x, W) = 1$. If state $\bar{W}$ is true, then the probability that the PoI's source is of type x is given by the probability that the pullover worn by the PoI, different from the control source, would be of type x. This probability can be represented by the term γ'. This probability is considered to be different from the general population proportion of the fibres characteristic described by x. The reason is that there may be information that indicates that the collection of textile habits of the suspect is not representative of the general population proportions of the various fibres types.

There are four other nodes in Fig. 1 which complete the factors of which account has to be taken in the evaluation of evidence.

C: This node is a child of the proposition node H and a parent node for the transfer node T. It describes all the three possibilities for the occurrence of a transfer.
T: This node is a child of the proposition node H and the contact node C and a parent node for the recovered evidence E_r. It notes whether or not the external fibres on the victim were transferred from the PoI or from some third party or through secondary transfer from the PoI via a third party.
E_r: This node is a child of the proposition node H, the transfer node T and the unobserved node that indicates whether or not the PoI wore the pullover (associated with the crime because of the similarity of its fibres to those found on the victim) at the time of the crime. It notes whether (x) or not ($\bar{x}$) the fibres on the victim's clothing are similar to those of the PoI's pullover. It is a parent node of the evidential report RE_r.
RE_r: This node represents the report of the recovered evidence on the victim and is a child of E_r and is a further extension of the context of (17). It allows for the possibilities of false positives and false negatives in the statement of similarity or otherwise in the characteristics of fibres found on the victim's clothing and the clothing of the PoI.

Table 3
Probabilities for the nodes of Fig. 1 as described in Table 1.

Node	Parent nodes or root node	Description (probability of)
C	H	c, d, t_a, t'_a, t_s, t'_s
E_r	H, T	Recovered evidence (x, $\bar{x}$) given transfer (T) and activity proposition (H), including values for b_0 and b_1
E_c	Root node	Probability of characteristics x, $\bar{x}$
H	Root node	$\Pr(H_p) = \Pr(H_d) = 0.5$
PS	E_c, W	$\Pr(PS|E_c, W)$, $\Pr(PS|E_c, \bar{W})$
RE_r	E_r	$\Pr(RE_r|E_r = x) = a$, $\Pr(RE_r|E_r = \bar{x}) = e$
T	H	f, t_1, t'_1, t''_1
W	Root node	w

The values for the prior probabilities for the activity proposition node H are not important. The ratio of posterior odds to prior odds is the Bayes' factor for the evaluation of the evidence. It is because the Bayes' factor is the ratio of odds that the initial values of the prior probabilities are not important. Any change in the prior odds feeds through to the posterior odds and is cancelled out when the ratio is taken. It is suggested the values for H_p and H_d be chosen as equal with value 0.5. The prior odds in favour of H_p are then 1. Probabilities are then assigned to the entries in the other probability tables according to the circumstances of the case. The network is then activated and the change in the probabilities for H to give posterior probabilities and posterior odds are noted.

This graphical structure describes a scenario involving a transfer from clothing of a PoI to a victim, but it can be adapted to describe transfer in the other direction, from victim to the PoI. By adapting the names of the nodes and the probabilities in the conditional probability tables, the structure of the Bayes' net may be used to deal with cases involving transfer from the victim. The problem of cross-transfer for extrinsic evidence is beyond the scope of this paper.

The conditional probabilities required to be assigned for Fig. 1 are given in Table 3.

7. Some illustrative examples of the generalised Bayes' factor

It is of interest to study the impact on the value of the evidence of different events that influence the calculation of the Bayes' factor. Such events include, for example, the uncertainty associated with the origin of the source and with the different transfer mechanisms. Consider three narratives for the circumstances of the activity. There are sixteen events for which probabilities are required as shown in Table 4. The context is the transfer of fibres to the victim from the PoI or from a third party, either directly or as a secondary transfer from the PoI. The propositions of interest are the PoI hit the victim (H_p) and the PoI did not hit the victim (H_d), respectively.

7.1. Narrative 1

The PoI and victim are close friends or close relations who are often in contact with welcoming and farewell hugs for example. They have similar outdoor lifestyles with little opportunity for a third party or secondary transfer of fibres. The activity of hitting was witnessed after a sports event one afternoon. The victim has a very high standard of cleanliness so there is a high probability of no transfer of background fibres. Probability values for the events of interest are listed in Table 5. The Bayes' factor, quantified by using (27) or through the Bayes network depicted in Fig. 1, equals 5 (using

Table 4
Definitions of probabilities and events associated with the generalisation of the Bayes' factor (27).

Prob	Event	Prob	Event
a	Report of recovered evidence of similarity given similarity of fibres	e	Report of recovered evidence of similarity given dissimilarity of fibres
b_0	No transfer of background fibres	w	Pullover worn at time of activity
t_p	Direct transfer of fibres $\|H_p$	t_d	Direct transfer of fibres $\|H_d$
c	Contact of clothing $\|H_p$	d	Contact of clothing $\|H_d$
t_1	Direct transfer of fibres $\|$ contact,H_p	f	Direct transfer of fibres $\|$ contact,H_d
t'_1	Third party transfer $\|H_p$, H_d	t''_1	Secondary transfer of fibres $\|H_p$, H_d
t_a	Contact, third party clothing and victim clothing $\|H_p$	t'_a	Contact, third party clothing and victim clothing $\|H_d$
t_s	Secondary contact of clothing $\|H_p$	t'_s	Secondary contact of clothing $\|H_d$
γ	Pop. characteristic x	γ'	PoI characteristic x

$\gamma = 0.02$) and 3 (using $\gamma = 0.2$). Assuming no uncertainty on the pullover worn at the time of the activity ($w = 1$), then the Bayes' factor increases to 53 (using $\gamma = 0.02$) and 5 (using $\gamma = 0.2$). The smaller the occurrence of the characteristic of interest in the relevant population, and the greater the probability the pullover was worn at time of the alleged activity, the greater will become the value for the Bayes' factor. The value of the evidence increases by a factor of over 10 from 5 to 53 when $\gamma = 0.02$.

7.2. Narrative 2

The PoI and victim are acquaintances with friends in common through which there may be secondary transfer. There is no evidence of opportunities for direct close contact between the PoI and the victim though there are people known to both and with whom both the PoI and victim may have close contact. The activity of hitting was witnessed to have happened during a social gathering of their common friends at which the PoI and the victim were both present. Assume a moderate standard of cleanliness for the victim so a moderate probability of transfer of background fibres. Probability values for the events of interest are listed in Table 6.

The Bayes' factor slightly increases by changing the values for the random occurrence of the characteristic of interest (x); the Bayes' factor increases from 2 to 4 (using $\gamma = 0.2$ and $\gamma = 0.02$, respectively). A greater change in the value of the evidence is obtained by fixing $w = 1$, meaning no uncertainty on the source at time of activity. The Bayes' factor increases from 4 (when $\gamma = 0.2$) to 37 (when $\gamma = 0.02$). The alternative transfer mechanisms, such as the secondary contact of clothing, have a greater impact under the defence hypothesis H_d. For this reason, the Bayes' factor under the narrative 2 scenario is smaller than under narrative 1. Moreover, the impact of the population proportion of the characteristic of interest does not play a fundamental role in the quantification of the value of the evidence; transfer probabilities are more relevant.

7.3. Narrative 3

The PoI and victim are strangers with no family, friends or acquaintances in common. The activity of hitting was witnessed in a street late at night by an acquaintance of the PoI. There was a group conflict with considerable opportunity for third party contact or secondary transfer. The victim has an unstructured lifestyle with much opportunity for transfer to their clothing from the general population. Probability values for the events of interest are listed in Table 7.

Transfer probabilities play an important role in this third narrative. Under the defence proposition H_d, the existence of a mechanism for a secondary or third party transfer is favourable to the defence strategy (i.e. t'_a and t'_s) as measured by its impact on the value of the Bayes' factor. Assuming $\gamma = 0.2$ and $\gamma = 0.02$, respectively, and more uncertainty of the source ($w = 0.6$), the Bayes' factor remains practically unchanged (1.3 and 1.6, respectively). The uncertainty on the source impacts more deeply on the value of the evidence; the value of the BF goes from 2 ($\gamma = 0.2$) to 16 ($\gamma = 0.02$) when $w = 1$.

These three narratives suggest the conclusion that the possible existence of secondary or third party transfer mechanisms, present under both propositions H_p and H_d, plays the main role in the quantification of the value of the evidence. Population proportion estimates play a lesser role in the evaluation of evidence when activity level propositions are under consideration than when source level propositions are considered. Values for the γ parameter are less important in the assessment of the Bayes' factor under activity propositions than under source propositions. Greater importance should be assigned to the uncertainty associated with the source of the transfer material. This importance is illustrated in the three narratives above by the changes in the value of the evidence arising from changes in the probability that the person of interest is associated with the source of the transfer material at the time of the relevant activity. Uncertainty associated with the origin of recovered

Table 5
Definitions and values (Val) of probabilities (Prob) and events associated with narrative 1.

Prob	Event	Val	Prob	Event	Val
a	Report of recovered evidence of similarity given similarity of fibres	0.95	e	Report of recovered evidence of similarity given dissimilarity of fibres	0.01
b_0	No transfer of background fibres	0.8	w	Pullover worn at time of activity	0.8
c	Contact of clothing $\|H_p$	0.7	d	Contact of clothing $\|H_d$	0.5
t_1	Direct transfer of fibres $\|$ contact,H_p	0.9	f	Direct transfer of fibres $\|$ contact,H_d	0.8
t'_1	Third party transfer $\|H_p$, H_d	0.5	t''_1	Secondary transfer of fibres $\|H_p$, H_d	0.4
t_a	Contact, third party clothing and victim clothing $\|H_p$	0.15	t'_a	Contact, third party clothing and victim clothing $\|H_d$	0.25
t_s	Secondary contact of clothing $\|H_p$	0.15	t'_s	Secondary contact of clothing $\|H_d$	0.25
γ	Pop. characteristic x	0.02 0.2	γ'	PoI characteristic x	0.1

Table 6
Definitions and values (Val) of probabilities (Prob) and events associated with narrative 2.

Prob	Event	Val	Prob	Event	Val
a	Report of recovered evidence of similarity given similarity of fibres	0.95	e	Report of recovered evidence of similarity given dissimilarity of fibres	0.01
b_0	Transfer of background fibres	0.5	w	Pullover worn at time of activity	0.8
c	Contact of clothing $\|H_p$	0.5	d	Contact of clothing $\|H_d$	0.1
t_1	Direct transfer of fibres | contact,H_p	0.8	f	Direct transfer of fibres | contact,H_d	0.1
t'_1	Third party transfer $\|H_p$, H_d	0.6	t''_1	Secondary transfer of fibres $\|H_p$, H_d	0.7
t_a	Contact, third party clothing and victim clothing $\|H_p$	0.2	t'_a	Contact, third party clothing and victim clothing $\|H_d$	0.45
t_s	Secondary contact of clothing $\|H_p$	0.3	t'_s	Secondary contact of clothing $\|H_d$	0.45
γ	Pop. characteristic x	0.02 0.2	γ'	PoI characteristic x	0.1

Table 7
Definitions and values (Val) of probabilities (Prob) and events associated with narrative 3.

Prob	Event	Val	Prob	Event	Val
a	Report of recovered evidence of similarity given similarity of fibres	0.95	e	Report of recovered evidence of similarity given dissimilarity of fibres	0.01
b_0	Transfer of background fibres	0.2	w	Pullover worn at time of activity	0.6
c	Contact of clothing $\|H_p$	0.5	d	Contact of clothing $\|H_d$	0.1
t_1	Direct transfer of fibres | contact,H_p	0.8	f	Direct transfer of fibres | contact,H_d	0.05
t'_1	Third party transfer $\|H_p$, H_d	0.7	t''_1	Secondary transfer of fibres $\|H_p$, H_d	0.6
t_a	Contact, third party clothing and victim clothing $\|H_p$	0.1	t'_a	Contact, third party clothing and victim clothing $\|H_d$	0.4
t_s	Secondary contact of clothing $\|H_p$	0.4	t'_s	Secondary contact of clothing $\|H_d$	0.5
γ	Pop. characteristic x	0.02 0.2	γ'	PoI characteristic x	0.1

material is an important consideration in the evaluation of extrinsic evidence under activity level propositions.

8. Conclusion

A generalised Bayes' factor and associated Bayesian network have been developed for the transfer of extrinsic evidence at the activity level. A strategy for the assessment of extrinsic evidence has been developed in stages with progressive increases in the complexity. The development has been illustrated with an example involving fibres from clothing and three narratives have provided numerical examples of possible practical applications.

Previous work on activity level evaluation, mainly for intrinsic evidence with examples from DNA evidence, has been extended to extrinsic evidence. The formal development generalises previous work to consider these extensions culminating in (27). This equation incorporates terms that allow for

- the role of various intermediate propositions such as transfer from a person other than the PoI;
- legitimate transfer from a known source;
- secondary transfer from the PoI via a third party;
- transfer by a third party;
- uncertainty associated with the origin of the source material and
- false positive and true positive associations.

The ideas are illustrated with three narratives involving fibres from clothing. The activity is that of a hit on a person. The hit may have been accidental and so not an offence. The important messages from consideration of the results from the analysis of the narratives can be summarised as follows: .

- the role of the probability γ, information on a population proportion, is minimal;
- the alternative transfer mechanisms under both propositions play the main role in evidence evaluation;
- care should be done in the assessment of probability w, the probability the PoI is associated with the putative source of the transfer material at the time of the relevant activities.

CRediT authorship contribution statement

F. Taroni: Conceptualization, Formal analysis, Writing - original draft. **P. Garbolino**: Validation, Writing. **C. Aitken**: Formal analysis, Writing, Validation, Supervision.

Declaration of Competing Interest

None.

Acknowledgment

The authors thank the reviewers for their careful reading of the manuscript and their helpful comments. The authors also thank the Swiss National Science Foundation for its support through grant n. IZSEZ0-19114.

Appendix A. Definitions of terms for probability.

Table 8.

Table 8
Definitions of some terms for probability for transfer and contact from a person of interest (PoI) to the victim.

Term	Definition - probability of
a	reported evidence RE_r given recovered evidence E_r
b_0	no groups of fibres transferred from general population to the victim
b_1	one group of fibres transferred from general population to the victim, corresponding to fibres associated with the PoI
c	clothing of the victim and clothing of the PoI have been in contact, given the PoI hit the victim
d	clothing of the victim and clothing of the PoI have been in contact given the PoI did not hit the victim
e	reported evidence RE_r given recovered evidence $\bar{E}_r$
f	transfer of a group of fibres from the PoI to the victim given a contact of clothing and PoI did not hit the victim
γ	(population proportion) for characteristics (x) of fibres of clothing
γ'	the PoI's source of fibres being of type x; considered to be different from the general population proportion of the fibres characteristic described by x
ρ	control material worn by PoI during the activity, given evidence of control material
t_p	transfer of a group of fibres from the PoI to the victim, given the PoI hit the victim
t_d	transfer of a group of fibres from the PoI to the victim, given the PoI did not hit the victim
t_1	transfer of a group of fibres from the PoI to the victim, given the PoI hit the victim and contact between PoI and victim
t'_1	transfer of a group of fibres from a third party to the victim, given there was contact between clothing of the victim and the clothing of the third party, whether or not the PoI hit the victim.
t''_1	transfer of group of fibres to the victim from the PoI, given there was secondary contact from clothing of the PoI via clothing of a third party to clothing of the victim, whether or not the PoI hit the victim.
t_a	contact between clothing of the victim and clothing of a third party given the PoI did hit the victim.
t'_a	contact between clothing of the victim and clothing of a third party given the PoI did not hit the victim.
t_s	occurrence of secondary contact from the clothing of the PoI via a third party to the clothing of the victim given the PoI hit the victim
t'_s	occurrence of secondary contact from the clothing of the PoI via a third party to the clothing of the victim given the PoI did not hit the victim
w	the PoI wore the pullover at the time of the relevant activity.
z	transfer of recovered fibres to the victim given the PoI hit the victim
$\bar{z}$	no transfer of recovered fibres to the victim given the PoI hit the victim
z'	transfer of recovered fibres to the victim given the PoI did not hit the victim
$\bar{z}'$	no transfer of recovered fibres to the victim given the PoI did not hit the victim

References

[1] C.G.G. Aitken, F. Taroni, P. Garbolino, A graphical model for the evaluation of cross-transfer evidence in DNA profiles, Theor. Popul. Biol. 63 (2003) 179–190.
[2] C.G.G. Aitken, F. Taroni, S. Bozza, Statistics and the Evaluation of Evidence for Forensic Scientists, third ed, John Wiley & Sons, Chichester, 2020.
[3] C. Champod, DNA transfer: informed judgment or mere guesswork? Front. Genet. 4 (2013) 300.
[4] R. Cook, I.W. Evett, J. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, Sci. Justice 38 (1998) 231–240.
[5] R.G. Cowell, A.P. Dawid, S.L. Lauritzen, D.J. Spiegelhalter, Probabilistic Networks and Expert Systems, Springer, New York, 1999.
[6] European Network of Forensic Science Institutes, Guideline for evaluative reporting in forensic science, Bruxelles, 2015.
[7] I.W. Evett, A quantitative theory for interpreting transfer evidence in criminal cases, Appl. Stat. 33 (1984) 25–32.

[8] I.W. Evett, Establishing the evidential value of a small quantity of material found at a crime scene, J. Forensic Sci. Soc. 33 (1993) 83–86.
[9] P. Garbolino, F. Taroni, Evaluation of scientific evidence using Bayesian networks, Forensic Sci. Int. 125 (2002) 149–155.
[10] P. Gill, T.N. Hicks, J. 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) Article 102186 (in press).
[11] T.N. Hicks, J.S. Buckleton, J.-A. Bright, D. Taylor, A framework for interpreting evidence, in: JS Buckleton, J.-A. Bright, D. Taylor (Eds.), Forensic DNA Evidence Interpretation, second ed, CRC Press, Boca Raton, 2016, pp. 37–86.
[12] D.A. Schum, Evidential Foundations of Probabilistic Reasoning, John Wiley & Sons, Inc., New York, 1994.
[13] F. Taroni, A. Biedermann, S. Bozza, J. Comte, P. Garbolino, Uncertainty about the true source - a note on the likelihood ratio at the activity level, Forensic Sci. Int. 220 (2012) 173–179.
[14] F. Taroni, A. Biedermann, J. Vuille, N. Morling, Whose DNA is this? This is not the relevant question (a note for forensic scientists), Forensic Sci. Int. Genet. 7 (2013) 467–470.
[15] F. Taroni, A. Biedermann, S. Bozza, P. Garbolino, C.G.G. Aitken, Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science, second ed, John Wiley &Sons, Chichester, 2014.
[16] F. Taroni, P. Juchli, C.G.G. Aitken, A probabilistic account of the concept of cross-transfer and inferential interactions for trace materials, Law Probab. Risk (2020) (submitted).
[17] D. Taylor, A. Biedermann, L. Samie, K.-M. Pun, T. Hicks, C. Champod, Helping to distinguish primary from secondary transfer events for trace DNA, Forensic Sci. Int. Genet. 28 (2017) 155–177.
[18] D. Taylor, B. Kokshoorn, A. Biedermann, Evaluation of forensic genetics findings given activity level propositions: a review, Forensic Sci. Int. Genet. 36 (2018) 34–49.
[19] D. Taylor, L. Samie, C. Champod, Using Bayesian networks to track DNA movement through complex transfer scenarios, Forensic Sci. Int. Genet. 42 (2019) 69–80.
[20] D. Taylor, B. Kokshoorn, T. Hicks, Structuring cases into propositions, assumptions, and undisputed case information, Forensic Sci. Int. Genet. 44 (2020) Article 102199 (in press).
[21] W.C. Thompson, F. Taroni, C.G.G. Aitken, How the probability of a false positive affects the value of DNA evidence, J. Forensic Sci. 48 (2003) 47–54.