Forensic Science International 361 (2024) 112086

Contents lists available at ScienceDirect

Forensic Science International

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

From frequented environments to the crime scene: Evaluating findings of fibre comparisons in complex transfer scenarios.

Yu Chen Lim-Hitchings [*], Franco Taroni, Geneviève Massonnet

University of Lausanne, Ecole des sciences criminelles, Batochime, Lausanne 1015, Switzerland

ARTICLE INFO

Keywords:
Trace evidence
Interpretation
Evaluative reporting
Secondary transfer
Background
Foreign fibre groups
Likelihood ratio

ABSTRACT

The evaluation of the results from a fibre comparison given activity level propositions is well established when considering only a single group of potential primary transfers. However, secondary transfers are less prevalent in the literature despite their potential value, especially in cases where the primary transfers are not sufficiently informative. In particular, one can consider the fibres from frequented environments of the person of interest (POI) identified in a struggle. If the POI did struggle with the complainant, these fibres can potentially be recovered in small quantities on the surface of the complainant as a result of secondary or higher order transfers. Therefore, these fibres may provide useful information that can resolve competing propositions involving struggles, as well as forensic intelligence in the form of linkages or investigative leads. If a non-differentiation is indeed found between recovered fibres and fibres from the frequented environments of the POI, these results need to be properly interpreted. In this paper, a model, based on an object oriented Bayesian network (OOBN), for evaluating such findings along with its implementation is proposed. Using available data from the literature and other sources, the model was then used to assess a few hypothetical scenarios involving secondary transfers. The results provided useful insights into secondary transfer that help to validate the model and demonstrate the potential utility that can be gained by considering transfers beyond the primary order. Moreover, these results can be used to help guide future research by identifying gaps in the literature. Finally, the direct application to a case study was conducted to demonstrate the practical aspects of such a model.

1. Introduction

One of the fundamental roles of forensic science is to conduct case-based, scientific study of traces left behind. Following which, it is tasked to help interpret the findings to aid decision makers (judges, jury, etc.) in the understanding of ambiguous events. These functions are as defined in the recent Sydney declaration [1] which has laid out seven principles as the guiding foundation on fulfilling these duties. In accordance with principle 1, these traces are a vestige of past events, and understanding their formation is fundamental to forensic science. Aside from their detection, recovery, and examination, almost all these principles stress the importance of the *interpretation* of these traces within the context of a case. Particularly of note, is the repeated emphasis on the need for knowledge (principle 3 and 5) and a sound framework (principle 2, 3, and 4) that is case specific (principle 7). The reason for this emphasis is due to the ever-present uncertainty, mentioned in principle 5, of the entire process, from the generation of the traces to their recovery and analysis. As such, to be truly informative, these uncertainties need to be quantified and understood, which requires a robust framework and application of available data.

These concepts of proper evaluation of analytical results and their importance, are detailed in the principles of interpretation [2–4]. In essence, the probabilities of observing the analytical results given appropriate propositions, formulated based on relevant disputed issues, should be assessed. These propositions should also be at a suitable level in the hierarchy of propositions to avoid ambiguity. In most cases the activity level, where some form of contact and its intensity is specified, is more suitable than the source level, where only the attribution of the origin of the trace is of concern [2–4]. The resulting probabilities can then be compared in the form of a likelihood ratio (LR) to see if the results support one proposition over another and by how much [4–7]. During case pre-assessment, such a step permits a transparent demonstration of potential value that future work may bring, thus allowing an informed decision on the allocation of resources [4]. If the results were to be brought to a trial, the evaluation gives the decision maker (be it judge, jury, or other) an appreciation of the bearing that it brings

---

[*] Corresponding author.
E-mail address: yuchen.lim-hitchings@unil.ch (Y.C. Lim-Hitchings).

https://doi.org/10.1016/j.forsciint.2024.112086
Received 26 January 2024; Received in revised form 13 May 2024; Accepted 3 June 2024

towards their decisions [5–7].

In terms of fibre analysis, the employment of such an evaluative framework is well documented for straightforward situations [8,9]. Namely, findings of a single group of fibres found to be non-differentiated to textiles that have had a direct contact specified in one or more propositions. In these scenarios, explanations for the result include that of primary transfer, whose mechanics have been substantially and appropriately studied [8], as well as the background, which has also been investigated considerably [10].

Consider the scenario where a complainant was purportedly assaulted violently, and a person of interest (POI) has been identified. Assuming that the existence of an assault is not under dispute, this would lead to the following two activity level propositions that will be referenced throughout the paper:

- Proposition 1, H_1: The POI was in a struggle with the complainant.
- Proposition 2, H_2: Someone other than the person of interest, known as the alternate offender (AO), was in a struggle with the complainant.

A typical manner of approaching such an incident is to consider the fibres that make up the clothing worn by the POI during the assault as target fibres. Fibres recovered on the complainant's clothing or body which cannot be differentiated from these target fibres after an appropriate sequence of analysis would be prioritised. The quantities of these non-differentiated fibres recovered then form the analytical findings, *E*, to be evaluated. In general, it can be described in the Bayesian network show in Fig. 1 [11].

The use of Bayesian networks for the evaluation of forensic evidence has already been extensively covered and affords many advantages [11, 12]. In this case, it is easy to see that the findings can be explained by either the transfer and subsequent persistence of these fibres during one of the propositions or from events completely unrelated that are globally considered as part of the background. A LR can then be obtained from such a model based on the assigned conditional probability distribution of each node and the instantiation of the findings node, E with the obtained analytical results.

However, target fibres are not only limited to ones that compose the clothing worn by the POI during the assault. One can also consider fibres of textiles from frequented environments of the POI, such as their home or place of work. Under H_1, these fibres could have been first primarily transferred onto the suspect's person during their day-to-day activities, then subsequently secondarily transferred during the assault. Although rarely empirically demonstrated, this population of fibres from a specific environment are often considered to be highly discriminating given the wide variety of fibres that can be found in different surroundings based on their context [13]. The consideration of these additional groups of fibres could offer a few advantages:

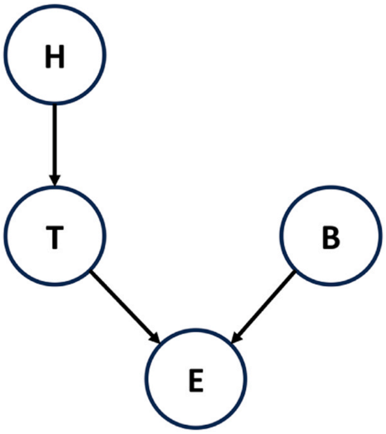

Fig. 1. General Bayesian network for the evaluation of fibre comparison results given activity level proposition. Where node H represents the proposition, node T represents the number of non-differentiated fibres transferred, persisted, and recovered given the proposition, node B represents the number of non-differentiated fibres present in the background, and node E represents the analytical findings.

- The clothing worn by the POI at the time of the assault may be unknown or unavailable. Hence, textiles recovered from the POI's environment offers alternative reference materials. Moreover, these textiles reside long term in these settings frequented by the POI where they potentially shed fibres. Therefore, it is likely that they would be consistently found on the POI, providing a stable pool of fibres for secondary transfer.
- Even if the clothing worn by the POI is known and available, the potential value from its comparison to recovered fibres may be limited. This is especially the case if the clothing is composed of common fibres such as white or denim cotton. Alternatively, the reference materials in question may be of low shedding [14,15], hence are expected to transfer only a limited number of fibres under H_1. The population of fibres from frequented environments can therefore provide additional avenues which may contribute value to the discrimination between the two given propositions.
- In addition to the case at hand, the consideration of these additional fibres also provides value in terms of forensic intelligence and investigative leads [13,16]. Such fibres can provide the means to create linkages between cases [17,18]. An offender may not wear the same article of clothing at each incident; however, their frequented environment likely presents certain constants in terms of textiles. Such target fibres recovered on separate complainants may be used to establish connections.

However, considering such target fibres requires substantial resources in terms of time and effort. From its detection to analytical comparison, the work can take up to thousands of hours due to the high volume of fibres to analyse [17,19]. Even though modern advances in technology such as the European SHUTTLE project[1] could in theory reduce the investment of resources, such automated initiatives may also create an unprecedented higher volume of traces that need to be treated and subsequently interpreted. A cohesive model for evaluating these findings against the given set of propositions is therefore essential. Such a model would initially be able to provide means to prioritise and justify the analytical work to be conducted with a case pre-assessment. Subsequently, its presentation in court settings allows a clear representation of its value to aid a decision point. Due to its complex nature, a suitable model would not be as straightforward as the simple framework presented previously. The challenges can be summarised as follows:

- When considering textiles from frequented environments, it is often the case where multiple target fibre groups are identified. An evaluative model for these findings therefore needs a coherent way to combine the findings of multiple groups. This adds a layer of complexity where the intricate dependencies of different parameters within need to be carefully considered.
- The number of fibres expected to secondarily transfer is low [17, 19–21], potentially overlapping with sizes of background fibre groups [10]. Hence, an evaluation of these findings requires the combined consideration of background and transfer. In the generic model presented previously, the findings are often considered to have come purely from transfer or background, with their combined contribution treated as negligible. This may be a fine approximation for evaluations concerning primary transfers but does not hold up when secondary transfers are considered.

---

[1] https://www.shuttle-pcp.eu/

- Finally, there is little data available for secondary transfer and background foreign fibre groups. Any model that can be designed is hence limited to uncertainty in these parameters.

These points, coupled with the tedious nature of fibre recovery and comparison, suggest that the execution of such work may not yet be empirically justified in terms of its value. Nonetheless the exploitation of such traces is not unheard of in practice [17,19,22] and would benefit from a more comprehensive interpretation framework to understand its worth. Hence, before going further, it would be wise to assess where we are and subsequently where we should go from here. In line with the principles stated in the Sydney declaration [1], and the issues raised by authors concerning trace evidence analysis [23–25], this paper aims to answer these questions by first proposing an evaluative model for such findings. Such a model should be flexible enough to be generalised to specific case-work scenarios and adapt based on varying contextual information and scientific findings. In order to achieve these objectives, the use of object-oriented Bayesian networks (OOBNs) was employed. Although other models that explore secondary transfer of fibres exist, none, as far as the authors are aware, appropriately address the combination of multiple groups or are suited to the scenario at hand. They either discuss secondary transfer as a defence strategy [26] or other propositions of interest [27]. Second, the current state of knowledge based on published literature was evaluated and implemented into the model. Finally, application of this model was carried out along with sensitivity analysis for both generic findings for a hypothetical scenario and a case study courtesy of the Institut National de Criminalistique et de Criminologie (INCC) in Belgium [19]. The results allow the appreciation of the value to be gained when considering these fibres from frequented environments as target fibre groups. This gives a global understanding of the implications based on current knowledge, along with gaps in the data that need to be filled.

2. Materials and methods

2.1. Proposed framework

A proposed evaluative model was built using the r programming language [28] using the packages bnlearn and Rgraphviz [29], with the computation handled by gRain [30]. The code for all functions created for this paper can be obtained by contacting the authors. To facilitate the process, the integrated development environment (IDE) Rstudio from Posit [31] was used. The model takes up the form of an OOBN based on the generic model detailed in Fig. 1. It features three classes: transfer, background, and evidence/findings that each handle an independent part of the modelling. Further details of the model will be presented in the *results and discussion* section below. The model permits the instantiation of findings from comparative fibre analysis and produces LRs based on them with respect to the two propositions involving contact, such as the ones detailed previously. The model was made to be sufficiently flexible such that it may incorporate varying findings and probability distributions depending on the context. As a function, it requires the input of the number of target fibre groups under consideration, *g*, as well as the conditional probability distributions of each relevant parameter. Node states were defined as categorical and take on a multinomial distribution. The overall structure of the model is shown in Fig. 2 where each rectangle represents a class. The respective classes have their own encapsulated architecture comprised of nodes joined by arcs that represent their dependencies. Each node has its own conditional probability distribution specified by their relevant parameters which are summarised in Table 1.

2.2. Current state of knowledge

Literature and data on fibres relevant to the parameters required for the proposed model (see Table 1) were obtained. This was done by searching for the keywords such as “transfer”, “secondary transfer”, “background”, “occurrence”, “frequency”, “population studies”, “colour

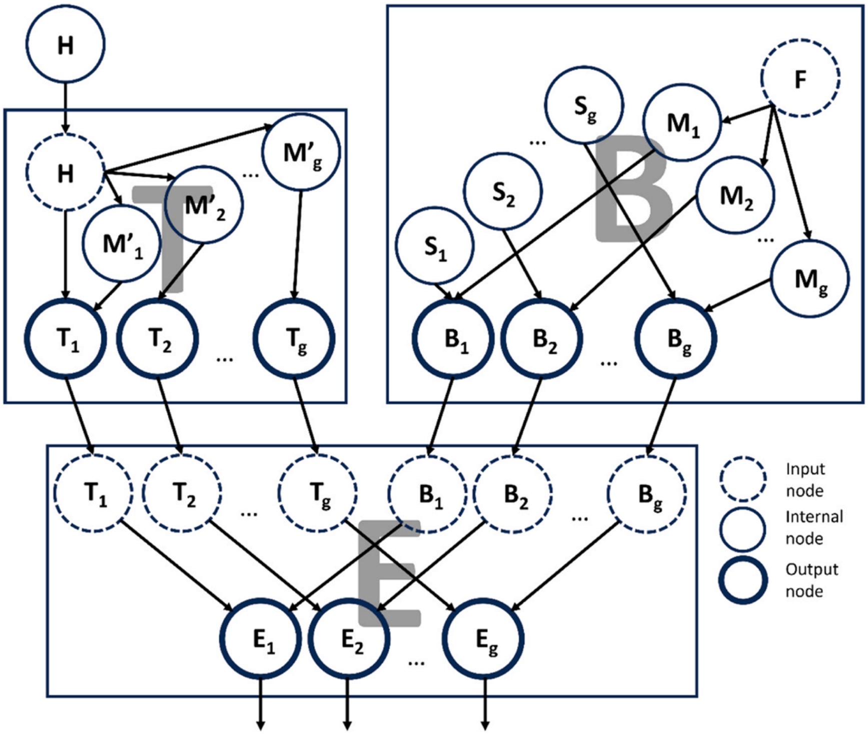

Fig. 2. Overall OOBN structure of proposed model for g target fibre groups, each rectangular node represents an entire class structure with its own independent sub-network structure. The classes are labelled in large block letters, T for transfer, B for background and E for evidence/findings.

Table 1
Node specification of each class in the proposed OOBN model, nodes with a subscript i indicate that a number equal to the number of target fibre groups of such nodes exist in the class.

Class	Node	Description	States	Parameters
E	E_i	Categorical number of fibres recovered that cannot be differentiated from the i^{th} target fibre group.	None, Small (1–5), Medium (6–50), Large (>50)	-
	B_i	See below		
	T_i			
B	F	Categorical number of foreign fibre groups (FFGs).	0, 1–10, 11–20, 21–30, 30–40, >40	$f = \{f_0, f_{1-10}, \dots, f_{>40}\}$
	M_i	One of the FFGs cannot be differentiated from the i^{th} target fibre group.	Cannot be differentiated, Can be differentiated	γ
	S_i	Categorical number of fibres in a given FFG.	Small (0–5), Medium (6–50), Large (>50)	$s = \{s_{small}, s_{medium}, s_{large}\}$
	B_i	Categorical number of fibres that cannot be differentiated from the i^{th} target fibre group in the background.	None, Small (0–5), Medium (6–50), Large (>50)	-
T	H	Propositions.	H_1, H_2	-
	M'_i	One fibre group on the AO cannot be differentiated from the i^{th} target fibre group.	Cannot be differentiated, Can be differentiated	γ'
	T_i	Categorical number of fibres that cannot be differentiated from the i^{th} target fibre group that were transferred (persisted and recovered).	None, Small (0–5), Medium (6–50), Large (>50)	$t = \{t_{none}, t_{small}, t_{medium}, t_{large}\}$ $t' = \{t_{none}, t'_{small}, t'_{medium}, t'_{large}\}$

block studies" along with the term "fibre". The information obtained was used to assign the probability distributions in the model with frequency data or using Dirichlet distributions where maximum likelihood estimates (MLE) can be obtained. These are summarised in Table 2. Two distributions were assigned for the transfer parameter, *t*, one for primary transfer, and one for secondary. Similarly, two distributions for the occurrence parameters were assigned, one for a generic common fibre and another for a generic rare fibre. The specific details for these assignments will be presented in the results and discussion section below.

2.3. Model implementation and testing

The probability assignments were integrated into the proposed model which were then used to evaluate both hypothetical findings and findings that pertain to a case study described by De Wael [19]. The former was used to judge the current state of knowledge and evaluate the potential value of such findings along with their limitations. The latter was conducted to deploy the model in a practical case-based scenario for demonstration. In both cases, the propositions were formulated as described in the introduction. The findings in both cases were instantiated in the evidence class of the model and a likelihood ratio (LR) was numerically obtained from the propositions node based on the ratio between posterior and prior odds [5].

Table 2
Parameter assignments for the model based on current state of knowledge.

Parameter	Assigned probability distribution
f	$f \sim Dir(1, 2, 4, 8, 16, 32)$
γ	$\gamma_{common} = 0.014$ $\gamma_{rare} = 0.00171$
s	$s \sim Dir(38, 10, 2)$
$t = t'$ (primary transfer)	$t_{none} = 0.001$ $t_{small} = 0.009$ $t_{medium} = 0.05$ $t_{large} = 0.94$
$t = t'$ (secondary transfer)	$t \sim Dir(13, 12, 4, 1)$
γ'	$\gamma'_{common} = 0.00566$ $\gamma'_{rare} = 0.000855$

2.3.1. Hypothetical findings

Two broad conditions were tested, in the first, only fibres corresponding to textiles from frequented environments were considered as target fibres. This condition represents cases where the clothing worn by the POI during the event were unavailable. In the second condition clothing worn by the POI during the struggle was also considered. These fibres were prescribed to be of common occurrence and evaluated under primary transfer. This condition represents the presence of typical findings explainable by primary transfer but are of low value. Concerning the target fibres which correspond to the textiles in the POI's environment, these were evaluated under secondary transfers. Between one and four such fibre groups of every possible combination of occurrence (common or rare) and sizes (small or medium) were recovered. These hypothetical findings are summarised in Table 3. The parameters were populated based on the assigned values shown in Table 2. In order to investigate the effect of the uncertainty on the resulting LRs, parameters assigned as Dirichlet distributions were simulated by drawing 1000 times from each respective distribution to obtain resulting LRs. The likelihood ratios were then converted to a log_{10} base for easy interpretation along with the use of a verbal scale proposed by Marquis et al. [32].

2.3.2. Sensitivity analysis

The evaluation was then repeated for the findings of four rare fibre groups of medium sized under condition A while respectively setting

Table 3
Various combinations of findings tested under each condition.

Condition	Number of target fibre groups, g	Findings evaluated (groups found to be non-differentiated to reference material)
A	4	- 1 rare (1 R) - 2 rare (2 R) - 3 rare (3 R) - 4 rare (4 R) - 1 common (1 C) - 2 common (2 C) - 3 common (3 C) - 4 common (4 C) - 1 rare + 1 common (1R1C) - 1 rare + 2 common (1R2C) - 1 rare + 3 common (1R3C) - 2 rare + 2 common (2R2C) - 3 rare + 1 common (3R1C) Evaluated as small and medium groups respectively as secondary transfers
B	5	1 large common group evaluated for primary transfer + - 1 rare (1 R) - 2 rare (2 R) - 3 rare (3 R) - 4 rare (4 R) - 1 common (1 C) - 2 common (2 C) - 3 common (3 C) - 4 common (4 C) - 1 rare + 1 common (1R1C) - 1 rare + 2 common (1R2C) - 1 rare + 3 common (1R3C) - 2 rare + 2 common (2R2C) - 3 rare + 1 common (3R1C) Evaluated as small and medium groups respectively as secondary transfers.

two of the three parameters (f, s, and t for secondary transfer) to their MLE and keeping the third as a distribution. This was carried out once for each of the parameters. The goal of this step was to evaluate the effect of each of these parameters on the results of the model and subsequently, the consequence of the uncertainties on these parameters.

2.3.3. Case study

Finally, the model was applied to findings of a case study described by De Wael [19] where two victims went missing from a fair in Belgium only to be found dead in a drainage canal 18 days later, immersed in running water. A POI was identified and had come forward at that time. Reference materials comprised of the clothing worn on the day of the disappearance were collected. Additionally, a duvet cover from the POI's partner's apartment, where they had spent significant time prior to the incident, was also seized. Experts at the INCC recovered fibres from the garments of the victims that corresponded to both the POI's clothing and the duvet cover. A summary of the recovered fibres on each of the victims, along with the assigned occurrence value from literature can be found in Table 4. Note that these occurrence values were assigned based on available data or from case information which were communicated from the experts.

Since the victims were submerged in running water for 18 days where heavy rains increased the flow rates on certain days, heavy loss of fibres was expected [38]. A loss of 85% was estimated and a negative binomial model was used to specify a probability distribution of the number of fibres before the deposition of the bodies. The two propositions defined at the start of this paper served as the ones these findings were evaluated against. The fibres which correspond to the POI's clothing were considered under primary transfer and those corresponding to the duvet cover under secondary transfer. Furthermore, since this was for a concrete case, a single LR should be obtained rather than a reported range [39]. Thus, a MLE was used for all parameters that were previously expressed as distributions. The findings were then evaluated first considering primary transfers only and subsequently all transfers to observe the added value of the latter.

3. Results and Discussion

3.1. Proposed framework

When taking fibres from the frequented environments of the POI as comparative reference material, analytical findings would take the form of at least one, but more likely multiple non-differentiated groups of varying sizes on the substrate of interest. These additional groups pose a challenge since the value from each group in terms of their LR cannot be assumed to be independent. As such, they cannot simply be multiplied together to obtain their combined value. Furthermore, under H_1, the secondary transfer of the fibres must be considered which adds more intricacies compared to a primary transfer. These added complexities render the general evaluative framework shown in Fig. 1 insufficient. Nonetheless, it lays the foundation for evaluating more complex scenarios such as the one of interest laid out in this paper.

Even though the reference materials from the POI's environment is not specified to have direct contact in either proposition of interest, each recovered non-differentiated group still has exhaustively two independent mechanisms that explain their presence. Either they were a direct consequence of the struggle with the POI or AO (transfer), or they were a consequence of an unrelated event (background). This is no different than the generic model first laid out in Fig. 1. Evaluations on such cases will still have to consider these three components of transfer, background, and findings. Since the models of each of these three parts are often repeated, this lends itself to be expressed with an OOBN architecture which has been previously explored in the DNA field for similar considerations [40,41].

An OOBN groups independent sub-models into objects known as classes for computational and visualisation advantages [40,42,43]. Each independent class can be reused and adapted to different scenarios. In the case of the proposed model, each of the three mentioned components assume a class and adapt to changes in the number of non-differentiated groups in the findings as well as the case circumstances that specify the different parameters. These groups would share a common structure based on their dependencies within their respective class. The three classes can then be combined based on their probabilistic dependencies to form the overall OOBN. The advantages of employing such a tool can be summarized as follows:

- Computationally efficient [11,42–44]
- Visually intuitive [40,41]
- Ease of construction [40,41,43,44]
- Amenability to generalisation [42,43]

This led to the development of the proposed model detailed in Fig. 2 and Table 1. When simplified to a single target fibre, it can be shown that the network reduces to the generic model shown in Fig. 1 and is mathematically equivalent to Eq. 1 described in previous works [11]. This serves as a demonstration of the model's validity. A formulaic derivation of the same model is also available upon request from the authors. The architecture of each class will be detailed presently in the following sections. In general, each class can adapt to the number of non-differentiated groups under consideration and different assignments to their respective parameters.

3.1.1. Evidence/findings class

Beginning with the evidence/findings class whose structure is shown in Fig. 2 in the rectangle labelled "E". Its role is to specify the probability distribution of obtaining the analytical findings, it must therefore consider all the possible combinations of recovered non-differentiated fibre groups. These are represented by distinct evidence nodes, E_i, where $1 < i < g$ and g is the total number of groups under consideration. The categorical intervals were chosen to maximise the discrimination

Table 4

Findings and assigned occurrences for the case study. Note that the citation next to occurrence values represents data used to make the assignment, if none exist, they were assigned based on personal communication with experts who worked on the case.

Target fibre group	Reference material (transfer considered)	Occurrence	Number recovered on Victim 1	Number recovered on Victim 2
Blue denim	POI's jeans (primary)	0.3 (extremely common)	566	1045
Blue polyester (with defect)	POI's jeans (primary)	0.0002 [33]	11	22
Grey polyester	POI's underwear (primary)	0.005 [34,35]	1	0
Purple polyamide	POI's shirt (primary)	0.005	0	6
Dark blue polyamide	POI's shirt (primary)	0.005	0	3
Blue-grey polyamide	POI's shirt (primary)	0.005	0	4
Grey-blue viscose	POI's shirt (primary)	0.005	0	2
Red cotton	Duvet cover (secondary)	0.015 [36,37]	2	2
Red Polyester	Duvet cover (secondary)	0.005	0	1
Black cotton	Duvet cover (secondary)	0.014 [36,37]	0	3

between transfers given the respective propositions and background levels. These intervals are flexible and can be adapted to suit varying case needs. As previously expressed, the state of such a node is dependent on the transfer and background of the specific fibre group in question. Hence, for each E_i node, there are two corresponding parent nodes of B_i and T_i that take up the same categorical states. These parents are input nodes and allow for the transfer of information from the two other classes into the evidence class. Each E_i node then specifies the sum of the probability distribution of the two parent nodes.

3.1.2. Background class

The general structure of a background class is shown in Fig. 2 in the rectangle labelled “B”. This class outputs the probability distribution for the categorical quantity of each non-differentiated target fibre group present which are unrelated to the propositions. As such, it has nodes B_i that can take the same four categorical states as the E_i nodes. For any target fibre group to exist in the background, they need to correspond to the analytical characteristics measured. This event is taken into account in a parent “match” node, M_i, for each respective target fibre group. Thes nodes can take one of two states of non-differentiation, M, or differentiation, $\overline{M}$, from the target fibre group based on the analytical sequence used to carry out the comparison. The probabilities of these events are dependent on the total number of foreign fibre groups (FFGs) [27,45]. The more FFGs are present in the background, the more likely that one of them would correspond with the target fibre group. As a result, all M_i nodes are parented by one common F node with categorical states that express the intervals of number of FFGs present in the background. This probability distribution is specified by the parameter f as shown in Table 1.

Concerning the conditional probability distribution of the respective M_i nodes, they would have to consider the occurrence of the measured characteristics of the target fibre group, γ_i. The distribution is also conditioned upon the number of FFGs present. This situation is comparable to the two-trace problem described by Evett [46] and further clarified by Aitken [5]. In terms of fibres, such a calculation has been explored by Champod [9] and Palmer [27]. In general, a desirable property of this derivation is that the probability should increase as more groups are present. Following this, as the number of groups approaches infinity, the probability of one group corresponding should approach 1. This is because intuitively, an infinite number of groups would be sure to encompass the group of interest which is a member of a finite set. Given these criteria, the following derivation of this probability distribution is proposed.

First, the mutually exclusive and exhaustive means in which the analytical characteristics can correspond to one of the m groups should be considered. If the FFGs are indexed from 1 to m, the probability of each group corresponding to the reference material can be enumerated as shown in Table 5. Concretely, for the correspondence of a given FFG, k, where $1 < k < m$, the probability of non-correspondence in preceding groups from 1 to $k-1$ must be included. This is the complement probability $(1-\gamma_i)^{k-1}$ described by Palmer [27]. Following which, the probability of correspondence at the k^{th} FFG is the probability of occurrence itself γ_i. Finally, the remaining groups must also not correspond to the characteristics given by the group i, this is the conditional probability of a non-correspondence in remaining groups from $k+1$ to m given that the k^{th} group had corresponded to i. After which, no other groups corresponding to i should occur, this probability is therefore 1. As such the general formula for the corresponding group to be the k^{th} group is as given at the bottom of Table 5.

Table 5
Table of match probabilities for the kth group corresponding given m FFGs.

k	Probability of k^{th} FFG corresponding to i
1	γ_i
2	$(1-\gamma_i)\gamma_i$
3	$(1-\gamma_i)^2\gamma_i$
…	…
k	$(1-\gamma_i)^{k-1}\gamma_i$

The total probability of one group corresponding to group i among m groups, M_i is then the summation of each of these k probabilities:

$$p(M_i|F=m) = \mu_{i|F=m} = \gamma_i \sum_{k=1}^{f} (1-\gamma_i)^{k-1} \tag{1}$$

This is a geometric series that can be shown to converge as m approaches infinity since, $0 < \gamma_i < 1$ to:

$$\gamma_i \sum_{k=1}^{\infty} (1-\gamma_i)^{k-1} = \frac{\gamma_i}{1-(1-\gamma_i)} = 1 \tag{2}$$

As such, as the number of FFGs approaches a large number, the probability of one of these groups corresponding to the reference material approaches 1 as desired. If only 1 FFG is present, this probability simply becomes the occurrence γ_i and with no FFGs present, the probability becomes 0. This appeals to the intuition that with no FFGs, there would be no fibres in the background and as a result, the characteristics cannot correspond to anything in the background.

Finally, the background class must also take into account the size of these non-differentiated background fibre groups. The probability distribution of each of these groups is specified in the parent S_i node for each respective B_i node. These nodes can take the same categorical states as the B_i node except for the “none” category. These distributions need to be assigned based on data and case circumstances. This is specified by the parameter s_i as shown in Table 1.

3.1.3. Transfer class

Moving on to the transfer class whose general structure is shown in Fig. 2 in the rectangle labelled “T”. This class outputs the probability distribution of categorical quantities of target fibre groups transferred (primary or secondary), persisted, and recovered with its respective T_i nodes. The transfers are dependent on the proposition, either from the POI or the AO during the alleged struggle. As such each T_i node is parented by the proposition node, H, which can take one of two states describing the competing propositions. For each T_i node, two probability distributions, one for each proposition needs to be assigned given data and case circumstances. It may be the case that the probability distribution of transfer given either proposition is equal, but they remain as separate events. These are respectively given by the parameters t_i under H_1 and $t\prime_i$ under H_2 as shown in Table 1.

Under the H_1, the non-differentiation of fibres is assumed to have a probability of 1. This is because the fibres recovered are given to be transferred from the reference material of the POI. However, under H_2, fibres transferred from the AO cannot be assumed to be non-differentiated by default. They will need to happen to have the same analytical characteristics as the reference material. As such the occurrence needs to be considered once again, but this time of the offender population rather than the background. Each T_i node is thus also parented by a $M\prime_i$ node which is in turn parented by the H node. The “$\prime$” in its superscript indicates that this node takes a different occurrence value than the one in the background class. This node can again take one of two states of non-differentiation or differentiation with the reference target fibre group. The probability of this solely depends on the occurrence, $\gamma\prime$ of the measured analytical characteristic used in comparison in the offender population.

Since the transfer class is the only one that encompasses the propositions, it is in this class in which the LR will be obtained. The findings of the analysis can be instantiated in the evidence class, which would then propagate information to the transfer class through the transfer nodes. The change in probabilities in the H node permit the calculation of the LR which describes the magnitude of support the findings have for one proposition over the other.

3.2. Current state of knowledge

In order to employ the model, the required parameters described for each class need to be assigned. Since all the node states are discrete and finite they may be modelled as categorical distributions where each possible state takes on a probability. This probability may be assigned as a straight value, such as with a maximum likelihood estimate from some distribution or using frequency data [11,27]. This is appropriate if the variability or uncertainty around this parameter is weak or not of interest. However, when the uncertainty around the parameter is of concern, the categorical distribution may itself be modelled as a Dirichlet distribution [11,27,47]. This permits the lack of knowledge to be taken into account in the modelling and reflected in the results. Furthermore, this allows the state of knowledge to be evaluated in the form of sensitivity analysis [11,47].

If based on the current state of knowledge, these assignments give insight into the current uncertainties and their effects on the overall evaluation. Once more data is available, they may also be easily integrated by using the assigned distribution as a prior distribution. This section will discuss the choice of assignment for each parameter which is summarised in Table 2. While it is acknowledged that a certain degree of subjectivity may be present in these assignments, the choices are justified in the discussion and transparent.

3.2.1. Background class parameters

The background class requires the assignment of FFG probabilities, f, occurrence probabilities, γ, and size category probabilities, s as shown in Table 1. These assignments can be made based on available data from a combination of population, colour block, and background studies. Through this discussion, it will hopefully become apparent that much of the current data is limited or incomplete with regards to f and s parameters. A certain level of judgement would be required to adapt this information into probability distributions.

3.2.1.1. FFG probability distribution. As far as the literature is concerned, there is a lack of reporting on the number of FFGs, with only three articles at the time of writing explicitly stating the number of FFGs recovered from the background [45,48,49]. Among which, only one concerns clothing, and its sample size is severely limited. While a uniform Dirichlet prior of all α taking on the value of 1 may be assigned, this is deemed inappropriate as the state of knowledge does allow some statements on the distribution of f to be made. A weakly informative prior is therefore more appropriate [50] where they are assigned based on data, but artificially widened to reflect uncertainty.

In general, it can be noted that it is highly unlikely that no FFGs are present as can be seen from the results of background studies [10] where fibres on various surfaces are always reported. Studies to date have yet to report a case in which no foreign fibres were found on a textile surface, which speaks to the ubiquitous nature of fibres in the background. Specifically concerning outerwear, Massonnet et al. [51] presented a study in which fourteen newly bought white t-shirts were found to have between 721 and 1134 foreign fibres, after washing and wearing, these numbers increased. Marnane et al. [45] reported between 9108 and 13, 925 foreign fibres recovered on each of six white t-shirts after one day of wear. Subsequently, on one of these t-shirts, over 1983 FFGs were reportedly identified. It remains doubtful that this many distinct classes of fibres can be distinguished based on purely microscopic means. Nonetheless, the considerable number of fibres recovered in these studies speak to how improbable it is that no FFGs are present in the background. These elevated numbers also suggest that it is highly probable that the number of FFGs on outerwear are high.

Based on the data so far, it is clear that the probability of no FFGs on a piece of clothing should be extremely low, and the probability of additional FFGs should increase accordingly. It seems therefore reasonable to attribute a distribution for the background number of FFGs as $f \sim Dir(1, 2, 4, 8, 16, 32)$. This distribution attributes a small probability for the case of no FFGs and increasingly higher probabilities for larger numbers of FFGs.

3.2.1.2. Occurrence probability distribution (background class). The probability of a non-differentiation, γ between a particular fibre group to a reference material, may be assigned based on the frequency of the corresponding characteristics in the relevant population. In contrast with data on FFGs, the literature is rich in data on this subject. While it is possible to model this parameter with a beta distribution, the use of frequency data to directly assign occurrence values has already been widely accepted [10]. Moreover, in the hypothetical scenario experiments, the idea of a non-differentiation is taken in a broad and overall sense. Any occurrence assigned here only requires an approximate notion of its rarity rather than a specific number. Since it is of interest to gain an understanding of the range of values that may be obtained, two extremes were considered. Specifically, the occurrence of a rare and common fibre was assigned. These two extremes allow one to appreciate the full scale of possible LRs based on varying occurrence values. When actual findings exist, as with the case study, each fibre group should be assigned their proper occurrence based on available data or experience.

The assignment of these occurrence values involves the combination of frequency data of the generic class and colour from relevant background population studies with colour block studies for the class and colour of the concerned fibre [10]. In combing through the literature, black cotton is consistently one of the most common fibres in many studies [10]. It was thus chosen as a representative “common” fibre group. On the other hand, concerning the choice of an archetypal “rare” fibre, orange cotton appears to be one of the more infrequent fibres, usually ranging in the $< 1\%$ range [10], where a colour block study has been done [52] and was chosen to represent the “rare” fibre group.

The most relevant population study of clothing was reported by Massonnet et al. [51] where fourteen white t-shirts were examined for foreign fibres. In this study the frequency of black cotton was reported at 24.0% and orange cotton at $< 1\%$. This is mostly consistent with other studies on the background discussed by Schnegg et al. [10], particularly with studies of surfaces in high contact with clothing, such as bus and cinema seats [34] and car seats [53]. Hence, an occurrence of 0.24 was assigned for black cotton and a conservative 0.01 for orange cotton.

In terms of colour block studies, the most applied technique for black cotton to further differentiate them is micro-spectrophotometry (MSP). Grieve et al. [54] applied this technique to 225 black cotton fibres, among which the largest non-differentiated group belongs to a group of 13 fibres. This corresponds to a frequency of 5.78%. Hence for a common fibre the following occurrence is assigned for the background:

$$\gamma_{common} = 0.24\text{x}0.0578 = 0.014$$

MSP is also the most applied method for orange cotton. Grieve et al. [52] conducted MSP analysis on 70 orange cotton fibres and reported that the smallest group that was non-differentiated was of size 12, giving these characteristics a frequency of 17.1%. Hence for a rare fibre the following occurrence is assigned for the background:

$$\gamma_{rare} = 0.01\text{x}0.171 = 0.00171$$

3.2.1.3. Size category probability distribution. The data on sizes of FFGs are once again scarce, with two main types of information. First, a few population studies do include the sizes recovered on each surface. Second, target fibre studies give insight into the fibre group sizes of specific fibre types after an analytical sequence has been applied. Therefore, the state of knowledge should again not be considered completely ignorant as some information is available.

In terms of population studies, the only relevant study concerning clothing which reports FFG sizes is from Marnane et al. [45]. However, this was only done for a sample size of one t-shirt. A huge concentration of these groups have sizes in the $1-5$ range, with 1802 out of 1983 groups being in this category. Despite the small sample size, it indicates that groups are much more likely to be small rather than large. This finding seems to be comparable to other background population studies in which most groups recovered are small [10].

Target fibre studies also support this notion where across literature [10], most report the largest recovered group on a single surface to be between $0-5$ with a few exceptions that are larger and none more than 50. Although these studies are typically done for only a select few fibre types, it does support the hypothesis that expected group size should be small. The data available may be limited in nature, however, in reviewing the overall information from the literature, they are still weakly informative. Hence, a distribution of $s \sim Dir(38, 10, 2)$ was assigned from the Marnane study by widening the distribution to reflect the uncertainty. This was done by maintaining the proportion of each size group from the study but reducing the total counts. This distribution gives the largest concentration of probability mass in the “small” size category followed by “medium” and “large” respectively.

3.2.2. Transfer class parameters

Finally, in the transfer class, two types of parameters need to be assigned (see Table 1). First the occurrence probabilities in the offender population, γ', which will follow the same process as with the probabilities assigned in the background class. Second, the transfer probabilities, t and t' under each proposition. Two further types of distributions are required here, one for primary transfers and the other for secondary transfers. The literature concerning the former event is quite well established with a multitude of information available. The latter event on the other hand, is lacking especially in data that is applicable to the given scenario.

3.2.2.1. Occurrence probabilities (transfer class). Occurrence probabilities in the transfer class, γ', can be assigned in the same way as detailed for the background class. As previously discussed, the use of these values has been well established and for testing a hypothetical scenario, exact values are not as important. As with the background class, black and orange cotton will be used as the archetypal common and rare fibres to gain an overall appreciation with values for specific target groups in the findings to be assigned in practice. The difference here is that the population of interest is no longer background surfaces but rather textiles in the offender population. It is postulated here that the types of textiles used by offenders of assault should not vary greatly from the general population. Since market studies offer a representation of the types of textiles available to the general public, data from such studies will be used as a benchmark for the frequencies of the generic fibre/class of fibres in the alternate offender population.

Muehlethaler and Albert’s publication [55] is the most recent market study to date. Utilising web-scraping methods, they were able to capture a snapshot of the generic fibre classes and colours that make up the clothing that can be purchased from an online store in 2020. While these do not represent the quantities sold for each clothing, this approximation was deemed appropriate for the purposes of this study. For black cotton, the frequency was reported to be 9.8 % and for orange cotton, it was 0.5 %. Combining these figures with their respective colour block studies, the following occurrence values of the transfer class can be assigned for a common and rare fibre:

$\gamma'_{common} = 0.098x0.0578 = 0.00566$ and

$\gamma'_{rare} = 0.005x0.171 = 0.000855$

3.2.2.2. Transfer probabilities. The transfer node, T can take on one of four states (see Table 1), with their probability specified by t or t' depending on the proposition. For the ease of specification at this stage, it will be assumed that $t = t'$. This is because it is not straightforward to adapt the available transfer data based on different donor and receiver characteristics. Moreover, these characteristics are unknown or unavailable in the scenarios of this study. As a result, for the time being, these transfer probabilities will be considered in general for the given activity without considering donor and receiver capacities.

Primary transfers are well documented in the literature, especially pertaining to assault activities. Lau et al. [56] and Sheridan et al. [57] both recently published experimental data in which assault activities were simulated by trained combat athletes. Both studies reported transfer of consistently in the order of magnitudes of hundreds to thousands of fibres primarily transferred. These results are consistent with what is known about primary transfer in the literature [8]. It would therefore be reasonable to assume that there is a disproportionately high probability of transfer of large groups, and decreasing probabilities for the “medium”, “small”, and “none” categories. As such, for the purposes of this study, the probability values are assigned for a primary transfer as shown in Table 2. Since the object of this study focuses on secondarily transferred fibres, these values are treated as reliable and will not be evaluated.

Secondary transfers on the other hand are lacking in data. Only two sources are considered here to be relevant to the matter at hand. First, Jackson and Lowrie [20] conducted an experiment on secondary transfer from person to person, where the first individual, wearing the donor, hugged a second individual wearing the intermediate for five seconds. Following which, this second person then went on to hug a third person wearing the final receiver for five seconds. This was done at a delay of fifteen minutes and one hour respectively with various textiles combining to a total of sixty different replicates. The results of this experiment are not directly representative for two reasons. First, the primary interaction of a five second hug is not representative of time spent in an environment, where interactions are much more prolonged, repeated and possibly intense. Additionally, given the closed nature of most environments, it is likely that the number of fibres primarily transferred would be higher. Similarly, the secondary interaction of a five second hug is also not representative of a violent struggle, which would suggest higher frictional forces and surface area in contact. As such, the number of fibres secondarily transferred during an assault would be expected to be higher than what is reported here. However, the data presented does offer a lower bound or a worst-case estimate for the scenario of interest. If taken at face value, the data suggests a concentration of probability mass at low to no fibres transferred given a struggle with close to zero probability for transfer of a medium or large sized group.

Another source of information is the master thesis work produced by Schmid [58]. In this work, fibres were first deposited onto a t-shirt. This was done by wearing the donor garment over a blanked t-shirt for an hour. This t-shirt was then donned by a volunteer who conducted a live simulation of a 60 second struggle with another volunteer wearing another t-shirt that acted as the receiver. This was repeated for sixteen trials under the same conditions. The numbers initially and secondarily transferred were both recorded which permitted the calculation of secondarily transferred fibres as a proportion of the primarily transferred. Although the activity tested here is complimentary with the scenario in question, the data produced in this work also has issues with applicability. This is mainly due to the way in which fibres were first deposited. Based on the methodology, it is likely that saturation has been reached during the initial transfer onto the intermediate. Furthermore, the secondary contact took place immediately after this primary deposition. Such conditions are unrealistic in two ways, first it is unlikely that the fibres from frequented environments approaches saturation based on daily interactions. Second, there is usually a lapse of time between leaving an environment and the assault in question. As such, these values are likely an over representation of a realistic scenario. This gives the opposite problem of the data provided by Jackson and Lowrie [59]. One way to correct for this over estimation is with the use of the reported proportion values. It is possible to artificially reduce the number of fibres from the primary transfer based on a loss function and apply these proportions to obtain more realistic counts of secondary transfer. This method is not foolproof, as it assumes that strongly persisting fibres would transfer at equal proportions to their weakly persisting counterparts. However, this offers an upper bound on the number of fibres expected to secondarily transfer. The process described here was carried out on the results presented by Schmid [58] assuming a 95% loss of fibres. This value was based on persistence curves in the literature [60] after more than 8 hours of wear. When taken at face value, the data suggests that probability of transfer would be concentrated on a medium sized group being transferred, followed by a large group and equally low probabilities for small sized and no fibres transferred.

Given the uncertainty surrounding the applicability of these data, and the fact that they represent two different extremes, the datasets were combined. The resulting distribution was also widened to reflect uncertainty. This was done by taking the weighted average of the two data sets and reducing the total counts. This gives the distribution $Dir \sim (13, 12, 4, 1)$ which appear consistent with what is known of secondary transfer. Namely that transfers most probably do occur but in small sizes, thus easily mistaken for the background. There is also a significant non-zero probability for the transfer of no fibres and the probability mass is lowest for the transfer of a larger group. When comparing to the number of fibres recovered from case studies where secondary transfers are evaluated, this distribution also appears to be sensible. De Wael et al. [19], reported numbers between $1-2$ recovered even after the recipient has been submerged in running water for several days. Similarly, Palmer [17] reports counts of between 1 and 26 on skin after the victims were submerged in water, these numbers fall into both the small and medium range. Note that for the moment a choice has been made to model the secondary transfer as absolute numbers. With the introduction of sufficient data, it may be more appropriate to model these as proportions of primary transfers from the environment with the use of additional nodes. However, it remains to be seen if this is possible.

Finally, these distributions do not take into account the persistence or recovery of the fibre groups. Although this can and should be integrated for specific cases as was done for the case study, they were kept not considered is for the hypothetical findings experiment. This was to maintain the genericity of the model as the objective is to examine the effects of considering secondary transfer. Persistence is thus out of the scope of this portion of the study and assumed to be of negligible influence in those instances.

3.3. Model implementation

3.3.1. Hypothetical findings

The results from condition A are plotted in Fig. 3. The corresponding verbal scale proposed by Marquis et al. [32] are indicated by green lines. At first glance, it appears that such findings can provide a certain value towards aiding the discrimination between the given propositions. This is contingent upon certain conditions being met. As observed, medium sized groups have LRs with stronger support of H_1 over H_2, as compared to small sized groups. This difference is more prominent in common groups than in rare groups. Intuitively, this comes as no surprise as smaller groups would have more overlap with the background sizes. Rare groups also provide stronger support in terms of their LR for H_1 over H_2. Again, this is intuitive since rarer groups have less of a chance to arise in the background or come from the AO. Finally, as more groups are considered, the LR also increases, indicating that based on the current state of knowledge, value is to be gained by considering additional groups. This increase appears to be exponential but also comes with an augmentation in variance as indicated by the greater spread in the box plots. Such a pattern suggests that the model is sensitive to the uncertainty in some parameters which compounds as the set of findings increase. As such, the lack of knowledge has an increased effect on the robustness of the evaluation as the number of target groups increases.

The results of condition B are shown in Fig. 4 and are similar to the results from condition A. The additional value for each additional group can be compared to the value of the findings evaluated at primary transfer labelled “base”. What is pertinent from these results is that the consideration of additional target groups from the POI’s environment have the potential to contribute additional probative value to findings based on the POI’s clothing. As with condition A, this is contingent upon the sizes of these groups being large enough and present in sufficiently rare combinations.

Aside from the generalities, these results demonstrate certain practical aspects, especially during case pre-assessment. Based on the values obtained, the expert and/or client would be able to set their priorities strategically based on the fibre types, group size, and number of target groups. This would allow an efficient investment of resources into

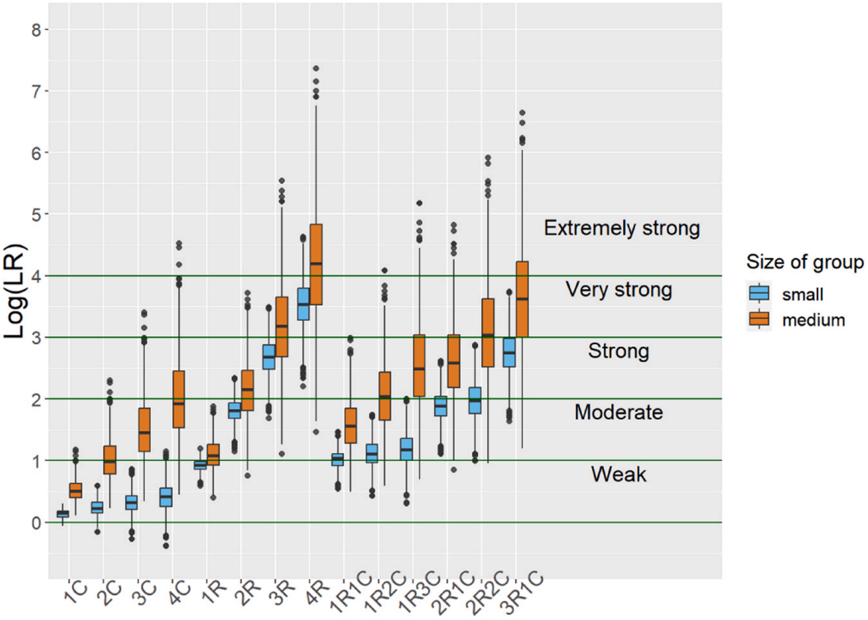

Fig. 3. Likelihood ratios in log form obtained from condition A. Fibre group occurrence combinations are abbreviated, ‘C’ for common and ‘R’ for rare. Preceding numbers indicate the number of such corresponding groups recovered as a finding.

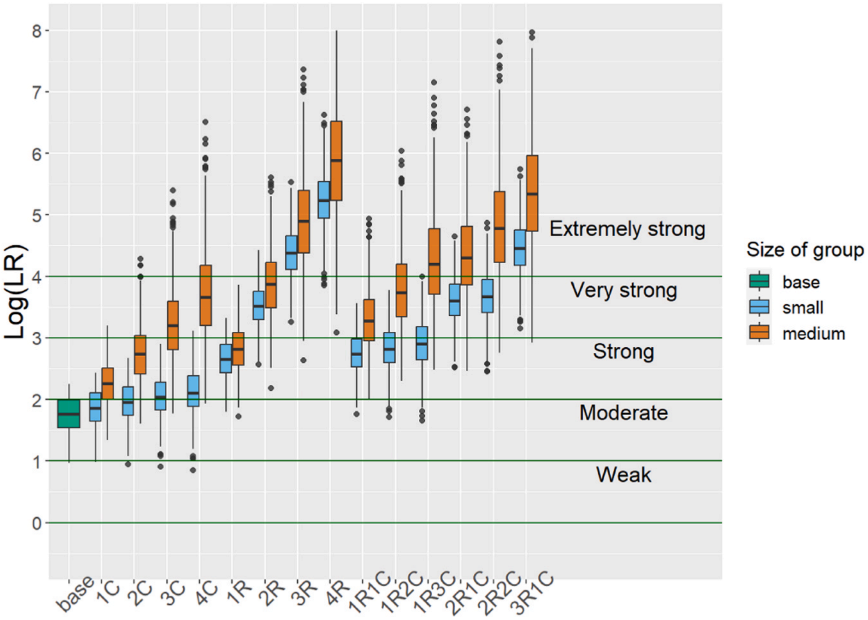

Fig. 4. Likelihood ratios in log form obtained from condition B. Fibre group occurrence combinations are abbreviated, ‘C’ for common and ‘R’ for rare. Preceding numbers indicate the number of such corresponding groups recovered as a finding. The LR for only one large group corresponding to the suspect’s garments at the time of the assault is included under the label ‘base’.

further analysis. For example, given the results, efforts would be better placed on rarer groups that are sufficiently large as they offer a stronger increase in value. Overall, the current state of knowledge appears to give grounds for the additional consideration of fibre groups that correspond to textiles in the POI’s environment provided that strategic choices are made. However, the results also highlight the need for more data, given the high sensitivity and variability. It seems justified and even essential to obtain more data on these parameters. A better understanding of these secondary transfer and background phenomena will permit a more accurate modelling of such events.

The results also expose some potential limitations of the model. The architecture of the model assumes certain independencies that have not been validated, namely between the following:

- Respective transfers of fibre groups given a proposition
- Transfer and occurrence of fibre groups
- Occurrence and size of fibre groups in the background

The exponential increase of LR observed with the increasing number of target groups gives some cause for concern. This increase was initially hypothesised to be logarithmic rather than exponential because additional groups were expected to add diminishing returns on the value provided. Based on the results in condition A, four rare corresponding groups push the LR to the order of magnitudes of above a hundred thousand. These numbers seem questionably high especially for results of fibre analysis. Any further increase, which is to be expected as more groups are considered, should reasonably raise some doubts. Therefore, it may well be the case that an assumption of independence between parameters may not be valid. On the other hand, it could also be plausible that the exponential increase holds but that it would be unrealistic to recover this many groups of fibres with such rarity. Finally, as with most evaluations of findings on activity level, only evidence that is present would be instantiated in this case. Findings relating to the absence of fibres that could not be differentiated to other reference materials are not taken into account. The use of absent evidence has been covered by Taroni et al. [54] and is out of the scope of this paper. Not enough information is available at the moment to make adjustments to the model. This represents a potential area in which research should be conducted to validate such models. With a better understanding of the dependencies, the effect of absent evidence, and more data, better models may be created.

3.3.2. *Sensitivity analysis*

The results of the sensitivity analysis are shown in Fig. 5. The results show the largest spread for the secondary transfer and size parameters, with simulated LRs spanning two orders of magnitude. Surprisingly, the simulated LRs appear to be quite stable to the FFG parameter. Globally, sensitivity analysis highlights the magnitude of the effect of a parameter and their uncertainties on the overall evaluation. These results suggest that the parameters of secondary transfers and size play an influential role in the evaluation. Given the small sample sizes in directly applicable data for these parameters, data collection should be prioritised for them. Such work would also permit the verification of the independence assumptions discussed previously to improve the model.

3.3.3. *Case study*

The results from the case study are reported in Table 6. Concerning both victims, the combination of fibres that could not be differentiated from the POI clothing (considered under primary transfer) already contribute substantially in value towards aiding the discrimination

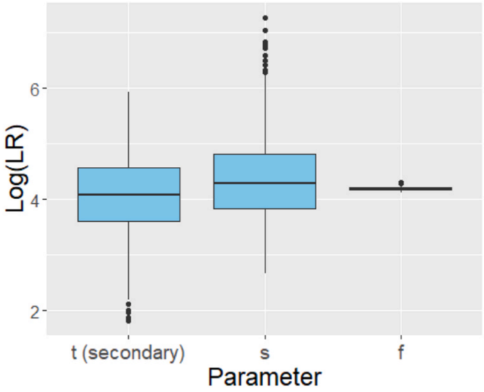

Fig. 5. Results from re-sampling from the respective prior distributions for secondary transfer (t), size (s), and number of FFGS (f) parameters.

Table 6
Results from case study.

H_1	LR (verbal equivalent) Primary transfers only	LR (verbal equivalent) All transfers
POI struggled with Victim 1	1020 (Very strong support for H_1 over H_2)	2200 (Very strong support for H_1 over H_2)
POI struggled with Victim 2	8500 (Very strong support for H_1 over H_2)	147,000 (Extremely strong support for H_1 over H_2)

between the propositions. As for victim 1, it was observed that the analysis of the additional 2 red cotton fibres did not contribute significantly to the overall value. Contrasted with victim 2, where the additional combination of fibres added much more substantial value. This case, as with most fibre cases of this magnitude, took considerable resources on the part of the experts [19]. A case pre-assessment that can provide such information would allow better priorities to be set before the undertaking of further tedious analysis. Such results could be presented to clients for them to make an informed decision on the investment of resources. Concerning this case study, it would be more beneficial to place emphasis on the additional fibres recovered on victim 2 than on victim 1. Finally, once the work is done, the model permits a precise valuation of the findings, such that they are not over or under-represented, permitting the diffusion of transparent information to the decision maker. The integration of persistence as a factor was also demonstrated in this application, showing the versatility of the classes to take up different distributions based on the case circumstances.

4. Where to from here and concluding remarks

The presently proposed model is by no means without its flaws, but it represents a step towards a more fine-grained evaluation of findings. This goes beyond generalised models suited to simplified cases which are difficult to apply. The OOBN architecture permits enough flexibility to adapt to specific case-based scenarios while still maintaining a suitable level of complexity required. It allows the adjustment for different scenarios such as varying findings, propositions, and other relevant contextual information. For example, different findings can be accounted for in each class by adjusting the number of groups, g, and specifying their occurrence. Alternate propositions or contextual information may also be taken into consideration by modifying the relevant parameters of transfer, size, or FFGs to reflect the situation at hand. These modifications can be made easily without having to start from scratch. The architecture can also be updated once more data is available to take into account even more factors like contact intensity, persistence, and donor or receiver characteristics, further increasing the versatility of the model.

To address the question of “where to from here?”, such models may already be implemented despite the lack of knowledge. Their use affords firstly the transparent and clear indication of value of findings based on the most current data. This information would be useful both during the case pre-assessment phase and for evaluative reporting after the analytical work has been carried out. The former gives a transparent indication to clients and experts such that informed decisions on resource allocation may be made. The latter permits a clear presentation of the value of findings to decision makers such as a judge or jury. This would help them in rendering a decision between two propositions and to place adequate weight on the scientific expert’s findings. Additionally, such models permit the evaluation of the current state of knowledge, along with its uncertainties, affording even more transparency for stakeholders and decision makers. Such results should guide the future research and drive data collection. In this case, it is apparent that the focus should be on the secondary transfer and size parameters, preferably with experiments that would grant the model more fine-grained complexities. Such research would be fundamental to the understanding of how fibre traces can contribute to specific case-based scenarios.

Work of this nature is already underway, and its result would serve to validate the current model by supporting or challenging the independence assumptions made. Adjustments to the model can then be made such that it may better serve its purpose. Additionally, this gain in knowledge may be exploited to serve further objectives of forensic science, such as for intelligence and investigative purposes.

Several principles published in the Sydney declaration emphasises the need for solid evaluation of forensic findings. This paper addresses this point with a focus on findings of fibre analysis, specifically the comparison to reference material that did not have direct contact with the recipient under any proposition. It proposed an evaluative model to interpret such findings and accessed the current state of knowledge. The model, based on an OOBN structure, was tested on both hypothetical findings and a case study that illustrates its utility in complex scenarios. Its parameters were assigned and populated using data that is currently available. The benefits of the model, as well as the value of considering findings beyond that which can be explained by primary transfer were demonstrated, especially if specific conditions are met. Namely, factors of rarity, group sizes, and number of groups play a role in the final probative value. These are considerations that should come into play from the beginning during case pre-assessment, and at the end should these findings be presented to a decision maker. It also highlighted the limitations and areas of research interest that should be explored. Specifically, the current state of knowledge still leaves a margin of uncertainty that is mainly driven by the lack of data in background group sizes and secondary transfer. Further research should focus on these two areas to refine and validate the model at hand.

Declaration of Competing Interest

None

Acknowledgements

The authors would like to extend their gratitude to Laurent Lepot, Kyra Lunstroot, and Fabrice Gason of the INCC who kindly provided information and guidance for the case study. Additionally, we would like to thank Christophe Champod at the Université de Lausanne who provided helpful advice during the early stages of the model development.

References

[1] C. Roux, R. Bucht, F. Crispino, P. De Forest, C. Lennard, P. Margot, M.D. Miranda, N. NicDaeid, O. Ribaux, A. Ross, S. Willis, The Sydney declaration – Revisiting the essence of forensic science through its fundamental principles, Forensic Sci. Int 332 (2022), https://doi.org/10.1016/j.forsciint.2022.111182.

[2] T. Hicks, J. Buckleton, V. Castella, I. Evett, G. Jackson, A Logical Framework for Forensic DNA Interpretation, Genes (Basel) 13 (2022) 957, https://doi.org/10.3390/genes13060957.

[3] European Network Of Forensic Institutes, Guideline for evaluative reporting in forensic science, (2015).

[4] R. Cook, I.W. Evett, J.A. Lambert, A model for case assessment, Sci. Justice 38 (1998) 151–156.

[5] C. Aitken, F. Taroni, S. Bozza. Statistics and the Evaluation of Evidence for Forensic Scientists 3e, 3rd ed., Wiley, Chichester, UK, 2021 https://doi.org/10.1002/9781119245438.

[6] F. Taroni, P. Juchli, C. Aitken, A probabilistic account of the concept of cross-transfer and inferential interactions for trace materials, Law, Probab. Risk (2020) 1–13, https://doi.org/10.1093/lpr/mgaa015.

[7] F. Taroni, P. Garbolino, S. Bozza, C. Aitken, The Bayes’ factor: the coherent measure for hypothesis confirmation, Law, Probab. Risk 20 (2022) 15–36, https://doi.org/10.1093/lpr/mgab007.

[8] M. Schnegg, R. Palmer, G. Massonnet, Les paramètres clés de l’interprétation des fibres textiles en sciences criminelles. Partie II: Transfert, persistance et détection, Journal of, Can. Soc. Forensic Sci. 51 (2018) 83–119, https://doi.org/10.1080/00085030.2018.1519769.

[9] C. Champod, F. Taroni, Bayesian framework for the evaluation of fibre transfer evidence, Sci. Justice 37 (1997) 75–83, https://doi.org/10.1016/S1355-0306(97)72151-8.

[10] M. Schnegg, R. Palmer, G. Massonnet, Les paramètres clés de l'interprétation des fibres textiles en sciences criminelles. Partie I: Occurrence et bruit de fond, Journal of, Can. Soc. Forensic Sci. 51 (2018) 1–25, https://doi.org/10.1080/00085030.2017.1379683.

[11] F. Taroni, A. Biedermann, S. Bozza, P. Garbolino, C. Aitken, Bayesian Networks and Probabilistic Inference in Forensic Science, Wiley, 2014.

[12] F. Taroni, A. Biedermann, P. Garbolino, C.G.G. Aitken, A general approach to Bayesian networks for the interpretation of evidence, Forensic Sci. Int 139 (2004) 5–16, https://doi.org/10.1016/j.forsciint.2003.08.004.

[13] N. Petraco, N.D.K. Petraco, A Guide to the Analysis of Forensic Household Dust Specimens and Their Statistical Significance, in: Forensic Science Handbook: Third Edition, 2017: pp. 201–244. https://doi.org/10.4324/9781315119939-4.

[14] A. Coxon, M. Grieve, J. Dunlop, A method of assessing the fibre shedding potential of fabrics, J. Forensic Sci. Soc. 32 (1992) 151–158, https://doi.org/10.1016/S0015-7368(92)73064-3.

[15] K. De Wael, L. Lepot, K. Lunstroot, F. Gason, Evaluation of the shedding potential of textile materials, Sci. Justice 50 (2010) 192–194, https://doi.org/10.1016/j.scijus.2010.06.001.

[16] O. Ribaux, Police scientifique Le renseignement par la trace, Presses Polytechniques et Universitaires Romandes, Lausanne, 2014.

[17] R. Palmer, The Ipswich Serial Killings The Fibre Evidence, in: 16th EFG Meeting, Budapest, 2008.

[18] S.A. Egger, Working Definition of Serial Murder and the Reduction of Linkage Blindness, J. Police Sci. Adm. 12 (1984) 348–357.

[19] K. De Wael, Fibres in the investigation of a double child murder Fibres in the investigation of a double child murder, in: 17th EFG Meeting, Berlin, 2009.

[20] C.N. Lowrie, G. Jackson, Secondary transfer of fibres, Forensic Sci. Int 64 (1994) 73–82, https://doi.org/10.1016/0379-0738(94)90215-1.

[21] M.C. Grieve, J. Dunlop, P.S. Haddock, Transfer experiments with acrylic fibres, Forensic Sci. Int 40 (1989) 267–277, https://doi.org/10.1016/0379-0738(89)90185-0.

[22] E. Davis, Case Study: A Secondary Fiber Transfer., in: 12th EFG Meeting., Prague, 2004.

[23] D.A. Stoney, P.L. Stoney, Critical review of forensic trace evidence analysis and the need for a new approach, Forensic Sci. Int 251 (2015) 159–170, https://doi.org/10.1016/j.forsciint.2015.03.022.

[24] J. Robertson, C. Roux, Trace evidence: Here today, gone tomorrow? Sci. Justice 50 (2010) 18–22, https://doi.org/10.1016/j.scijus.2009.11.005.

[25] C. Roux, B. Talbot-Wright, J. Robertson, F. Crispino, O. Ribaux, The end of the (forensic science) world as we know it? The example of trace evidence, Philos. Trans. R. Soc. B: Biol. Sci. 370 (2015), https://doi.org/10.1098/rstb.2014.0260.

[26] F. Taroni, P. Garbolino, C. Aitken, A generalised Bayes' factor formula for evidence evaluation under activity level propositions: Variations around a fibres scenario, Forensic Sci. Int 322 (2021), https://doi.org/10.1016/j.forsciint.2021.110750.

[27] R. Palmer, The evaluation of fibre evidence in the investigation of serious crime, 2016. ⟨http://serval.unil.chhttp//serval.unil.ch⟩.

[28] R. Core Team, R: A Lang. Environ. Stat. Comput. (2022).

[29] ScutariPadova, Learning Bayesian Networks wirth bnlearn R package 35 (2010).

[30] S.H. jsgaard, Graphical Independence Networks with the **gRain** Package for *R*, J. Stat. Softw. 46 (2012), https://doi.org/10.18637/jss.v046.i10.

[31] RStudio Team, RStudio: Integr. Dev. R. (2020).

[32] R. Marquis, A. Biedermann, L. Cadola, C. Champod, L. Gueissaz, G. Massonnet, W. D. Mazzella, F. Taroni, T. Hicks, Discussion on how to implement a verbal scale in a forensic laboratory: Benefits, pitfalls and suggestions to avoid misunderstandings, Sci. Justice 56 (2016) 364–370, https://doi.org/10.1016/j.scijus.2016.05.009.

[33] K. De Wael, C. Baes, L. Lepot, F. Gason, On the frequency of occurrence of a peculiar polyester fibre type found in blue denim textiles, Sci. Justice 51 (2011) 154–162, https://doi.org/10.1016/j.scijus.2011.02.001.

[34] J. Wąs-Gubała, Comparative population studies of fibres secured in Poland, Czech Republic and Germany, Probl. Forensic Sci. 60 (2004) 58–77.

[35] M.C. Grieve, T.W. Biermann, K. Schaub, The individuality of fibres used to provide forensic evidence - Not all blue polyesters are the same, Sci. Justice 45 (2005) 13–28, https://doi.org/10.1016/S1355-0306(05)71616-6.

[36] M.C. Grieve, J. Dunlop, A practical aspect of the Bayesian interpretation of fibre evidence, J. Forensic Sci. Soc. 32 (1992) 169–175, https://doi.org/10.1016/S0015-7368(92)73066-7.

[37] M.C. Grieve, J. Dunlop, P. Haddock, An Assessment of the Value of Blue, Red, and Black Cotton Fibers as Target Fibers in Forensic Science Investigations, J. Forensic Sci. 33 (1988) 12577J, https://doi.org/10.1520/jfs12577j.

[38] L. Lepot, T. Vanden Driessche, Fibre persistence on immersed garment - Influence of water flow and stay in running water, Sci. Justice 55 (2015) 431–436, https://doi.org/10.1016/j.scijus.2015.09.003.

[39] F. Taroni, S. Bozza, A. Biedermann, C. Aitken, Dismissal of the illusion of uncertainty in the assessment of a likelihood ratio, Law, Probab. Risk 15 (2016) 1–16, https://doi.org/10.1093/lpr/mgv008.

[40] 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, https://doi.org/10.1016/j.fsigen.2017.02.008.

[41] 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, https://doi.org/10.1016/j.fsigen.2019.06.006.

[42] Y. Xiang, D. Poole, M.P. Beddoes, Multiply Sectioned Bayesian Networks and Junction Forests for Large Knowledge-Based Systems, Comput. Intell. 9 (1993) 171–220, https://doi.org/10.1111/j.1467-8640.1993.tb00306.x.

[43] O. Bangsø, P.-H. Wuillemin, Object Oriented Bayesian Networks A Framework for Topdown Specification of Large Bayesian Networks and Repetitive Structures, (2000).

[44] D. Koller, N. Friedman, Proablistic Graphical models, The MIT Press, Cambridge, Massachusetts, 2009.

[45] R.N. Marnane, D.A. Elliot, S.A. Coulson, A pilot study to determine the background population of foreign fibre groups on a cotton/polyester T-shirt, Sci. Justice 46 (2006) 215–220, https://doi.org/10.1016/S1355-0306(06)71601-X.

[46] I.W. Evett, On meaningful questions: a two-trace transfer problem, J. Forensic Sci. Soc. 27 (1987) 375–381, https://doi.org/10.1016/S0015-7368(87)72785-6.

[47] F. Taroni, S. Bozza, A. Biedermann, P. Garbolino, C. Aitken, Data Analysis in Forensic Science, Wiley, 2010, https://doi.org/10.1002/9780470665084.

[48] C. Roux, P. Margot, An attempt to assess the relevance of textile fibres recovered from car seats, Sci. Justice 37 (1997) 225–230, https://doi.org/10.1016/S1355-0306(97)72195-6.

[49] V. Cammarota, M. Schnegg, G. Massonnet, A study of background population of fibres on knife blades, Forensic Sci. Int 296 (2019) 132–143, https://doi.org/10.1016/j.forsciint.2019.01.006.

[50] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin. Bayesian data analysis, 3rd ed., Chapman and Hall/CRC, 2013.

[51] G. Massonnet, M. Schisser, C. Champod, Population of textile fibres on white T-shirts, in: 6th European Fibres Group Meeting, Dundee, Scoteland, 1998.

[52] M.C. Grieve, T. Biermann, M. Davignon, The occurrence and individuality of orange and green cotton fibres, Sci. Justice 43 (2003) 5–22, https://doi.org/10.1016/S1355-0306(03)71736-5.

[53] C. Roux, P. Margot, The population of textile fibres on car seats, Sci. Justice 37 (1997) 25–30, https://doi.org/10.1016/S1355-0306(97)72137-3.

[54] M.C. Grieve, T.W. Biermann, M. Davignon, The evidential value of black cotton fibres, Sci. Justice 46 (2001) 198–202.

[55] C. Muehlethaler, R. Albert, Collecting data on textiles from the internet using web crawling and web scraping tools, Forensic Sci. Int 322 (2021) 110753, https://doi.org/10.1016/j.forsciint.2021.110753.

[56] V. Lau, X. Spindler, C. Roux, The Influence of Physical Activity Upon the Population of Extraneous Fibres on Cotton T-shirts, in: 17th ETHG Meeting, The Hague, Netherlands, 2017.

[57] K. Sheridan, R. Palmer, D. Charlton, J. Beckett, K. Bellerby, L. Brown, E. Donaghy, A. Finlayson, C. Graham, B. Robertson, L. Taylor, M. Gallidabino, Transfer in Assault Scenarios, in: 28th ETHG Annual Meeting, On-line, 2021.

[58] P. Schmid, Le transfert secondaire de fibres textiles, 2022.

[59] C.N. Lowrie, G. Jackson, Recovery of transferred fibres, Forensic Sci. Int 50 (1991) 111–119, https://doi.org/10.1016/0379-0738(91)90139-A.

[60] C.A. Pounds, K.W. Smalldon, The Transfer of Fibres between Clothing Materials During Simulated Contacts and their Persistence During Wear: Part II—Fibre Persistence, J. Forensic Sci. Soc. 15 (1975) 29–37, https://doi.org/10.1016/S0015-7368(75)70933-7.