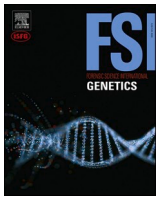
Forensic Science International: Genetics 73 (2024) 103122

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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



Accounting for site-to-site DNA transfer on a packaged exhibit in an evaluation given activity level propositions

Duncan Taylor [a,b,*], Luke Volgin [a], Bas Kokshoorn [c,d]

[a] Forensic Science SA, GPO Box 2790, Adelaide, SA 5001, Australia
[b] School of Biological Sciences, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia
[c] Forensic Trace Dynamics, Faculty of Technology, Amsterdam University of Applied Sciences, Amsterdam, the Netherlands
[d] Netherlands Forensic Institute, The Hague, the Netherlands

ARTICLE INFO

Keywords:
Activity level evaluation
Bayesian Network
DNA transfer
Site to site transfer
Exhibit packaging and transport, contamination

ABSTRACT

Considering activity level propositions in the evaluation of forensic biology findings is becoming more common place. There are increasing numbers of publications demonstrating different transfer mechanisms that can occur under a variety of circumstances. Some of these publications have shown the possibility of DNA transfer from site to site on an exhibit, for instance as a result of packaging and transport. If such a possibility exists, and the case circumstances are such that the area on an exhibit where DNA is present or absent is an observation that is an important diagnostic characteristic given the propositions, then site to site transfer should be taken into account during the evaluation of observations. In this work we demonstrate the ways in which site to site transfer can be built into Bayesian networks when carrying out activity level evaluations of forensic biology findings. We explore the effects of considering qualitative vs quantitative categorisation of DNA results. We also show the importance of taking into account multiple individual's DNA being transferred (such as unknown or wearer DNA), even if the main focus of the evaluation is the activity of one individual.

1. Introduction

Numerous bodies have published guidelines on evaluation of forensic findings given activity level propositions [1–3]. New works on matters of DNA transfer, persistence, prevalence and recovery (TPPR) are being published at a rapid rate (see reviews [4,5]). The Netherlands Register of Court Experts (NRGD; https://english.nrgd.nl/) has introduced a category of expert competency testing and registration that deals specifically with evaluations at activity level. Feedback from court-attending forensic biology scientists is that the topics they are questioned on most (over 60 % of the time) relate to TPPR issues (see Fig. 1). These findings mirror those of Yoon et al. [6] in the US and Prinz et al. [7] internationally. At the same time, anecdotally, the authors have noticed the number of challenges to the evaluation of DNA profiles given sub-source level evaluations is decreasing. This is likely due in part to the rising use of probabilistic genotyping (see [8]) increasing the standardisation of DNA profile interpretation and the many published validation articles now available (see the number of validation publications tabled in [9]). These facts make the point that the evaluation of forensic biology observations is shifting focus from DNA source to activity.

Generating the data on TPPR issues underpins the ability to carry out activity level evaluations. However, just as important as having the data, is being able to evaluate it appropriately. To this end, there have been several articles that seek to provide guidance on different aspects of evaluation [2,10–16]. A common tool used to assist in carrying out evaluations are Bayesian networks (BN) [17]. BNs are graphical depictions of probabilistic reasoning and can be preferable to hand-deriving likelihood ratio (*LR*) formulae, particularly for complex evaluations that involve multiple conditional probabilities [18].

Studies have shown that there are multiple factors that can have an effect on probability of DNA TPPR [4]. Consequently, the structure of a BN being used to evaluate forensic observations will depend on the framework of case circumstances that applied to the relevant exhibits. One factor that can play an important role in an evaluation is the location of DNA on an item. For example, Zuidberg et al. [19] conducted experiments that found that different factors about the framework of circumstances will affect the positions that are contacted on a body that is being moved. Ramos et al. [20] found that depending on the type of activity performed, underwear will be contacted in different areas. De

* Correspondence to: Forensic Science SA, GPO Box 2790, Adelaide, South Australia 5001, Australia.
E-mail address: Duncan.Taylor@flinders.edu.au (D. Taylor).

https://doi.org/10.1016/j.fsigen.2024.103122
Received 9 April 2024; Received in revised form 1 August 2024; Accepted 2 August 2024
Available online 10 August 2024

Ronde et al. have shown that:

- Different areas of a pillowcase are contacted when the pillow is used to smother someone as opposed to having a pillowcase changed [21].
- Different patterns of contact may be found on a sheet of paper given either writing or reading a letter [22].
- Different activities will lead to fingerprint deposition on knifes in different locations [23].

These studies all have experimental designs whereby the exhibit is sampled directly after the action has occurred. This differs from the way in which many exhibits are handled in criminal casework. Frequently, rather than being sampled at the scene of the incident, an exhibit would be collected and placed in an exhibit bag, which would then be transported to a forensic facility. Under such conditions DNA can transfer from one site on an item to another site on that same item, due to contact with itself, and the bag during transport [24]. Even before the forensic process, an item may be handled or transported in a way that introduces a potential for site-to-site re-location of trace materials. This is referred to as bi-directional DNA transfer in [4] or cross-transfer [25]. The idea of cross-transfer in this way has been shown in BN constructions [26,27] and shows the dependency of the transfer between the two items. Site-to-site transfer can be considered as a type of cross transfer. Most literature considers cross transfer as the transfer of two different DNA sources from two different objects, however there is no need for the definition of cross-transfer to be restricted in either of these ways. Cross transfer could be considered a transfer of DNA from one individual between two sites on the same object i.e. site-to-site transfer.

The work presented in this paper deals with site-to-site transfer in two ways. First, we develop a theoretical framework for incorporating site-to-site transfer within an activity level evaluation using BNs. This includes a consideration of DNA transfer in a binary fashion (i.e. the presence or absence of DNA) as well as using DNA amounts. We also show how to incorporate results from the sub-source level evaluation of DNA profiles, and how to extend the model to consider the presence of DNA of multiple individuals. Furthermore, we set the conditional probability tables in the BN using previously published data on the amount of DNA transferred from different body fluids and latent DNA within packaging under different transport and handling conditions.

2. Modelling site-to-site transfer through packaging using bayesian networks

Throughout the paper we show examples of BNs. These have been constructed using the software HUGIN V9.5 [28]. We provide the final constructed BN as supplementary material to this paper, and a document detailing the conditional probability tables.

2.1. A case example of site-to-site transfer

Consider the following simplified case example:

The complainant (C) and defendant (D) met for the first time at a ball and danced together. During the dance D placed his hand on C's hip, contacting her underwear waistband (which was exposed above her skirt). C claims that D put his hand down the front of her skirt and underwear and attempted to digitally penetrate her vagina shortly after the dance. D denies this occurred. The underwear was seized by police, placed in an evidence bag, and sent to the local forensic facility. At the forensic facility, DNA samples were taken from the inner groin and outer waistband of C's underwear. D's DNA was present on both samples.

The propositions and case information (laid out as per [29]) for this scenario are:

Prosecution proposition, Hp)

- D placed his hands down the front of C's underwear

Defence proposition, Hd)

- No one placed their hands down the front of C's underwear

Undisputed case information, I)

- D and C were dancing during which D touched the waistband of C's underwear
- D and C have not interacted prior to the dance

Assumptions, A)

- D's DNA is present on the waistband and inner groin of C's underwear (noting that the sub-source level evaluation would not have provided a conclusion of identity, but rather provided an LR supporting the presence of D's DNA being present compared to an unknown person's DNA being present)

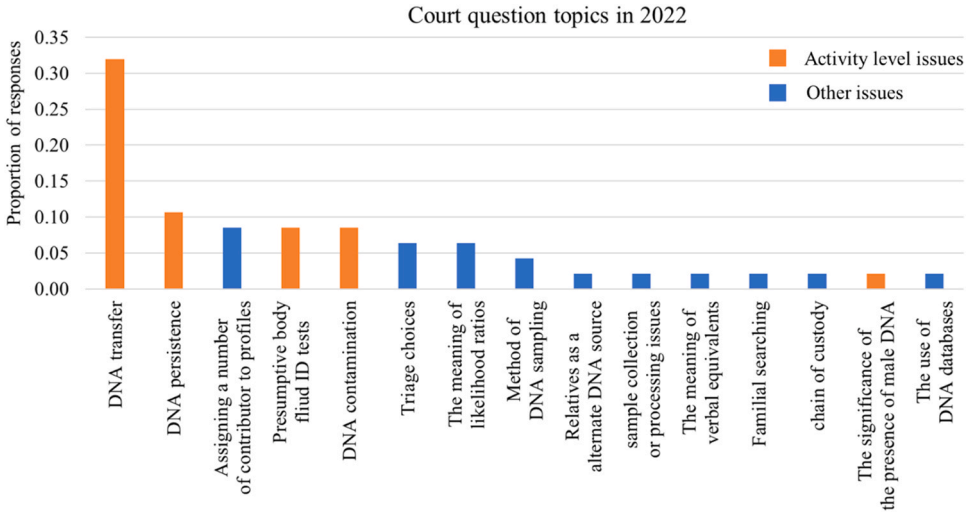


Fig. 1. Summary of topics that court testifying forensic biologists in South Australia were questioned on when testifying during 2022 (Forensic Science SA internal data capture on 47 self-reported court testimonies).

These propositions, case information, and assumption will remain constant for all evaluations used throughout the paper. The assumption made by the reporting scientist indicates that the dispute in court is not presence or absence of D's DNA. One possible BN construction is shown in Fig. 2, but without yet adding in consideration for site-to-site transfer during the underwear's time in the exhibit bag. Note that the case example is simplified to more clearly communicate the issue of cross-transfer between sites on the underwear. In an actual casework situation, the sampling strategy may have been different (for instance it may be that different areas of the waistband, like front and sides, as well as inside and outside, would have been sampled separately). Also, some DNA transfer to the waistband would also potentially occur from digital penetration. While it is crucial to model all relevant mechanisms for DNA transfer in a formal evaluation of case findings, we do not take this mechanism into account in the BN seen in Fig. 2, purely due to the fact that it visually complicated the BN and detracts from the main point we are making about dealing with site-to-site DNA transfer. In casework we would expect that it would be sensible to take this possibility into account by adding another child node of 'D places hand down the front of C's underwear' which was 'D DNA transfer to waistband of C underwear from digital penetration' and that this would yield another pathway for D's DNA to be present on the waistband.

In the BN shown in Fig. 2 the architecture utilises:

'Hp/Hd': A proposition node that has states that represent the two propositions

'D placed hands down front of C's underwear' and 'D touched C's underwear during dancing': Two activity nodes that represent the two activities being considered in the evaluation. Both activity nodes are children of the proposition node and so probability assignments (typically taking values of 0 or 1) will dictate which activity occurs according to each proposition.

'D DNA transfer from D hands to C underwear inner groin from touching' and 'D DNA transfer from D hands to C underwear waistband from dancing': are nodes in which the probability of DNA being transferred, persisting, and being recovered given the state of the parent activity nodes are assigned. Typically, the probabilities assigned will take values between 0 and 1 and will be determined by available literature and expert knowledge.

'D DNA observed on C underwear inner groin' and 'D DNA observed on C underwear inner groin': are results nodes, which will be instantiated by the analyst depending on the DNA profiling observations from the samples taken from the underwear. The probabilities assigned to states in these nodes will again typically take values of 0 or 1 depending on the presence or absence of DNA in the parent nodes.

We do not, at this stage, provide the probability assignment used in the construction of the conditional probability tables in the BN in Fig. 2. Assuming these are assigned in a reasonable manner, there is nothing illogical about the construction of the BN, however there are practical issues that would mean it is inappropriately constructed for use in a case evaluation. The first point to note is that the DNA results from the waistband (as shown in the construction of the BN in Fig. 2) are not related to the alleged criminal activity (D placing his hand down C's underwear). In fact, the activity node related to D dancing with C would have the same probability assignment under both propositions, meaning that it is independent of the propositions and hence has no effect on the evaluation. The three right hand nodes could therefore be removed from the BN, except that the BN shown in Fig. 2 is a steppingstone to considering DNA transfer between sites in exhibit packaging (where the presence of those nodes will become important to the evaluation). The second point is that as the criminal activity is the only explanation given in the BN structure for the presence of D's DNA on the groin of C's underwear, then the instantiation of D's DNA being present will lead to

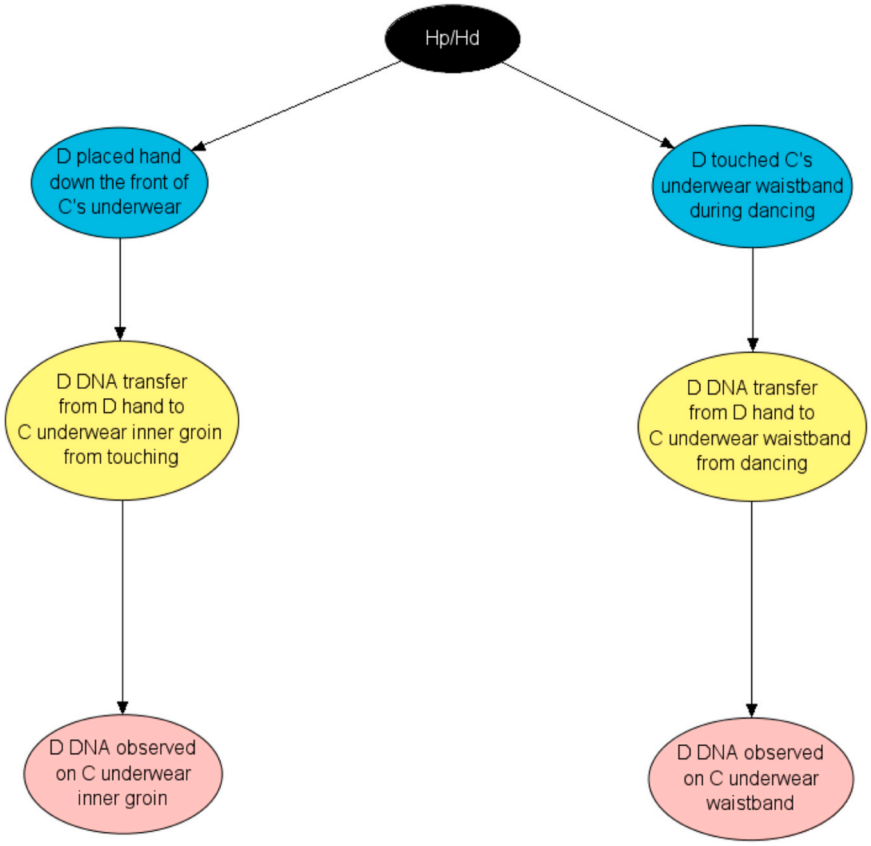


Fig. 2. BN constructed to assist in the evaluation of DNA result in the case example. Blue nodes are activities, yellow nodes are TPPR nodes and red nodes are results nodes. Black nodes are proposition nodes.

certainty in Hp, which is not an acceptable evaluation result to provide. Both of these issues are dealt with by considering the possibility of site-to-site transfer between the waistband and outer groin of C's underwear whilst in the bag during transport from the crime scene to the forensic facility. We also do not consider the possibility here for dancing to lead to DNA on the inner groin of the underwear without the exhibit site-to-site transfer. This is a simplifying assumption made for our example only and may not be justified given the case context in regular casework evaluation (for example if C went to the bathroom shortly after dancing, then this potential transfer route for D's DNA to the inner underwear groin, for instance via her hands, may be included in the evaluation).

When two areas, each of which may possess a trace material, come into contact with each other, there exists the possibility of transfer of that material in both directions. This is referred to as bi-directional DNA transfer in [4]. The idea of cross-transfer in this way has been shown in BN constructions [26,27] and shows the dependency of the transfer between the two items. The form of cross transfer we deal with our paper is slightly different to the cross-transfer example given by Taroni et al. [27]. While cross-transfer is certainly being considered, the important distinction is that it is not a transfer of two different sources of the trace i.e. in DNA terms, we are interested in the cross transfer of two traces, both of which have come from the same person. The BN shown in Fig. 3 shows the extension of the BN in Fig. 2 with the consideration of cross-transfer.

The BN in Fig. 3 has a number of new nodes that have been added (compared to the BN in Fig. 2) that specifically deal with DNA transfer between sites on the underwear.

'properties of contact between C underwear groin and waistband': in which states can be set by the analyst to reflect different types of contact that occur between sites of the underwear (we discuss this in more detail later)

'D DNA transfer from C underwear groin / waist band to C underwear waistband / groin from cross transfer': These are transfer nodes where the analyst assigns probabilities of DNA transferring, and being recovered, from site to site, depending on the properties of the contact. Probabilities will typically be assigned with values between 0 and 1.

'D DNA on C underwear groin' and 'D DNA on C underwear waistband': These are the white accumulation nodes in Fig. 3 which accumulate the two possible sources of D's DNA on each area of the underwear. These nodes will have probability values assigned as 0 or 1.

The BN seen in Fig. 3 does not include a node that details whether a contact between the groin and the waistband of C's underwear occurred. This node, if added, could be placed as a parent of the 'properties of contact…' nodes. The presence or absence of this additional 'contact occurred' node depends on the manner in which the data used to inform the probability assignment is collected. We discuss these two options of data collection below.

Option 1: contact and transfer combined

In this experimental setup the two factors of DNA transfer, and whether a contact occurred are combined within the final results. Consider an experimental design where one area of an exhibit is spiked with DNA and then the item is placed into a bag and transported. The item is then removed from the bag and sampled in an area that was not spiked with DNA. This experiment is repeated some large number of times. The number of times that DNA is observed on a non-spiked area is a combination of the number of times those areas came into contact with

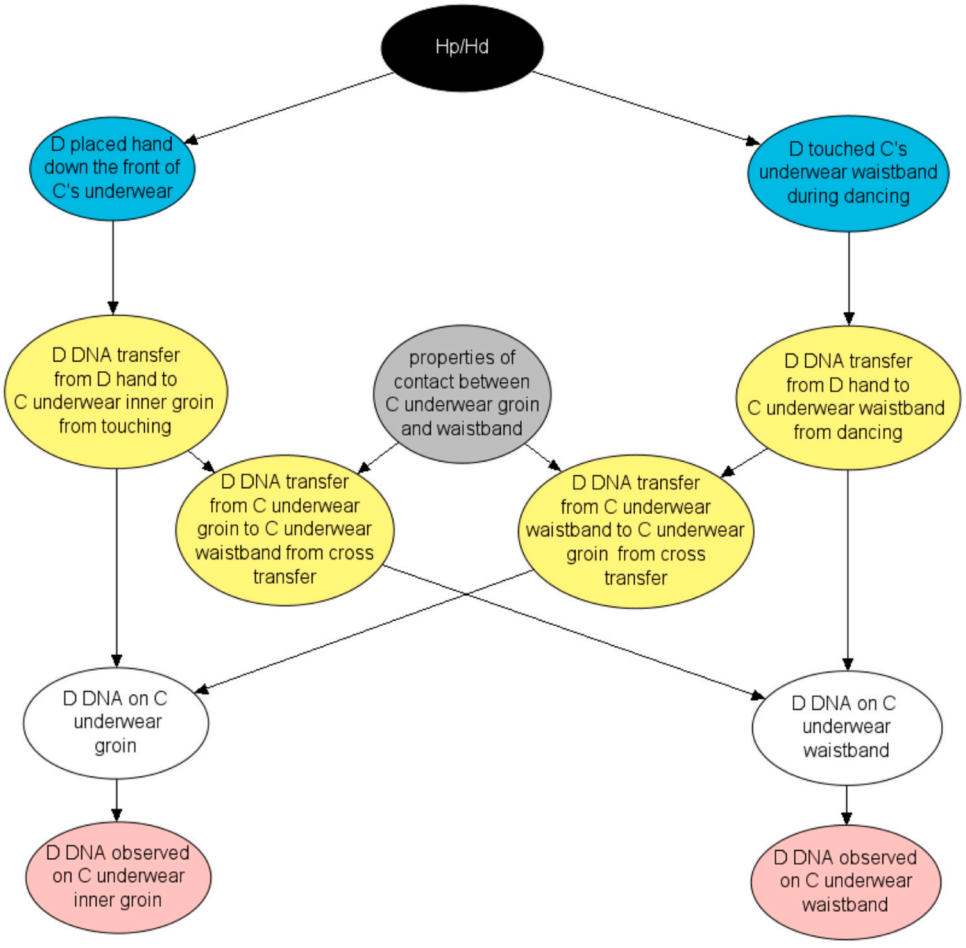


Fig. 3. Extension of the BN in Fig. 2 to consider cross-transfer between the waistband and outer groin of C's underwear. White nodes are accumulation nodes (i.e., those that do not have probabilities assigned based on data and are filled with probabilities of 1 or 0).

the DNA-spiked area and the number of times transfer occurred (and was later recovered) when the contact occurred. Simultaneously, the loss of DNA from the spiking location to the packaging or other contact areas may be measured by sampling this area and comparing the outcome to the amount of spiked material. This is the most common style of experimental setup in current forensic publications [24,30].

Option 2: contact and transfer separate

In this experimental setup the two factors of DNA transfer and whether contact occurred are separated. The probability of DNA transfer can be readily obtained through controlled experiments whereby cloth spiked with DNA is knowingly contacted with clean cloth. More difficult is it to carry out an experiment that would determine how often different areas of cloth come into contact with each other within a bag. Possibly, the probability of contact could be inferred using the probability of DNA transfer obtained from the controlled experiment compared to the probability of DNA contact and transfer together from an experiment carried out as per option 1 (as long as the properties of the contact in the controlled experiments could be carried out in a way mimicking the properties of the contact within a bag). However the experiment is approached, if the probability of contact can be obtained then it is appropriate to have separate nodes for contact, and properties of contact.

2.2. using information about other individual's DNA to assist the evaluation

When using the results of DNA profiling analyses in activity level evaluations, it is often the case that results which are typically considered not useful at the sub-source level, do provide additional discrimination power at activity level. Consider the case example set out in Section 2.1, and that typically the result of finding C's DNA on her own underwear (which she was wearing) would not be considered useful information to assist the court. However, now consider that the presence of C's DNA could be used as a type of calibration as to whether contact between the outside waistband and inner groin occurred. For example, consider that in the DNA profiling that C's DNA was found on the waistband, but not on the inner groin. This may suggest that a transfer from the waistband to the inner groin is less probable than what was considered without that information, i.e., had a contact between these two areas occurred whilst the underwear was in the exhibit bag then we might expect to have seen C's DNA on the inner groin as well as D's DNA. In other words, in a situation where there is a mixture of DNA from C and D on the waistband, we would expect that if a DNA transfer occurred that it would be most probably for both individual's DNA to transfer, or neither, but not one without the other. This, of course, does depend on the relative amounts of DNA of each individual, which we will address later, but for the moment consider them to be in roughly equal proportions. While we have given one example to show when the consideration of additional individual's DNA is important, the concept expands to many configurations of other sources of DNA.

In the current evaluation, the example given above (whilst demonstrating a point) is somewhat artificial, given that high levels of DNA from a wearer of underwear would be expected on both waistband and groin of underwear. However, the same concept can be applied to unknown DNA on an item. There may be a greater expectation of unknown DNA on the waistband of underwear than the inside groin, as the waistband is more frequently touched by the wearer and is generally more exposed than the inside groin. In the same thinking as for C's DNA, the relative levels of unknown DNA on the two areas will affect the posterior probability of DNA having transferred and hence the outcome of the evaluation. This is in contrast to evaluations where the activity is in question (as opposed to the actor) where the presence of unknown DNA is typically irrelevant to the evaluation. We do not go to the complexity of modelling whether the unknown DNA on the two sites are supported as being the same unknown (although this could be carried out as shown in [13]) and make the simplifying assumption for this example that they are. The idea of co-transfer could be extended further, for example to a co-habiting partner, as the needs of the case dictate.

In the extension of the BN to handle additional individual's DNA there are sub-sections of the BN that repeat the same architecture of nodes and arcs. Specifically, there are nodes relating to the transfer of D's DNA from site-to-site, and the same structure of nodes for considering C's DNA (and the same structure would be repeated again for any additional individual being explicitly modelled in the evaluation). These repeated structures (or 'idioms') can be compiled into their own mini-BN called a 'class network'. Then the class network can be utilised multiple times within a main BN to create an object-oriented-BN (OOBN). While not necessary, the advantages of this practice are that it creates a visually simpler BN, but also the creation of the BN is less prone to errors as there is no need to repeatedly populate the same values in conditional probability tables (see section 7.6 of [31] for more information on the theory and construction of OOBNs).

The BN can be extended to consider this new information by including the presence of C's DNA and adding dependencies for co-transfer of C and D as shown in Fig. 4. Note that the formulation of the BN in Fig. 4 makes the assumption that the site-to-site transfer is occurring between two areas of the one object, which consist of the same material (e.g. such as two areas of the one piece of underwear) and therefore the same class network can be used. If the contact is being modelled between two areas that have different surface properties, and different transfer rates (or amounts) are to be applied for the two directions of transfer then the BN would have to be extended to accommodate this (which we do not do in this paper). Another instance in which different surface types may be accounted for is if DNA loss from the item to the inner surface of the packaging is accounted for in the modelling specifically (which would require data to be collected as per option 2. As discussed, we do not extend the BNs in this paper to do so).

The complexity of the BN in Fig. 4 has, necessarily, increased with the additional consideration of C's DNA on the underwear. Comparing to earlier BN architectures (Figs. 2 and 3) the following nodes have been added:

'C wore C's underwear' (node 4): Is an activity node which, like other activity nodes, is a child of the proposition node (node 1) and will have probability values of 0 or 1 assigned depending on which proposition the activity has occurred under. Here the activity is not disputed, hence the probability of the activity occurring is 1 under both propositions.

'C DNA transfer from C to C underwear waistband / groin from wearing' (nodes 10 and 11): These are the new transfer nodes that consider the probability of DNA transfer from C to her underwear from wearing. They will have probability values assigned between 0 and 1, again based on literature and expert knowledge when the parent activity has occurred.

'BG DNA on C waistband' (node 14) and 'BG DNA on C groin' (node 15): These are two nodes that consider the presence of background (i.e. unknown) DNA on the different areas of the underwear. We have chosen not to link these nodes, meaning that the presence of background DNA on one site of the underwear does not affect the probability of background DNA on the other site. This is an assumption that could be informed with a study on background DNA on underwear, tested at multiple sites and if a dependency does exist then this could be built into the BN architecture. Given current literature that could inform of such a dependency [20,32,33], such a dependency has not yet manifested.

'D / U / C DNA transfer site to site on C underwear' (unnumbered): These white, rounded rectangles in Fig. 4A represent instances of an Object Oriented Bayesian Network (OOBN) class (shown in Fig. 4B). They consider the transfer of a specific source of DNA from site to site, given the probability of DNA being transferred from site to site (nodes 8 and 9).

There are a number of points that can be taken from the additional BN architecture:

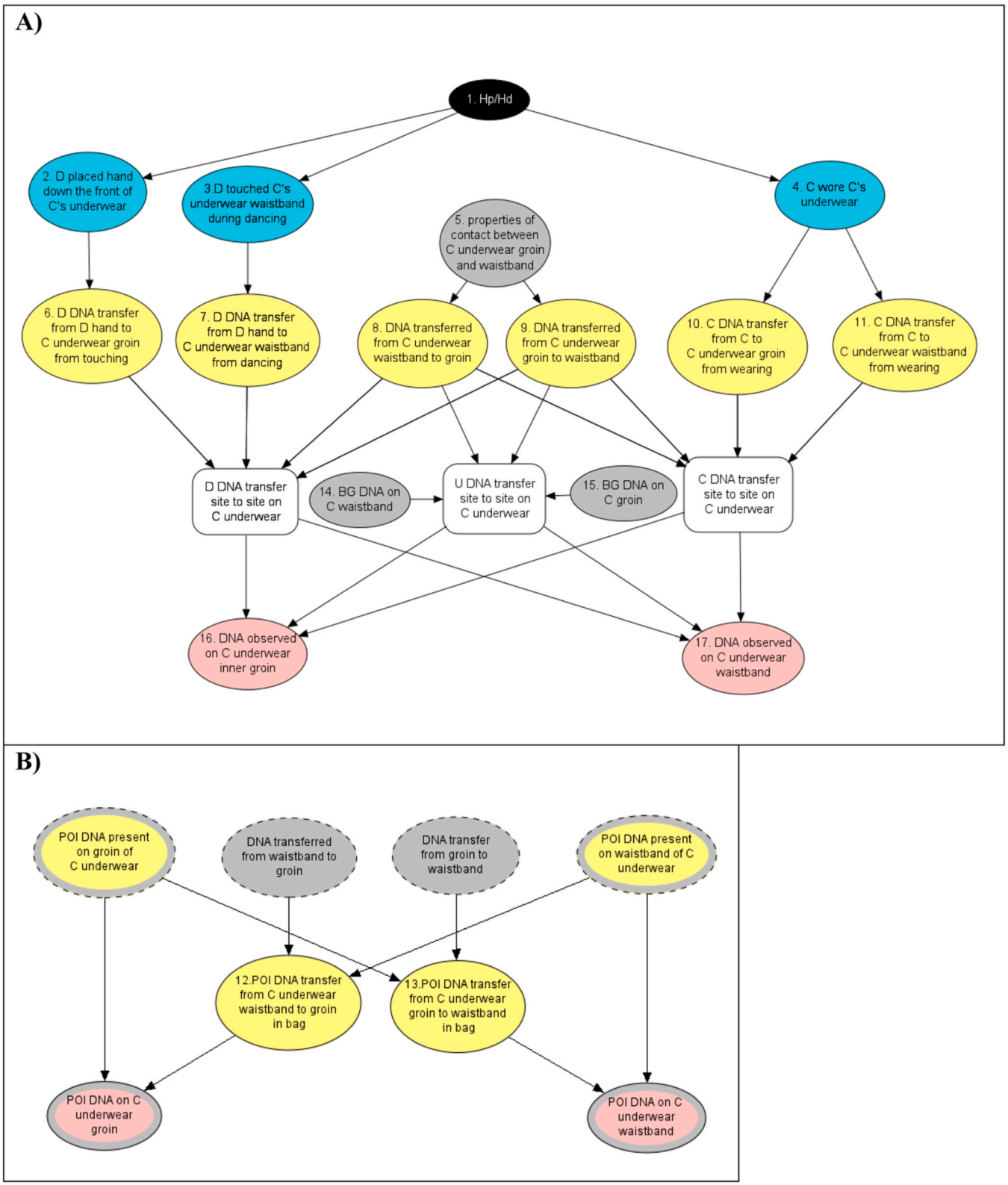


Fig. 4. A) extension of BN from Fig. 3 to consider the presence or absence of C's DNA on the areas of the underwear, with class network (represented by an unnumbered, white box with rounded corners) and B) the class network. POI, U and C correspond to person of interest, unknown and complainant.

- An additional activity node, 'C wore C's underwear' (node 4) has been added to consider the wearing of C's underwear by C. This is not in contention and the activity node itself could be omitted, with the probabilities assigned in lower nodes done under the assumption that C wore her own underwear. While this would simplify the BN, we choose to include node 4 for the reasons set out in [10] i.e., it makes explicit, purely from a view of the BN architecture, the sources of DNA and transfer mechanisms being considered in the evaluation (although the arc between the propositions node and the activity node may be omitted, as it could for node 3).
- With the addition of C's DNA to the evaluation, the results nodes now detail DNA observations rather than the presence of a specific person's DNA (D in previous examples). Therefore, these results nodes will possess states of 'C+D+U', 'C+D' 'C+U', 'D+U', 'C', 'D', 'U', 'none' (if the states are being assigned in a presence or absence of DNA paradigm), or they will contain combinations of DNA amounts of C, D and U (if the states are being assigned in a DNA amount paradigm). If DNA amounts were being used in the evaluation it may be preferable to split the results into six nodes; 'D DNA on waistband', 'C DNA on waistband', 'U DNA on waistband', 'D DNA on inner groin', 'C DNA on inner groin' and 'U DNA on inner groin'. This would avoid an overly large conditional probability table in the results nodes in the case of fine-grained DNA amount modelling.
- The 'properties of the contact' node (node 5) is still a parent to the DNA transfer nodes (nodes 8 and 9), however these DNA transfer nodes now relate to the transfer of any DNA. They are the parents of the transfer nodes that relate to specific individual's DNA. In Fig. 4, node 12 now models the probability that a specific person of interest's (POI) DNA (who in this BN is either C, D or U) has transferred from one site on C's underwear to another, given that there has been some DNA transfer. For example, if the intent is to specify a model where DNA transfer is improbable then nodes 8 and 9 will have a value close to 0 assigned for the 'yes' state. However, once a transfer has occurred then it might be expected that all sources of DNA are likely to transfer together, therefore in node 12 and 13 a value close to 1 would be assigned to the 'yes' state. In other words, the overall DNA transfer probability can be maintained by assignments in the general transfer nodes (nodes 8 and 9), but the dependency between the two separate sources of DNA that are present potentially together on the item (e.g. C and D) in the same area can be accounted for (in nodes 12 and 13).

The consideration of an additional person's DNA on the underwear can occur for any person in the same way that we have demonstrated in Fig. 4 with C's DNA. For example, if the scenario was that shortly before the alleged assault the complainant's partner touched C on the inner underwear (consensually) then the consideration of the presence or absence of the partner's DNA from both sites could be considered in the same way.

2.3. *modelling DNA amounts or DNA presence*

The BN structure seen in Fig. 4 does not dictate the number or format of the states within each node. At the simplest these states can be binary i.e., for nodes involving DNA TPPR the states would be that DNA was present (because it had transferred or persisted) or absent (because it had not transferred or persisted). Ultimately, the states in TPPR nodes could represent DNA amounts, with states that each define a small window of the possible range. When transitioning from a presence/absence paradigm to a DNA amount paradigm there are differences in the way that data is used and the specific meaning of nodes. Specifically, the probabilities assigned to the DNA transfer nodes in a presence/absence paradigm represent a rate of transfer i.e., how often has transfer occurred, whereas in a DNA amount paradigm they represent a proportion of transfer i.e., how much DNA was transferred. While the proportion of transfer could be a single value, a more common way to model a proportion of DNA transfer is using a distribution (such as in [34]).

When modelling DNA amounts, the distribution used for modelling is a crucial point to consider. The distribution chosen should have properties that align with the properties of the factor being measured. In the case of DNA amounts this means a distribution that is bound at lower levels by 0 (i.e. the model cannot allow negative DNA amounts due to the obvious logic flaws that present themselves when translating to real life), can in theory take any value above 0 (although in reality there are upper limits that are sensible to consider) and is likely to be heavily left skewed (i.e. there are commonly many very low amounts of DNA recovered but the occasional high amount). A log-normal distribution fits these criteria, however it is common to transform the DNA amount data so that it is described by a well-known distribution, i.e., the normal distribution. A simple transformation that best achieves this is the log base-10, which is commonly seen in DNA transfer publications. However, when modelling DNA on a transformed scale it is likely that at some point (or multiple points) during the evaluation there will need to be transformations back to natural scale so that accumulations of DNA can occur (i.e. numbers cannot be added on a natural scale while in their transformed state without going back through their natural state in between). In keeping with the previous comment that an ideal distribution has properties of the data it is modelling; a beta distribution is an ideal distribution for transfer proportion data as it has properties that align well with transfer proportions i.e., it only exists between 0 and 1 and can have a skewed shape.

Fig. 5 shows the transfer BN set up for consideration of DNA amounts. There are some additional nodes that have been added in order to make this transition from presence and absence of DNA to DNA amounts:

Nodes 16–21: These are results nodes that have split out the single results node per site (from the BN in Fig. 4A) to results nodes per site and per contributor.

Within the OOBN structure (Fig. 5B) the single node from Fig. 4B that held the probability of DNA transfer have been split into three nodes:

'proportion of POI DNA transfer from [site 1] to [site 2] in bag': This is the proportion equivalent of the rate node (12 and 13 in Fig. 4). It captures the proportion of the total amount of DNA from the POI being transferred. Note that we use generic terminology 'POI' within the class network as the nodes relate to either D, U or C in the overall BN.

'amount of POI DNA transfer from [site 1] to [site 2] in bag': This node combines the information about the proportion of DNA transfer and the total amount of DNA of the POI in order to calculate the amount of DNA of the POI being transferred.

'amount of POI DNA left on [site]': This node considers the starting amount of DNA and the amount of DNA transferred in order to calculate the amount of DNA of the POI left on the original site.

We have chosen to model a transfer proportion separately for both directions. In many cases we could make the simplifying assumption that if the garment is made of one material type that the transfers in both directions should be the same and this would be reflected in the BN by having only a single proportion of DNA transferred from site-to-site node (i.e. amalgamating the current nodes 8 and 9 in Fig. 5). However, by having two separate nodes (even though they will possess the same distribution in our example) it sets up a BN structure that will allow different transfer proportions for the different directions if ever needed (i.e. if the two sites had different surface properties).

In nodes 12a and 13a we allow the transfer proportion for individuals within a mixture to vary from the overall transfer proportion. Nodes 12b and 13b then use the proportion of DNA transferred and the amount of starting DNA to model the amount of DNA transferred from site-to-site within the packaging. Nodes 12c and 13c model the amount of DNA remaining on each site, given that some has transferred.

Finally, in the main network, the results (nodes 16–21) accumulate the starting DNA amount on a site and the amount that has been transferred to that site from another.

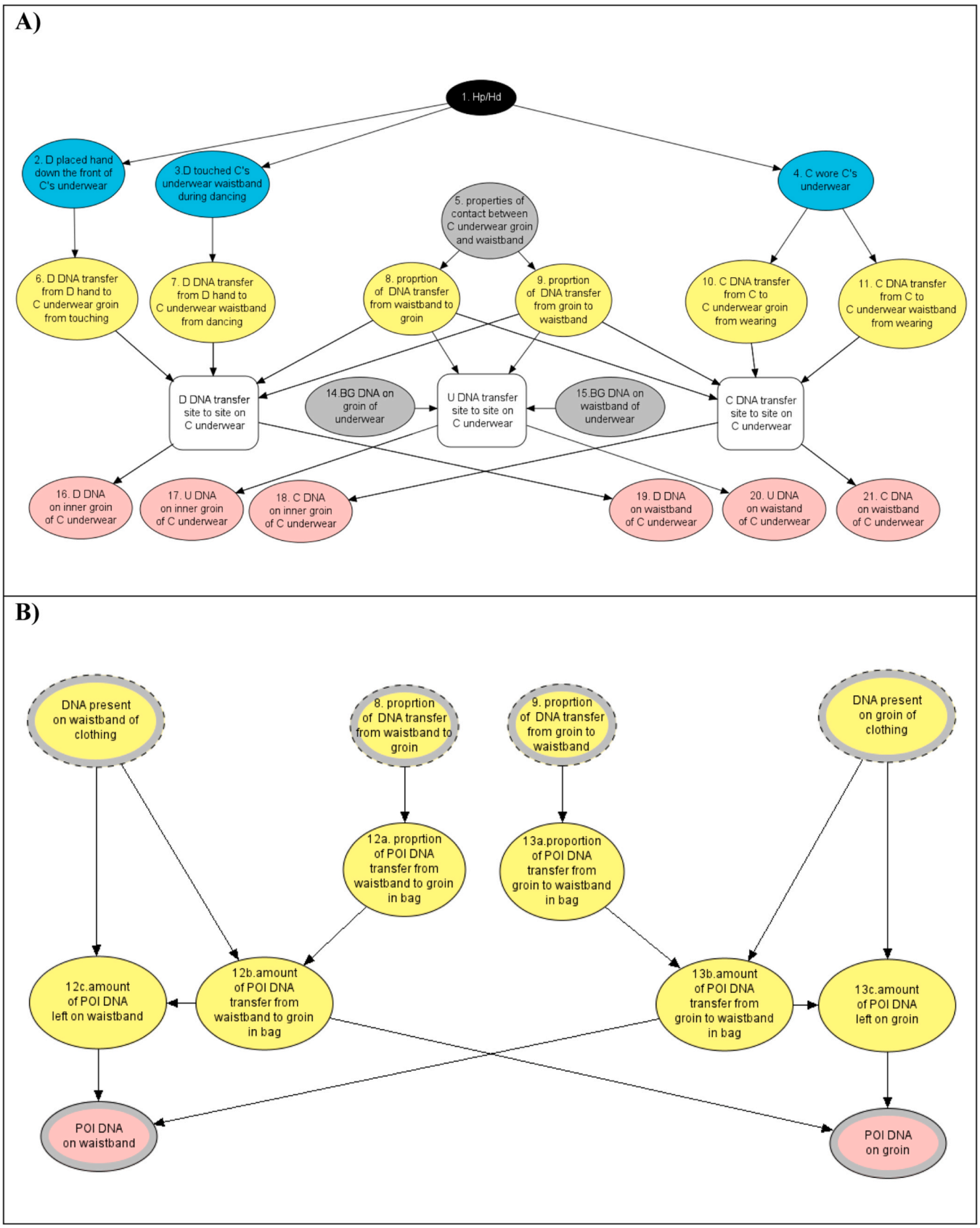


Fig. 5. A) extension of BN from Fig. 4 to consider DNA amounts and B) class network for DNA transfer.

3. Bayesian network method

3.1. proportion of DNA transfer

We use the supplementary material from Gléonec [30] to assign probabilities associated with DNA transfer, and contacts within packaging (as is about to be described).[1] The study found that one of 19 experiments exhibited a detectable site-to-site transfer from waistband to groin of underwear. In this sample an approximate DNA transfer proportion of 0.002 was observed.

To provide more information we turn to the work of Goray et al. [35], who investigated DNA transfer between substrates (importantly to our work including cloth to cloth) under very controlled conditions where contact is known to have occurred. In one of their experiments that is closely relevant to ours, they deposited 50uL of saliva or pure DNA (at a concentration of 5 ng/ul) onto cotton, allowed it to dry, and the cotton was then contacted against another cotton surface in either a passive, pressure or frictional way. The results of their experiment are shown in Table 1 (raw data was provided to us on request by the authors of Goray et al. [35]). Note that we combine the findings from Goray et al. for DNA and saliva samples. Inherent in this combination is the assumption that the two biological materials transfer in approximately equivalent rates. With further research on this point the assumption may be found not to be reasonable, in which case modelling can be done using only the most relevant data.

Also shown in Table 1 are the counts of samples where DNA was detected in the experiments. Note that Goray et al. [35] also carried out experiments on wool, but only using blood and so we do not include those results in this work (however the same process we carried out here could be done for blood if the case circumstances warranted it).

Although the conditions within the packaging do not exactly align with the contact types of Goray et al., here we assume that the passive mechanism is most closely aligned, as it is not likely that the package was subjected to 1 kg of weight, or (in addition to the weight) rubbed together. This may of course depend on the situation faced by a laboratory and the way in which exhibit packaging and handling tends to occur. For the passive results the total number of samples is 7, the number of samples when DNA was noted on the secondary substrate was 4.

Given these findings, even when transfer did occur, a very low proportion of DNA was transferred. The proportion expected to transfer means that none of the DNA samples in the Gleonec study would have recorded transfer, even if contact had occurred, as it would be below the limit of detection of the quantitation system. We therefore only concentrate on the saliva samples for modelling. Let Pr(T|C) be the probability that DNA transfer occurs when direct contact by touching is known to have occurred and let Pr(T|P) be the probability that DNA transfer has occurred within packaging. The probability that contact will have occurred within the packaging can be approximated by:

$$\Pr\left(C|P\right) = \frac{\Pr(T|P)}{\Pr(T|C)}$$

For example, if experiments carried out showed that when two surfaces contacted there was a probability of 0.5 that DNA would transfer, and experiments looking at DNA transfer in packaging found that DNA transferred occurred with probability 0.1 then we could surmise that contact occurred in packaging 0.1/0.5 = 0.2 of the time, and in the 20 % of the time contact occurred, DNA transferred occurred 50 % of the time, leading to the 10 % observation.

From Goray et al. [35] we have Pr(T|C) = 4/7=0.571. From the work of Gléonec [30] we have Pr(T|P) = 1/19 = 0.053, therefore contact within packaging occurs with probability 0.092. Note that in assigning the probability in this way we are making an assumption that the probability of contact occurring is independent of the contact type. This means that we assume that the probability that contact occurs within packaging if the contact is a passive type is the same as the probability that contact occurs within packaging if the contact is a pressure (or friction) type. Intuitively we may feel that if a friction contact type has occurred, it tells us something about the way that the package is being handled and this then affects the probability of contact. We do not explore this further, and at this point make the simplifying assumption that contact type does not affect contact probability, but note this is an area for further investigation.

No data have been published at this time to inform the prior probability assignments for the contact types in node 5. We can assign the value for no contact as the value for 1-Pr(C|P) that was determined. Given a contact has occurred, we here assume that the most common type of contact (when it did occur) would align with the passive contact type in Goray et al. [35], followed by the pressure type and then the friction type. We subjectively assign probabilities seen in Table 2 for node 5. These values may be assigned differently if the circumstances around the transport of the item are known i.e., it may be known that the bag holding the item was stored underneath a heavy exhibit and so friction and pressure may be assigned higher values. Given a contact has occurred we assign probabilities of 0.75, 0.2 and 0.05 for passive, pressure and friction types, and given that the probability that no contact has occurred is 1 - Pr(C|P) then each of these must be multiplied by Pr(C|P) to obtain the values seen in Table 2.

Table 1
Results of Goray et al. [35] study on DNA transfer between cotton surfaces.

		Contact type		
Deposit material		Passive	Pressure	Friction
DNA	Transfer % from original study Mean (sd)	0.03 (0.02)	0.06 (0.04)	0.49 (0.47)
	Total samples	N = 3	N = 4	N = 4
	DNA transferred	N = 3	N = 4	N = 4
Saliva	Transfer % from original study Mean (sd)	0.01 (0.02)	0	0.57 (0.18)
	Total samples	N = 4	N = 4	N = 4
	DNA transferred	N = 1	N = 0	N = 4

[1] To demonstrate the use of these findings we consider the results from the experiments that used brown paper bags and breathable evidence bags combined. We make the assumption here that contact occurring between two sites on clothing is not affected by bag type as long as this type of contact is not mechanically prevented (as is the case with the use of a sterile sheet). We consider that any quantification indicating below 0.0001 ng/μl being present as being an indication that there was no DNA present (at this level even if DNA was present, it is not likely to show any signs of amplified alleles in an STR profile). We only use the results for saliva deposit samples, and only those where the amount of DNA in the original deposit location was above the 0.0001 ng/ μl level.

3.2. Assigning prior probabilities to nodes that model continuous distributions

It is common when using count data to assign probabilities that prior counts of 1 are used for each of the 'I' categories (for example see [36]).

Table 2
priors for different contact properties used in node 5.

5. Properties of DNA transfer	None	1 - Pr(C|P) = 0.908
	Passive	Pr(C|P) x 0.75 = 0.069
	Pressure	Pr(C|P) x 0.2 = 0.018
	Friction	Pr(C|P) x 0.05 = 0.005

If state i of category k has n_{ik} observations, then the posterior probabilities are calculated by:

$$p_{i,k} = \frac{n_{i,k} + 1}{I + \sum_i n_{i,k}}$$

This process has worked well for discrete state nodes, but for nodes with probabilities assigned from a discretised continuous distribution, the practice does not translate well. One of the issues is that the arbitrary choice in the number of brackets used in the discretisation will affect the way the prior acts on the data (with the more brackets used leading to the prior eventually swamping out the modelled observations).

In order to rectify this situation a prior of $1/I$ counts can be added to each bracket (rather than 1), where I is the number of brackets. This has the effect of considering that prior to any observations, you have a belief that if a single experiment were performed that the result could fall into any bracket with equal probability. Therefore, the more brackets you divide the continuous distribution into, the smaller the range that each bracket represents and the proportionally smaller the amount of prior that is added to that bracket.

If the data used to model the probability distribution came from N observations, then the posterior probability for state i, given that $p_{i,dist}$ of the modelled distribution falls within the bounds of bracket i is calculated by:

$$p_{i,k} = \frac{N \times p_{i,dist} + \frac{(i_{upper} - i_{lower})}{\sum_i (i_{upper} - i_{lower})}}{N + 1}$$

Where i_{upper} and i_{lower} are the upper and lower bounds of bracket i, and if uniformly sized brackets are used then:

$$\frac{(i_{upper} - i_{lower})}{\sum_i (i_{upper} - i_{lower})} = \frac{1}{I}$$

for all brackets. As expected, the prior probability has little effect on the probability derived from the observations when the observational probability was high and had a much greater relative effect when the observational probability was low. Also, the prior probability used in the calculation of posterior probability has a greater effect when the number of observations used to model the distribution was low.

3.3. proportion of DNA transfer

We can now model the transfer proportion when contact has occurred by fitting a beta distribution. To do so, we combine the observed transfer from the saliva results from Gléonec [30] with the results from Goray et al. [35]. There are several stages to assigning probabilities in a BN for a DNA transfer proportion. The first part is modelling the observations in two parts; the proportion of the time that no transfer is observed, and the modelling of DNA transfer proportions when it does occur. The latter stages involve discretising the continuous distribution and applying priors, which we talk about later).

In order to calculate the proportion of results where DNA transfer was not observed this is simply the number of experiments where no DNA transfer was observed, divided by the total number of experiments. The distribution is then fit to the remaining observations where a transfer was observed. Beta distributions were fit to data using R V4.3.2 [37], with package MASS V7.3.60 [38].

For the results from Goray et al. [35] (again from the combined DNA and saliva for passive contact type) a beta distribution was fit. We also fit a beta distribution to the DNA proportions transferred to the combined DNA and saliva results for the pressure contact type (but do not include the results from Gléonec [30]), and for the friction contact type (and again do not include the results from Gléonec [30]). A summary of the transfer proportion modelling is given in Table 3 and the fitted distribution are shown in Fig. 6.

We now have all the information required to fill out nodes 8 and 9 in the BN shown in Fig. 5.

To be able to use the beta distribution to populate the conditional probability tables in a Bayesian Network model, they need to be discretised. A decision must be made on the manner in which the proportion nodes will be discretised. There are no specific rules around how to break up a continuous distribution, however in general, the coarser the discretisation the more information loss there will be. In other words, if a very course break-up of proportions was used (for example if the node was broken into two states 0 – 0.5 and 0.5 – 1) then the fine scale information about DNA transfer amount will be lost. On the other end of the spectrum if the discretisation is too fine then this will result in overly complex computation that requires an unnecessarily high amount of computer power, and is likely to represent a level of resolution not supported by the data. A happy intermediate solution is required. For our work we choose to use uneven bracket sizes due to the fact that site-to-site DNA transfer proportions (as seen in Fig. 6) are so low that even breaking the proportion node into 100 even brackets would mean that almost all the probability would be assigned to the first bracket. The brackets we chose are shown in Table 4. We also include one bracket at the lower end of 0–0 that signifies no transfer having taken place.

Putting all of these concepts together the probabilities assigned for node 8 are shown in Table 4. The process for contact type '*k*' involved:

- Determining the proportion of non-transfers that occurred, *Pr(no contact)*
- Modelling the transfer proportions (N of them) when transfer did occur using a beta distribution, $B(\alpha_k, \beta_k)$
- Discretising the beta distribution into brackets, with each bracket gaining a probability of $p_{i,dist} = \int_x^{x+y} B(\alpha_k, \beta_k)dx$, where bracket i spans from x to $x + y$
- Adding the no transfer proportion to the first category
- Applying a prior to the bracket i so that the probability $p_{i,k}$ (as seen in Table 4) is given by $p_{i,k} = \frac{N \times p_{i,dist} + \frac{(i_{upper} - i_{lower})}{\sum_i (i_{upper} - i_{lower})}}{N+1}$

Note that this last step will reduce the 0 proportion category probability from its experimentally observed level by a factor of $N/(N+1)$ due to the fact that the bracket has no width.

3.4. individual's DNA transfer and proportion

If a surface is sampled and subjected to DNA profiling, a mixture of DNA may be obtained. However, this mixture does not reflect the distribution of biological materials of different sources over the sampled surface. As this distribution may be uneven, the proportion of their DNA that is transferred may vary, depending on the areas that are in contact. The less homogenous the distribution over the contacting surfaces, the higher the variation that may be expected in the proportions of DNA of different sources that are being transferred.

In nodes 12a and 13a we allow the transfer proportion for individuals within a mixture to vary from the overall transfer proportion. In order to do this, we specify parameters that given a value, allow some distribution either side. Regardless of the values coming into nodes 12a and 13a, the distribution of transfer proportion values must be within the range [0,1] and so we model the individual transfer proportion (given an incoming total proportion) using a beta distribution, $B(\alpha_{12a}, \beta_{12a})$. Given an incoming transfer proportion value from node 8 (or 9), (which we will denote as μ_8, and call a 'mean' as we will treat it as a mean in further modelling), the distribution in node 12a is modelled maintaining a mean of μ_8, by varying parameter α_{12a} and setting:

$$\beta_{12a} = \alpha_{12a}\left(\frac{1}{\mu_8} - 1\right)$$

Table 3
Summary of DNA transfer modelling used in nodes 8 and 9.

Contact type	Data source	Number of samples	Proportion of no transfer	Observed transfer proportions	Modelling distribution given transfer occurred
None	[30], [35]	NA	1	NA	NA
Passive	[30], [35]	N = 7	0.43	0.000342884 5.33617E−05 0.000497025 0.000456191, 0.002	B(1.213, 3000.002)
Pressure	[35]	N = 8	0.5	0.000582043 0.000153199 0.001101837 0.000621264	B(2.224, 3515.176)
Friction	[35]	N = 8	0	0.002523432 0.003337257 0.011914167 0.001708122 0.007718255 0.004675383 0.003821462 0.006722195	B(4.286, 847.108)

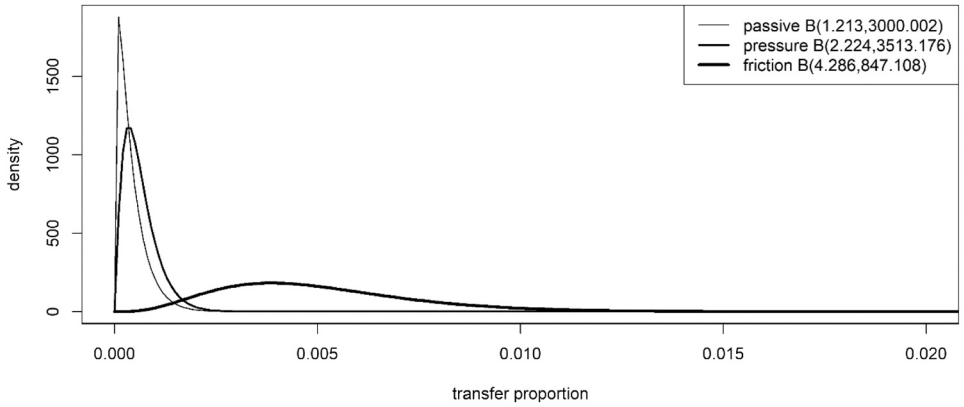


Fig. 6. Beta distributions fitted to Goray et al. [35] transfer proportion data.

Table 4
conditional probabilities for nodes 8 and 9.

Properties of contact (node 5)		None	Passive	Pressure	Friction
proportion	0	1	0.3750	0.4444	0.0000
	0–0.001	0	0.4638	0.3693	0.0061
	0.001–0.002	0	0.0344	0.0709	0.0548
	0.002–0.003	0	0.0021	0.0044	0.1222
	0.003–0.004	0	0.0002	0.0003	0.1583
	0.004–0.005	0	0.0001	0.0001	0.1553
	0.005–0.006	0	0.0001	0.0001	0.1289
	0.006–0.007	0	0.0001	0.0001	0.0957
	0.007–0.008	0	0.0001	0.0001	0.0655
	0.008–0.009	0	0.0001	0.0001	0.0423
	0.009–0.01	0	0.0001	0.0001	0.0260
	0.01–0.02	0	0.0013	0.0011	0.0359
	0.02–0.03	0	0.0013	0.0011	0.0012
	0.03–0.04	0	0.0013	0.0011	0.0012
	0.04 – 0.05				
	…				
	0.09–0.1				
	0.1–0.2	0	0.0125	0.0111	0.0111
	0.2 – 0.3				
	…				
	0.9 – 1.0				

The higher the value for α_{12a} the smaller the variance of the distribution, and hence the amount by which the expected proportion of any individual's DNA can vary from the expected total DNA transfer proportion. Note that we could construct parent nodes for 12a and 13a which define the value of α, and which can be instantiated, but we have not done so in this construction. Doing so would allow the effect of the choice for α on the evaluation (a sensitivity analysis) to be explored by the user immediately through instantiation. Fig. 7 shows an example of $\mu_8 = 0.05$, for three values of α_{12a}, 2, 20 and 200.

When considering transfer of DNA where the source is a mixture, the amount that any individual transfers relative to the amount of the entire mixture that transfers has currently not been studied, and so at this stage remains unknown. Experiments are needed where the amount of individuals within a mixture are transferred relative each other, as well as additional samples of the sampled areas or their immediate surroundings to provide an estimate of the homogeneity of the distribution of biological sources. Without this information we initially provide the subjective assignment of alpha = 20 based on the belief that we expect the transfer of each donor of DNA during a contact to occur in proportion relatively closely tied to the total proportion transferred (hence not choosing alpha = 2), however we still wish to allow some movement of each individual transfer from the overall transfer (hence not choosing alpha = 200).

The proportion of individual DNA transfer nodes can be calculated by (again, showing the node 8–12 path as an example):

$$\mu_{12a} = \begin{cases} 0 & \mu_8 = 0 \\ B(\alpha_{12a}, \beta_{12a}) & \mu_8 > 0 \end{cases}$$

Given the size of the conditional probability table we do not provide this in the paper text (but see supplementary data), but again there will be multiple beta distributions being used to calculate probability distributions for the node, one for each incoming value of μ_8.

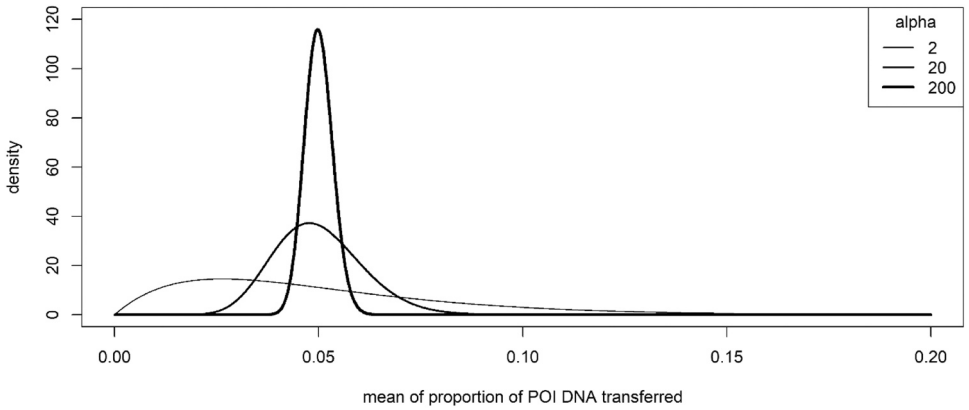


Fig. 7. Graph showing the expected transfer level of any individual's DNA given an expected total DNA transferred value of 0.05 for varying levels of α_{12a}.

3.5. Amount of DNA transferred from activities

We take a step back to nodes in the main network that describe the amount of DNA transferred during activities. Specifically needed are amounts of DNA transfer:

- From hands to inner groin of underwear during attempted digital penetration (node 6)
- From hands to waistband of underwear from touching while dancing (node 7)
- To the waistband of underwear from wearing (node 11)
- To the groin of underwear from wearing (node 12)

Again, a choice must be made for modelling and discretisation. As mentioned earlier, we model the $\log_{10}$ transformed DNA amounts from literature with normal distributions. We will however use a natural scale in the BN, but the brackets will be of unequal size. The reason to do this is to avoid transformation of DNA amounts, which can lead to imprecision (depending on the sampling algorithms that the BN software uses to handle expressions).

To achieve appropriate coverage, the distribution of DNA amounts (on a $\log_{10}$ scale) used that span from -3 to 3. The reason for choosing the limits of -3 and 3 are as follows:

- Lower limit of -3: The limit of detection (when a DNA concentration will be detected at least 95 % of the time) of the quantification instruments is typically around 0.0005 ng/μL (data not shown), which corresponds to approximately 0.05 ng in the DNA extract once multiplied by the extract volume (a typical extract volume can range from 50uL to 200 μL, but we use 100uL for an example). A value of 0.03 on a $\log_{10}$ scale is approximately -1.3, which falls into the -2 to -1 bracket. Therefore, we extend the brackets to -3 to -2 so that any instances of DNA not being detected can be captured in the -3 to -2 bracket.
- Upper limit of 3: This represents an upper DNA amount that is typically seen on high DNA yield items. It is not sensible to consider DNA amounts greater than 1000 ng for the type of exhibits we are considering in this evaluation. If values larger than this were expected (or observed), then the states of the DNA amounts nodes in the BN could be constructed so that the final bracket effectively represented anything greater than 100 ng.

We take the same steps of probability assignment for these nodes as was undertaken for the DNA transfer proportion nodes 8 and 9 i.e. calculate the proportion of no transfer, model the transfer when it does occur, and applying a prior.

To model the amount of DNA transferred from digital penetration we use the work of Ramos et al. [20]. In Fig. 9 of Ramos et al. [20] the results of DNA testing underwear involved in simulated sexual assaults are show. Simulation 2 relates to a digital penetration and from the internal underwear samples shown in the upper right panel of Fig. 9 of Ramos et al. [20], we see that there were 12 experiments performed, of which 9 yielded some DNA being detected. The leads to no DNA transfer 0.25 of the time. For the remaining observations, a normal distribution can be fit to the $\log_{10}$ transferred DNA amounts obtained (across all donors) and this yields a N(-0.89, 0.79) distribution. The calculation steps using this information to obtain probabilities are shown in Table 5.

To model the amount of DNA transferred by a touch to underwear (node 7) we use the work of Daly et al. [39]. In part of their study, they asked volunteers to grab a cloth item. The resulting amount of DNA transferred is given in their Table 1. It is not clear whether there were any samples with no DNA detected as the smallest bracket of results in their table is $<$0.65 ng, which accounts for 0.53 of their findings. We arbitrarily choose to consider that half of this is undetected (N = 26.5), and half is detected, but below 0.65 ng. As exact values are not given for each of the 100 samples, we evenly space the values within each bracket across the range of the bracket for modelling. Modelling their findings (transformed by $\log_{10}$) yields a normal distribution, N(0.00, 0.57). Work by Breathnach et al. [40] also examines the transfer from grabbing, and in the context of underwear, which would make the study more directly applicable to our example case. However, this data could not be used as it was not analysed and presented in a way that provides DNA amounts for the contributors to samples. Instead, total DNA amount is given, and a breakdown of result based on interpretive criteria, such as whether an individual represented a major (or minor) component or whether a profile was interpretable or uninterpretable. While this is a valid manner of analysing and presenting the data, it does not provide the required information to be able to use its findings in a DNA amount-based evaluation.

For the wearer DNA transfer to garments we use the work of Doole et al. [33]. One part of the study by Doole et al. tested the levels of male DNA on underwear worn by female volunteers. Multiple areas of the underwear were tested, including the groin and waistband. Total DNA and male DNA were measured. We use the total DNA component minus the male DNA as the DNA attributed to the wearer. For the groin (node 10) there was 0 / 10 samples where no DNA was detected, and the data was modelled (on a $\log_{10}$ scale) with normal distribution N(1.93, 0.49). For the waistband (node 11) there was 0 / 10 samples where no DNA was detected, and the data was modelled (on a $\log_{10}$ scale) with normal distribution N(0.53, 0.57).

All DNA transfer amounts are given in Table 6.

3.6. Background DNA

Nodes 14 and 15 relate to background levels of DNA expected on underwear in the groin and waistband. We again use the work of Doole

Table 5
steps in calculating the probabilities for node 6.

DNA bracket	DNA bracket bounds (ng), b_i		Expected probability	Adding consideration of proportion of no DNA results, p_0	Expected with p_0 and prior
i	Lower $b_{i,lower}$	Upper $b_{i,upper}$	$p_i'' = \int_{\log10(b_{i,lower})}^{\log10(b_{i,upper})} N(\mu, \sigma^2)$	$p_i' = \begin{cases} p_0 + (1-p_0)p_i'' & i = 1 \\ (1-p_0)p_i'' & i > 1 \end{cases}$	$p_i = \frac{p_i' N + I^{-1}}{N+1}$
1	0	0.01	0.079122	0.309341	0.298366
2	0.01	0.1	0.365402	0.274051	0.265791
3	0.1	1	0.426656	0.319992	0.308198
4	1	10	0.120701	0.090526	0.096383
5	10	100	0.008001	0.006001	0.018360
6	100	1000	0.000119	0.000089	0.012903

Table 6
DNA transfer proportions for various nodes in the BN from Fig. 5.

	DNA (log_{10}) ng	DNA ng	D DNA transfer from D hand to C underwear groin from touching (Node 6)	D DNA transfer from D hand to C underwear waistband from dancing (Node 7)	C DNA transfer from C to C underwear groin from wearing (Node 10)	C DNA transfer from C to C underwear waistband from wearing (Node 11)	BG DNA on groin of underwear (Node 14)	BG DNA on waistband of underwear (Node 15)
Amount of DNA transferred / present	−3 to −2	0.001 – 0.01	0.298366	0.264178	0.015152	0.015156	0.218176	0.161207
	−2 to −1	0.01 – 0.1	0.265791	0.029725	0.015152	0.018713	0.345987	0.426821
	−1 to 0	0.1 – 1	0.308198	0.337076	0.015183	0.174678	0.365010	0.340265
	0 to 1	1 – 10	0.096383	0.337405	0.039932	0.574647	0.040435	0.041229
	1 to 2	10 – 100	0.018360	0.029813	0.493453	0.196956	0.015241	0.015325
	2 to 3	100 – 1000	0.012903	0.001803	0.408189	0.019842	0.015152	0.015152

et al. [33] to inform these nodes. We note that background DNA can be male or female, however the amount of female DNA in the study cannot be used as it relates to the wearer of the underwear. An adjustment was made for this by assuming that background DNA of male and females are in equal amount and multiplying the male DNA only amounts by two. For nodes in Section 3.3, we accounted for a non-event (i.e. non transfer) by removing no-transfer events, modelling the remaining data and then adding the proportion of non-DNA event back. The same modelling occurs for background DNA, although rather than removing non-transfer events we remove samples with no DNA detected. In all other aspects the modelling is carried out in the same way. For the groin (node 14) there was 1 / 10 samples where no DNA was detected, and the remaining 0.9 of the data was modelled (on a log_{10} scale) with normal distribution N(-1.1, 0.60). For the waistband (node 15) there was 2 / 10 samples where no DNA was detected, and the remaining 0.8 of the data was modelled (on a log_{10} scale) with normal distribution N(-0.98, 0.54).

Background DNA amounts are given in Table 6.

3.7. DNA site-to-site transfer and DNA remaining after site-to-site transfer

The distribution of DNA transferred is given in node 12b, DNA_{12b}, (and 13b). Given a starting amount in the 'DNA present on site 1 of clothing' node, DNA_{site1}, and the proportion of DNA transferred for the POI, P_{12a}, the amount of DNA from the POI transferred is obtained by:

$$DNA_{12b} = DNA_{site1} \times P_{12a}$$

Finally, we model the fact that if some DNA has transferred from site to site, then it is no longer on the site it originated, and nodes 12c and 13c model the DNA of the POI remaining by subtracting the amount transferred from the starting amount.

$$DNA_{12c} = DNA_{site1} - DNA_{12b}$$

Note that this type of construction is similar to the general setup of the OOBN described by Taylor et al. [11]. Modelling this loss of DNA amount was not something that could be easily carried out in the presence/absence model BN seen in Fig. 4. This means there is an implicit assumption using the BN in Fig. 4 that not all DNA has transferred (which is likely to be a reasonable assumption to make based on the findings in [35], although there were rare instances of 100 % transfer noted in [24]).

Finally, the results nodes are an accumulation of the amount of DNA remaining at a site of an item (after some has potentially transferred to another site) and the amount of DNA transferred to that site from other locations.

$$DNA_{POI,site} = DNA_{POI\_\text{remaining},site} + DNA_{POI\_\text{transferred},site}$$

Note that in our model we make the assumption that none of the DNA from either site has been lost to other areas of the item, or the packaging. We do not provide table for nodes 12b, 12c, 13b, 13c or the results nodes in the paper, but they are provided as supplementary material.

4. Bayesian network performance

4.1. The effect of different DNA observations

Having constructed the architecture of the BN and populated its conditional probability tables with values, it is possible to explore the effects of DNA profiling results on the LR. In Fig. 8 one possible instantiation is shown, resulting in an LR of approximately 200. The results being considered in Fig. 8 are:

- D DNA on inner groin: 1 – 10 ng
- U DNA on inner groin: 0.01 – 0.1 ng
- C DNA on inner groin: 10 – 100 ng
- D DNA on waistband: 0.1 – 1 ng
- U DNA on waistband: 0.1 – 1 ng
- C DNA on waistband: 100 – 1000 ng

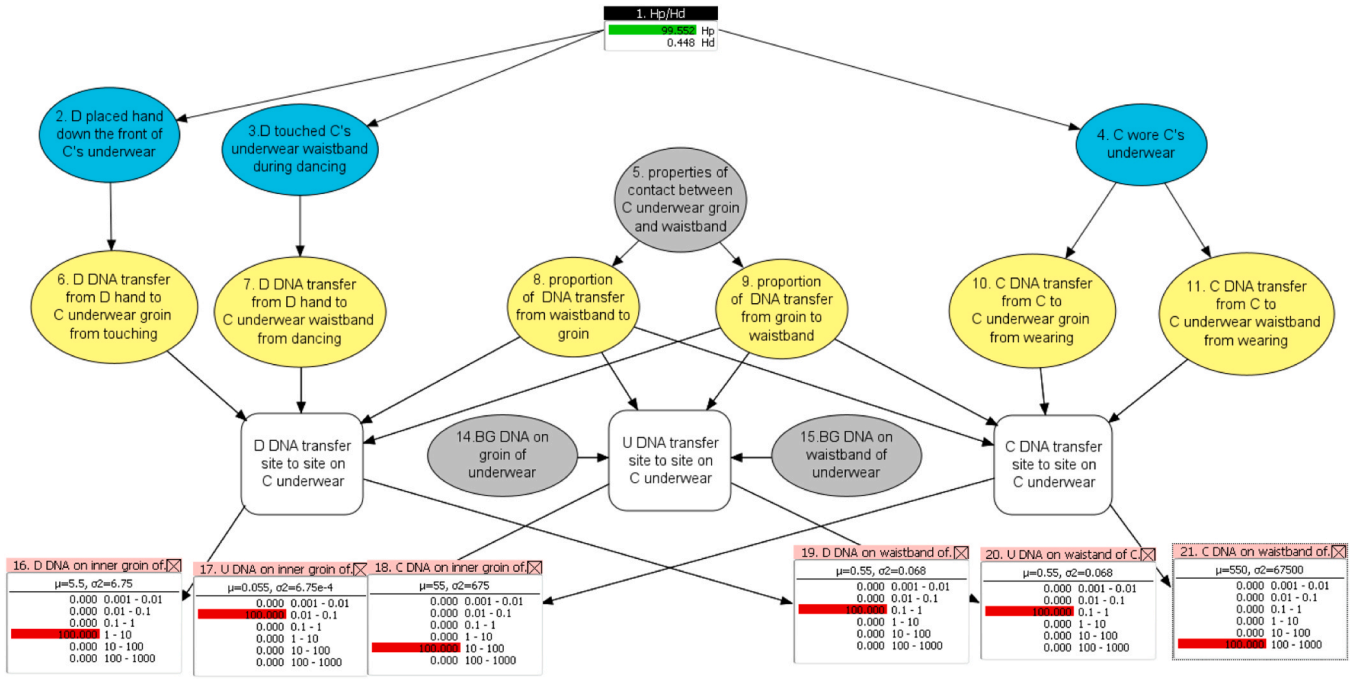


Fig. 8. An example of an instantiation of the BN shown in Fig. 4 when all nodes have been populated with probabilities.

Consider that each of the six results nodes can either not be instantiated, or instantiated with one of six possible DNA amount categories. This leads to 117,648 combinations of instantiation, and hence LRs (this comes from 7 possible states of 6 nodes, $7^6 = 117{,}649$, minus 1 as we do not consider no instantiation of any node as a result). The BN from Fig. 5 was programmed in R V4.2.3 [37] using packages BNleanrn V4.8.1 [41] and gRain V1.3.13 [42], allowing all combinations of results to be investigated.

Fig. 9 shows the DNA results that give rise to LRs up to 10^{10}. It only shows the summary of

46,655 LRs that result from instantiation of all 6 results nodes. Some of the LR values are above 10^{10}, however the area of interest is the transition of the LR from supporting Hd over Hp, to neutral, to supporting Hp over Hd and so we expand that area. Each stacked column in Fig. 9 shows an average value of the DNA amounts across a bracket of LRs, instantiated across the 6 results nodes. The range of $\log_{10}(LR)$ values from -2 to 10 has been broken into 100 brackets. The different colours represent the three sources of DNA being modelled, D, U and C. The height of each component of the stack relates to the amount of DNA that is contributed by that person. Results above the horizontal line represent DNA on the groin of the underwear, while results below the line represent results on the waistband of the underwear.

The results in Fig. 9 show a clear pattern produced by the colours as they range from low to high LR values. At the lower LR range, particularly below the line of neutrality ($\log_{10}(LR) = 0$) a predominance of blue (D's DNA) can be seen below the line (on the waistband), and an absence of blue above (on the inside groin area). As results are considered that move through neutral LRs and into low support for Hp over Hd Fig. 9 shows that the presence of D's DNA on the groin area of the underwear increases (but still retaining a high level on the waistband as well). At these lower LR levels there are also high levels of other sources of DNA (complainant and background DNA) on both regions of the underwear. As results are considered that move into LRs that strongly support Hp over Hd, Fig. 9 shows that the amount of D's DNA increases on the groin, and decreases on the waistband. There is also a steady decline of other sources of DNA on the groin and an increase of other sources of DNA on the waistband.

While Fig. 9 shows the effect of DNA amount on particular locations on the LR, it does not give an indication of the effect the consideration of DNA other than D's is having on the evaluation. In Fig. 10 the spread of LRs is shown for each of the 36 combinations of node states with amounts of D's DNA that can be on the groin or the waistband of C's

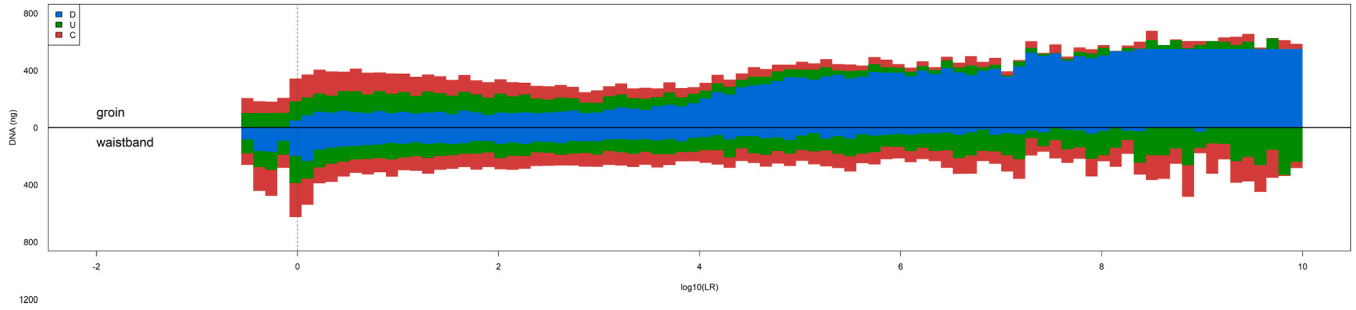


Fig. 9. Column chart summarising the 46655 LRs resulting from different combinations of instantiating all six results nodes. The x-axis has been capped at log(LR) = 5 as beyond this level the points become sparser and do not add to the informativeness of the graph. The position of the columns relate to the log10(LR) obtained from the instantiation. The height of the columns represents the amount of DNA of each individual. Columns above the centre line relate to DNA results on the groin of C's underwear. Columns below the centre line relate to DNA results on the waistband of C's underwear. The DNA amounts shown are averaged across 100 LR brackets from $\log_{10}(LR) = -2$ to 10.

underwear. The spread of LRs within each box in Fig. 10, comes from the different combinations of U and C DNA that can be present for each combination of D's DNA. Also shown in Fig. 10 (with a red dashed line) is the LR that is obtained when only the results of D's DNA are instantiated i.e., when the evaluation carried out without any consideration of other individuals DNA. A large spread of LRs that ranges above and below the red dashed line can be observed for all combinations of D's DNA. The greatest level of underestimation of the LR when considering only D's DNA occurs when there are abundant levels of D's DNA on the groin and very little on the waistband. Conversely, the greatest overestimation of the LR when considering only D's DNA occurs when there is very little of D's DNA on the groin and abundant on the waistband (although noting the overestimation is very slight, and still supports Hd over Hp).

Fig. 10 shows that considering DNA other than D's can have a large effect on the LR, and we can see from Fig. 9 that this sensitivity is about the same to both C or U (by the fact that both C and U show a similar trend across the LR range). This make sense, as any source of DNA other than D's which we consider in the evaluation is having the same effect i.e. providing additional information about the potential site-to-site transfer.

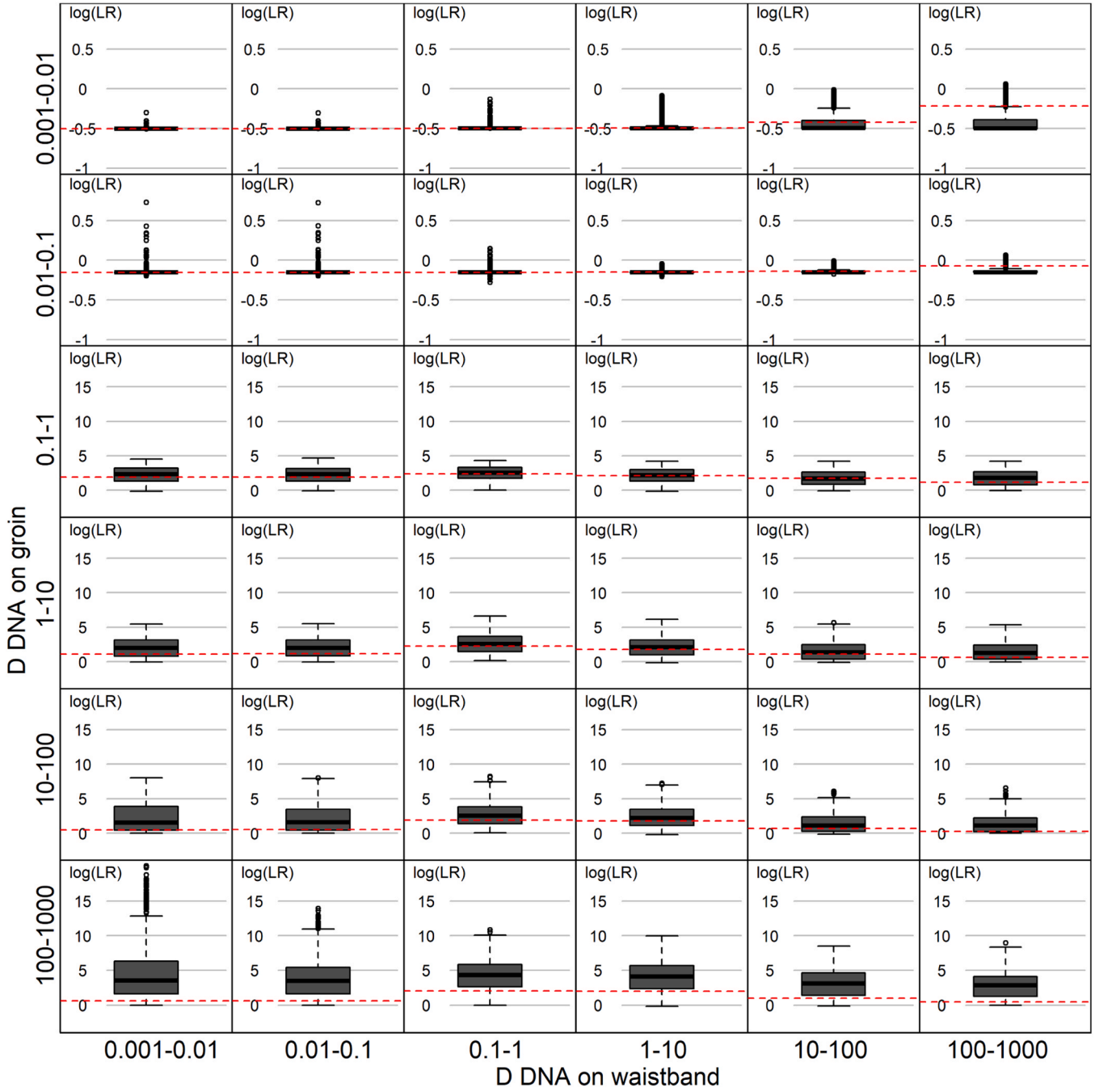


Fig. 10. boxplots showing the 36 different combinations of instantiation that can be carried out for just the presence of D's DNA on either the groin or waistband of C's underwear. The red dashed line represents the log10(LR) obtained when only the results of D's DNA are instantiated. The boxplots shows the spread of log10(LR) values obtained when D's DNA is instantiated (according to the panel in which the plot resides), and all combinations of C's DNA and U's DNA are also instantiated. Note that the first two rows have a different range for the y-axis in order to visualise the spread of results.

4.2. The dependence of the two bi-directional transfers

In the BN shown in Fig. 4 the 'properties of contact' node tie the proportion of DNA transferred in the two directions together, but only weakly. There is still the ability to consider highly imbalanced transfer proportions if the data strongly push the posterior probabilities in that direction. Fig. 11 A shows the mean transfer proportion (i.e. the single value summary of the distribution of values from the proportion of DNA transfer nodes 8 and 9) from waistband to groin compared to that from groin to waistband. Fig. 11A shows the effects of the weak dependence that has been applied. It is not clear from literature whether the transfer in both directions should be tied together i.e., if 10 % of DNA transferred from site 1 to site 2 during a contact do we also expect 10 % of DNA from site 2 to transfer to site 1? And how does differing starting DNA amounts on site 1 and site 2 affect this expectation? If further study reveals that the transfer proportions should be considered highly dependent, then a slight modification of the BN from Fig. 4 can be constructed so that the individual DNA transfers are treated much like different individual's DNA within a transfer are treated now. This would be achieved by making the 'properties of contact' node a parent of a single 'proportion of DNA transfer' node, and then this would be a parent of the current nodes 8 and 9. In this architecture the probabilities of nodes 8 and 9 would then be assigned by allowing a slight variation from the main expectation of transfer proportion (similar to the manner that is graphically shown in Fig. 7).

Another way to investigate the effect of transfer proportion on the evaluation is to consider how different DNA results lead to different posterior probabilities for transfer and then ultimately to different LRs. This is shown in Fig. 11B for the transfer proportions from waistband to groin (the main transfer direction that will be driving the magnitude of the LR). As can be seen in Fig. 11B the greater the amount of D's DNA on the groin of the underwear the more the effect of DNA transfer will have on the LR. If the probability of DNA transfer from the waistband to the groin is very high then the LR is driven towards neutrality, as would be expected by the scenario for this case. When there is very little of D's DNA on the groin of C's underwear then transfer from the waistband has very little effect shifting the LR from its position of neutrality.

4.3. The effect of DNA discretisation

The effect that considering DNA amounts (Fig. 5) has on the evaluation, compared to considering the binary presence or absence of DNA, can be explored by populating the nodes in Fig. 4. We do not go into detail on the assignment of probabilities in the conditional probability tables of Fig. 4, but in general each transfer or DNA amount node from Fig. 5 is reduced to a binary transfer/no-transfer or DNA/no-DNA state. The nodes for the BN considering DNA amounts were all set up so that the lowest categories corresponded to no transfer or no DNA. This meant that populating the nodes in the binary BN is a simple matter of taking the probability value for the lowest category of the node in the BN considering DNA amount and assigning it as 'No' (transfer/DNA) and then assigning the complement to 'Yes'.

The only exception to the above scheme of populating node probabilities is for nodes 12 and 13 in Fig. 4, which have to be set in a subjective way (as they were for their equivalents in the DNA amount BN). In the DNA amount BN we considered that, given a proportion of total DNA had transferred, there could be some variation in the proportion of an individual's DNA having transferred, from the total proportion. This was described by a beta distribution with parameters that were designed to maintain the mean proportion. In the binary BN the situation is different. There is a probability that some transfer will occur in parent nodes 8 and 9, and now in nodes 12 and 13 we must consider the probability that some DNA of the individuals will transfer given that some total DNA has transferred. Without any specific studies to rely on for the probability of transfer we assign an arbitrary value that ties the individual DNA transfer closely to the total but does not force it to occur. The value chosen for an individual POI DNA transfer given the mixture of DNA they are in has transferred is assigned as 0.95. There is one other consideration that must be made for node 12 and 13 and that is the probability that some DNA transfer of a POI has occurred when there is none present on the starting site. At first it may seem obvious that the probability of such an event should be 0 (after all, how can DNA from a POI have transferred if there was none to begin with), however we must take into account that the fact that we may see none of the POIs DNA on the original site because it all transferred to the secondary site (or lost to other locations, but this is mechanism is not explicit in our model). This is automatically taken into account when dealing with DNA amounts when a transfer proportion of 100 % is considered, but this is not so for a presence/absence setup. If no account is made for such an event, then instantiating the BN so that there is a presence of D's DNA on the groin of the underwear and an absence of D's DNA on the waistband will provide complete support to Hp i.e., Hd becomes impossible (testing the BN for such complete support, and addressing its occurrence is suggested by Taylor et al. [10]). Therefore, we assign a small probability of 0.01 to the

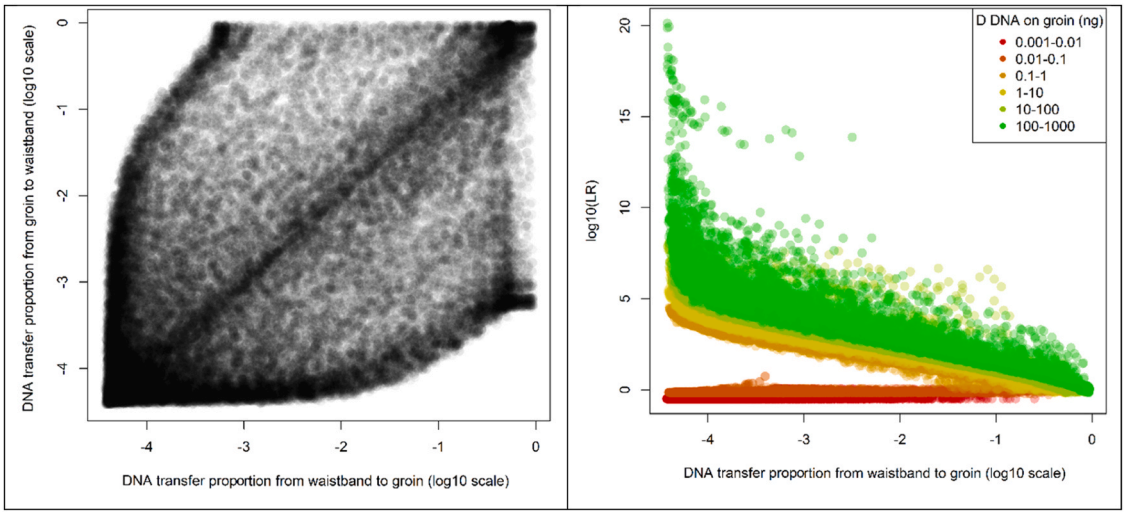


Fig. 11. for all combinations of instantiating all six results nodes A) shows the relationship between the proportion of DNA transferred from waistband to groin and from groin to waistband and B) shows the relationship between the LR obtained and the proportion of DNA transferred from waistband to groin (coloured by amount of D DNA present on the groin).

occurrence of transfer of a POI from site 1 to site 2 when there is no DNA from the POI on site 1. The general structure of the table for nodes 12 and 13 is shown in Table 7.

Having populated the binary BN, a similar exercise can be carried out as that which lead to Fig. 10. The presence or absence of D's DNA only can be considered on the waistband and groin of the underwear, and then the effect of taking into account others DNA can be explored. This is shown in Fig. 12.

The results seen in Fig. 12 for the BN with qualitative states (presence/absence of DNA) are similar to those seen in Fig. 10 for the BN with quantitative states. When considering only the results from D there is limited power in the evaluation to discriminate between propositions. When the presence or absence of DNA from other individuals is taken into account then there is generally an increased ability to discriminate between propositions. While the consideration of D's DNA only leads to LRs of a similar magnitude in either qualitative or quantitative evaluations (red dashed lines in Fig. 10 and 14), the effect that the presence of other individuals DNA can have on the evaluation is much greater when DNA amounts are considered. Particularly when dealing with extremes of DNA amounts the LRs in the quantitative evaluation can reach 20 orders or magnitude, whereas they are limited to 5 in the qualitative evaluation. Therefore, the benefits of both undertaking a quantitative evaluation, and accounting for other sources of DNA can be seen. However, while this demonstrates the increase in discrimination power provided by using DNA amounts, the larger question remains whether LRs of such magnitude can be justified in activity level evaluations. (for example, see commentary of [43,44] who question the appropriateness of providing LRs dealing with sub-source level propositions that reach values beyond what would ever be empirically testable). Whether or not the LR that is calculated is 'robust' needs to be assessed on a case-by-case basis. While small datasets may provide high LRs for findings given activity level propositions (e.g. see [45]), such LRs may or may not be considered robust [46].

5. Discussion

Note that the magnitude of the LR that can be obtained ranges up to approximately 10^{20}. This occurs when there is a very high level of D's DNA on the groin of the underwear and a low level of D's DNA on the waistband. This can be described in two ways by the model; either Hp occurred, or Hd occurred and there was a large transfer of DNA from the waistband to the groin. By itself, instantiating the results of only D's DNA (or only taking account of D's DNA in the model) would bound the LR at the upper range by the probability that such a large transfer could occur. Table 4 shows that the probability assigned to transfers of such a high proportion lie around 0.01. This corresponds to the result seen in Fig. 10 that shows the LR considering only D's DNA (the red dashed line) sit around the log(LR) value of 1–2. This is true even when the highest levels of D's DNA on the groin of C's underwear is observed. However, when adding in the results of other individual's DNA on the areas of the underwear the magnitude of the LR can rise dramatically above its previous bound. The reason for this rise is that if large amounts of other DNA is observed on the waistband of the underwear and very little is observed on the groin, it makes the probability of a large transfer extremely unlikely. This is a result of tying the transfer proportions of individual DNA components together as shown in Fig. 7.

We note that when such high levels of support are obtained in an activity level evaluation, alternate pathways should be considered that are often omitted due to having negligible effects on the LR. For example, in the example BN shown in Fig. 4, a limiting factor to the size of the LR may be the possibility of laboratory contamination, or chance matching DNA between the defendant and the background DNA. The former would tend to set a bound to the activity level LR to the inverse of the probability of contamination occurring (as we are dealing with only a single item, see [31] Chapter 6, case 1, for an example of including contamination in an BN), and the latter will be bound by the level of support for the defendant being a donor of DNA to the DNA profiles. In this latter case it is possible to build the chance matching probability into the BN, and hence the activity level evaluation (for an example of this see [13]). As our example is simplified (as was discussed in Section 2.2), other mechanisms for transfer (e.g. indirect transfer of DNA through the hand of the complainant to the underwear when taking them off) have not been considered which would be essential to include in casework, regardless of the LRs that are obtained with a simplified model.

The results in this study demonstrate some expected behaviours of the models. For example, as more detailed information is provided to an evaluation (in this case the use of DNA amounts relative to the presence or absence of DNA) then there is more ability for that information to discriminate between propositions. Additionally, as more sophisticated modelling of the results is used (in this case the addition of modelling the transfer of multiple individual's DNA relative to modelling just a single POI's DNA) then again, the additional information provided by the model is better able to discriminate between propositions. In a more case-specific sense the behaviour of the model is intuitive, with regards to providing higher probabilities for transfer from larger DNA amounts on the donating surface to smaller amounts on the receiving surface, and higher probabilities for co-transfer of multiple individual's DNA when it is on the donating surface, as opposed to a large differential in the individual's transfer rates.

When the level of support for one of the propositions is extremely high with a simplified model, then other factors (which were not considered in the simplified model) will tend to limit the strength of support (such as the probability of laboratory contamination, or background DNA that matches one of the references). The extent of the simplification of a model needs to be considered on a case-by-case basis. In a casework situation it is crucial to consider all mechanisms of transfer of biological material that are relevant given the set of information on the case circumstances.

6. Conclusion

For some cases it can be important to consider site-to-site DNA transfer. Particularly when there is no other reasonable Hd alternative transfer route (i.e. omitting the possibility becomes an issue of non-exhaustive propositions). There are also instances where the defence may claim site-to-site transfer as a possibility. Much like when contamination is claimed to be the potential cause of a result, these actions (site-to-site transfer, or contamination) do not make up part of the propositions, but rather are taken into account as part of the evaluation of results (and the structure of the BN). Clearly there must also be a scenario where the presence or absence of DNA at different locations on an item are considered informative to the propositions. Given this conditions, site-to-site transfer could always be considered in an evaluation, but in many cases other factors in the evaluation make for more probable transfer routes and so (for simplicity) site-to-site transfer may be deemed unnecessary to include.

Depending on modelling choices there can be a large effect on the LR. This is both true of the choice of whether to consider DNA amounts (quantitative) vs DNA presence/absence (qualitative) evaluation, and the choice of whether to consider all individual's DNA in the potential transfer (and not just the defendant's DNA).

We have shown several methods of taking into account site to site

Table 7
General construction of conditional probability table for nodes 12 and 13.

Some DNA transferred from site 1 to site 2		Yes		No	
POI DNA present on site 1		Yes	No	Yes	No
POI DNA transferred from site 1 to site 2	Yes	0.95	0.01	0	0
	No	0.05	0.99	1	1

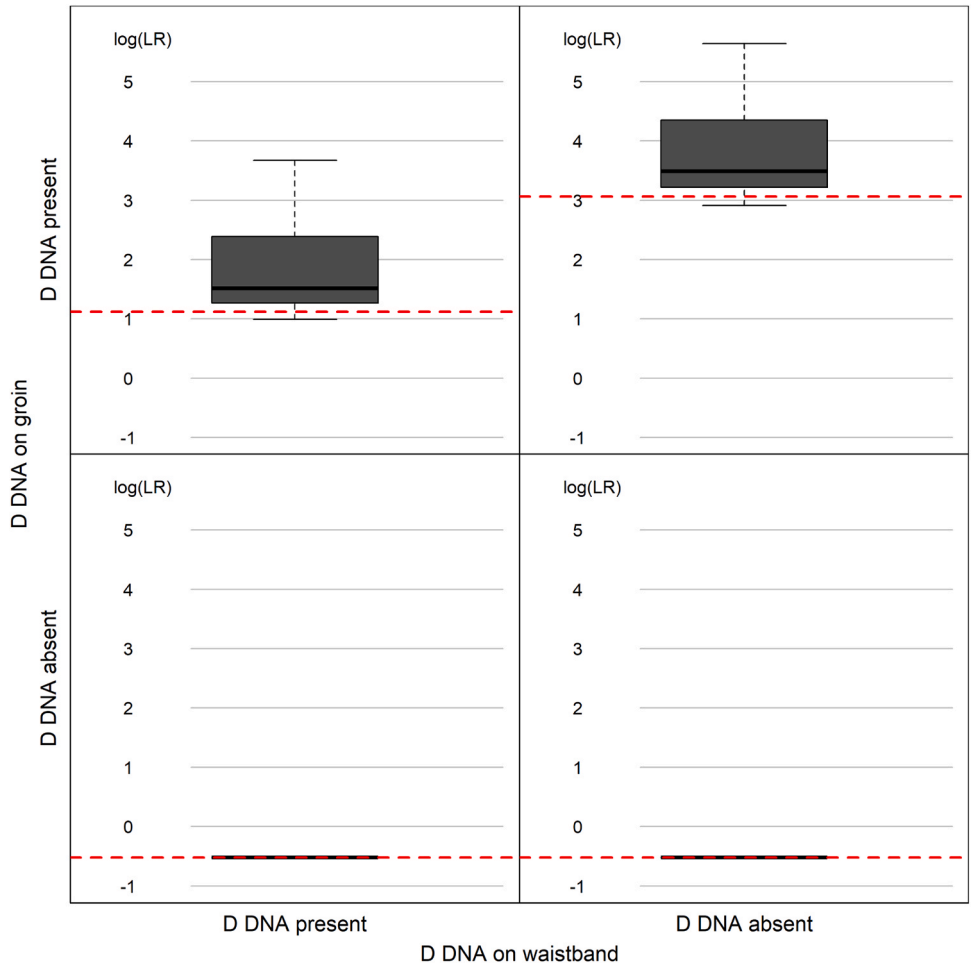


Fig. 12. boxplots showing the 4 different combinations of instantiation that can be carried out for just the presence of D's DNA on either the groin or waistband of C's underwear. The red dashed line represents the log10(LR) obtained when only the results of D's DNA are instantiated. The boxplots shows the spread of log10(LR) values obtained when D's DNA is instantiated (according to the panel in which the plot resides), and all combinations of C's DNA and U's DNA are also instantiated.

transfer within an evaluation, ranging from simplistic to complex. These should suit a variety of case circumstances and evaluation complexities as preferred by a laboratory. The choice of how complex the model should be is one of balance between practicality, data availability and ability. In some situations, there will not be enough data to model transfers in a manner that considers DNA amounts. Either the data will not exist, not be presented in the right format, or will be too sparse. In these cases, there is little choice but for the analyst to consider their evaluation in a DNA presence or absence format. While we have shown that discretising the data in this way leads to lower discrimination power, if the data is not sufficient to begin with then there is nothing to gain in trying to extend the binary model. If data is available to allow DNA amounts to be considered, then we have shown the advantages to discrimination power of doing so. The choice of whether to include the full complexity of the models shown in Fig. 5 will depend on whether the case circumstances warrant it (i.e. it maybe that there are much more probable explanations for the presence of DNA on multiple sites of the item than cross-transfer). It is not always the case that all possible mechanisms for explaining DNA results are included (such as contamination and chance matching profiles as discussed earlier).

One aspect which has not yet been explored in this study is the sensitivity of the LRs to the modelling assumptions and datasets used. Sensitivity analyses seek to determine how much the LR will change based on underlying assumption, approximations, simplification or data choices, and it is becoming more commonplace to see them attached to activity level evaluations [12,36,46]. Modelling assumptions for which the sensitivity of the analysis could be tested include the choice:

- Of how closely to tie individual's DNA transfer together during a transfer of a mixed DNA source (i.e. the choice of alpha in Fig. 7),
- Of whether or not to further tie the transfer in one direction to the transfer occurring in the opposite direction of a site-to-site transfer,
- Relating to discretisation of continuous distributions (see [31] for a discussion on this point), or distributions used to model transfer of DNA amounts
- To combine results of experiments including either DNA or saliva in the transfer modelling shown in Table 1

We hope to explore these sensitivity analyses in future work.

CRediT authorship contribution statement

Luke Volgin: Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Duncan Taylor:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Conceptualization. **Bas Kokshoorn:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Acknowledgements

We gratefully acknowledge the assistance of Oanelle Gléonec for the discussions about her work and her comments on this paper. The opinions or assertions in this document are the private views of the authors and are not to be considered the official opinions of the authors' organisations.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fsigen.2024.103122.

References

[1] S.M. Willis, L. McKenna, S. McDermott, G. O'Donell, A. Barrett, B. Rasmusson, et al., ENFSI guideline for evaluative reporting in forensic science (available at http://enfsi.eu/sites/default/files/documents/external_publications/m1_guideline.pdf), Eur. Netw. Forensic Sci. Inst. (2015) (available at http://enfsi.eu/sites/default/files/documents/external_publications/m1_guideline.pdf).
[2] P. Gill, T. Hicks, J.M. Butler, E. Connolly, L. Gusmão, B. Kokshoorn, et al., DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence - guidelines highlighting the importance of propositions. Part ii: evaluation of biological traces considering activity level propositions, Forensic Sci. Int.: Genet. 44 (2020) 102186.
[3] J.M. Butler, H. Iyer, R. Press, M.K. Taylor, P.M. Vallone, S. Willis, DNA mixture interpretation: a NIST scientific foundation review (NISTIR 8351-DRAFT), Gaithersburg, Md., U. S.: Natl. Inst. Stand. Technol.: U. S. Dep. Commer. (2021).
[4] R.A.H. van Oorschot, B. Szkuta, G.E. Meakin, B. Kokshoorn, M. Goray, DNA transfer in forensic science: a review, Forensic Sci. Int.: Genet. 38 (2019) 140, 66.
[5] R.A.H. van Oorschot, G.E. Meakin, B. Kokshoorn, M. Goray, B. Szkuta, DNA transfer in forensic science: recent progress towards meeting challenges, Genes 12 (2021) 1766.
[6] Y.J. Yang, M. Prinz, H. McKiernan, F. Oldoni, American forensic DNA practitioners' opinion on activity level evaluative reporting, J. Forensic Sci. 67 (2022) 1357, 69.
[7] M. Prinz, D. Pirtle, F. Oldoni, Global survey on evaluative reporting on DNA evidence with regard to activity-level propositions, J. Forensic Sci. (2024) (n/a).
[8] M. Coble, J.-A. Bright, Probabilistic genotyping software: an overview, Forensic Sci. Int.: Genet. 38 (2019) 219, 24.
[9] P. Gill, C. Benschop, J. Buckleton, Ø. Bleka, D. Taylor, A review of probabilistic genotyping systems: EuroForMix, DNAStatistX and STRmix™, Genes 12 (2021) 1559.
[10] D. Taylor, A. Biedermann, T. Hicks, C. Champod, A template for constructing Bayesian networks in forensic biology cases when considering activity level propositions, Forensic Sci. Int.: Genet. 33 (2018) 136, 46.
[11] 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.
[12] L. Samie, C. Champod, D. Taylor, F. Taroni, The use of Bayesian Networks and simulation methods to identify the variables impacting the value of evidence assessed under activity level propositions in stabbing cases, Forensic Sci. Int.: Genet. 48 (2020) 102334.
[13] D. Taylor, L. Volgin, B. Kokshoorn, C. Champod, The importance of considering common sources of unknown DNA when evaluating findings given activity level propositions, Forensic Sci. Int.: Genet. 53 (2021) 102518.
[14] P. Gill, Ø. Bleka, A. Roseth, A.E. Fonneløp, L.R. An, framework incorporating sensitivity analysis to model multiple direct and secondary transfer events on skin surface, Forensic Sci. Int.: Genet. 53 (2021) 102509.
[15] B. Kokshoorn, B.J. Blankers, J. de Zoete, C.E.H. Berger, Activity level DNA evidence evaluation: On propositions addressing the actor or the activity, Forensic Sci. Int. 278 (2017) 115, 24.
[16] I.W. Evett, P.D. Gill, G. Jackson, J. Whitaker, C. Champod, Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks, J. Forensic Sci. 47 (2002) 520, 30.
[17] F. Taroni, C. Aitken, P. Garbolino, A. Biedermann, Bayesian Networks and Probabilistic Inference in Forensic Science, John Wiley & Sons, Ltd, Chichester, 2006.
[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] M. Zuidberg, M. Bettman, L. Aarts, M. Sjerps, B. Kokshoorn, Targeting relevant sampling areas for human biological traces: where to sample displaced bodies for offender DNA? Sci. Justice 59 (2019) 153–161.
[20] P. Ramos, O. Handt, D. Taylor, Investigating the position and level of DNA transfer to undergarments during digital sexual assault, Forensic Sci. Int.: Genet. 47 (2020) 102316.
[21] A. Ronde, Mv Aken, M. Puit, Poot, Cd. A study into fingermarks at activity level on pillowcases, Forensic Sci. Int. 295 (2019) 113, 20.
[22] A. de Ronde, M. van Aken, C.J. de Poot, M. de Puit, A study into evaluating the location of fingermarks on letters given activity level propositions, Forensic Sci. Int. 315 (2020) 110443.
[23] A. de Ronde, B. Kokshoorn, M. de Puit, C.J. de Poot, Using case specific experiments to evaluate fingermarks on knives given activity level propositions, Forensic Sci. Int. 320 (2021) 110710.
[24] M. Goray, Rv Oorschot, J. Mitchell, DNA transfer within forensic exhibit packaging: potential for DNA loss and relocation, Forensic Sci. Int.: Genet. 6 (2012) 158, 66.
[25] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, Case pre-assessment and review in a two-way transfer case, Sci. Justice 39 (1999) 103, 11.
[26] 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, 90.
[27] 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).
[28] S.K. Anderson, K.G. Olesen, F.V. Jennings, F. Jensen, HUGIN - a shell for building Bayesian belief universes for expert systems, Proc. Elev. Int. Jt. Conf. Artif. Intell. (1989) 1080, 5.
[29] D. Taylor, B. Kokshoorn, T. Hicks, Structuring cases into propositions, assumptions, and undisputed case information, Forensic Sci. Int.: Genet. 44 (2020) 102199.
[30] O. Gléonec, DNA loss and relocation during exhibit packaging and transport of undergarments: faculteit der Natuurwetenschappen, Wiskd, En. Inform. (FNWI) (2023). https://scripties.uba.uva.nl/search?id=record_53004.
[31] D. Taylor, B. Kokshoorn, Forensic DNA Trace Evidence Interpretation: Activity Level Propositions and Likelihood Ratios (1st ed.), CRC Press, 2023.
[32] R. Roersma, J. Storm, J. Joling, B. Kokshoorn, Prevalence of male DNA on female worn undergarments. How small datasets may support robust opinions in activity level evaluative reporting, Forensic Sci. Int. 361 (2024) 112097.
[33] S. Doole, C. Hargreaves, G. Davidson, K. Alexander, M. Boyce, I. Fraser, et al., Examining the transfer of male DNA onto female underwear in simulated non sexual social interaction versus digital penetration to assist in the evaluation of results in casework scenarios, Sci. Justice 63 (2023) 90–108.
[34] 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.
[35] M. Goray, E. Eken, R.J. Mitchell, R.A.H. van Oorschot, Secondary DNA transfer of biological substances under varying test conditions, Forensic Sci. Int.: Genet. 4 (2010) 62–67.
[36] D. Taylor, T. Hicks, C. Champod, Using sensitivity analyses in Bayesian networks to highlight the impact of data paucity and direct future analyses: a contribution to the debate on measuring and reporting the precision of likelihood ratios, Sci. Justice 56 (2016) 402–410.
[37] R. Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria 2013.
[38] W.N. Venables, B.D. Ripley. Modern Applied Statistics with S, Fourth Ed, Springer, USA, 2002.
[39] D.J. Daly, C. Murphy, S.D. McDermott, The transfer of touch DNA from hands to glass, fabric and wood, Forensic Sci. Int.: Genet. 6 (2013) 41–46.
[40] M. Breathnach, L. Williams, L. McKenna, E. Moore, Probability of detection of DNA deposited by habitual wearer and/or the second individual who touched the garment, Forensic Sci. Int.: Genet. 20 (2016) 53–60.
[41] M. Scutari, Learning Bayesian networks with the bnlearn R Package, J. Stat. Softw. 35 (2010) 1–22.
[42] S. Højsgaard, Graphical Independence Networks with the gRain Package for R, J. Stat. Softw. 46 (2012) 1–26.
[43] P. Vergeer, A. van Es, A. de Jongh, I. Alberink, R. Stoel, Numerical likelihood ratios outputted by LR systems are often based on extrapolation: When to stop extrapolating? Sci. Justice 56 (2016) 482–491.
[44] L.A. Foreman, I.W. Evett, Statistical Analysis to Support Forensic Interpretation of a New Ten-Locus STR Profiling System, Int. J. Leg. Med. 114 (2001) 147–155.
[45] H. Johannessen, P. Gill, G. Shanthan, A.E. Fonneløp, Transfer, persistence and recovery of DNA and mRNA vaginal mucosa markers after intimate and social contact with Bayesian network analysis for activity level reporting, Forensic Sci. Int.: Genet. 60 (2022).
[46] D. Taylor, B. Kokshoorn, C. Champod, A practical treatment of sensitivity analyses in activity level evaluations, Forensic Sci. Int. 355 (2024) 111944.