View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Serveur académique lausannois

1 **Post-print (final draft post-refereeing)**
2 Final version published in:
3 *Forensic Science International: Genetics*
4 Vol. 28, Pages 155–177, 2017
5 http://dx.doi.org/10.1016/j.fsigen.2017.02.008
6
7 **TITLE:**
8 Helping to distinguish primary from secondary transfer events for trace DNA
9
10 **AUTHOR:**
11 Duncan Taylor[1,2], Alex Biedermann[3], Lydie Samie[3], Ka-Man Pun[3], Tacha Hicks[4], Christophe
12 Champod[3]
13 1. Forensic Science South Australia, 21 Divett Place, Adelaide, SA 5000, Australia
14 2. School of Biological Sciences, Flinders University, GPO Box 2100 Adelaide SA, Australia
15 5001
16 3. Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice,
17 University of Lausanne, Lausanne-Dorigny, Switzerland
18 4. Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice
19 and Fondation pour la formation continue UNIL-EPFL, University of Lausanne, Lausanne-
20 Dorigny, Switzerland
21
22
23 **CORRESPONDING AUTHOR:**
24 Duncan A. Taylor, PhD
25 Forensic Science South Australia
26 21 Divett Place
27 Adelaide SA
28 Australia 5000
29 Phone: +61-8 8226 7700
30 Fax: +61-8 8226 7777
31 Email: Duncan.Taylor@sa.gov.au
32
33 **KEY WORDS:**
34 Primary transfer; secondary transfer; Bayesian networks; likelihood ratio; data; activity level
35 propositions.

ABSTRACT:

DNA is routinely recovered in criminal investigations. The sensitivity of laboratory equipment and DNA profiling kits means that it is possible to generate DNA profiles from very small amounts of cellular material. As a consequence, it has been shown that DNA we detect may not have arisen from a direct contact with an item, but rather through one or more intermediaries. Naturally the questions arising in court, particularly when considering trace DNA, are of how DNA may have come to be on an item. While scientists cannot directly answer this question, forensic biological results can help in discriminating between alleged activities. Much experimental research has been published showing the transfer and persistence of DNA under varying conditions, but as of yet the results of these studies have not been combined to deal with broad questions about transfer mechanisms. In this work we use published data and Bayesian networks to develop a statistical logical framework by which questions of transfer mechanism can be approached probabilistically. We also identify a number of areas where further work could be carried out in order to improve our knowledge base when helping to address questions about transfer mechanisms. Finally, we apply the constructed Bayesian network to ground truth known data to determine if, with current knowledge, there is any power in DNA quantities to distinguish primary and secondary transfer events.

INTRODUCTION:

DNA profiling evidence is commonplace in the courtroom for a variety of criminal offences. Powerful tools [1-7] exist that can help evaluate questions of whether or not the DNA of a person of interest (POI), or a combination of DNA from multiple POIs, is present in a particular DNA extract. Increasing so, with the advent of ever more sensitive DNA profiling systems and laboratory hardware, the value of such queries about DNA is being questioned. In fact, what is being questioned is not the reliability of the DNA profiling results, nor the evaluation of the DNA profile, but rather the significance of those findings in relation to how they support different activity-level propositions suggested by, for example, the competing assertions of the prosecution and defence. Such propositions are described as being activity level within the concept of the hierarchy of propositions [8]. One common question is whether the DNA that has been detected from an examined item was deposited by being directly handled (known as a primary transfer event) or whether there was an intermediary object that acted as a vector to transfer the DNA from the POI ultimately to the item in

question (known as a secondary transfer event). Naturally, scenarios that explain the presence of DNA on an object need not be limited to primary and secondary transfers, and there has been literature that documents instances of tertiary and even quaternary transfer events [9]. The mode of transfer by which DNA came to be on an item has profound implications on the way the DNA profiling results are considered in light of questioned activities. In order to assess the biological results given the alleged activities one needs to understand the factors that affect DNA transfer and persistence on differing target surfaces. Also required is knowledge of case specific details such as the amount of genetic material the POI is likely to shed, and the alleged activities (e.g., timing, type of contact with the objects).

To help address questions of transfer and persistence there have been numerous publications that consider transfer rates under varying conditions [9-14]. Often these studies replicate conditions of a specific case, or are very specific to the hardware and wetware used. This means that it can be difficult to apply their findings in a probabilistic manner to situations that are somewhat removed from those used in the study. We believe this may stem from the fact that researchers design studies and describe results without having a logical framework of interpretation in which to place them. Additionally, it is difficult to separate the factors of transfer to an object, persistence on the object and recovery in the laboratory and so they are often considered jointly, which again complicates the ability to apply the results more broadly to other cases.

Another complicating factor in the evaluation of transfer events is that there is a high degree of variability that exists in seemingly multiple factors, not the least of which is whether the individual is prone to shedding or retaining their DNA [15]. All of these difficulties were presented in a review of DNA transfer by Meakin *et al.* [11], who concluded that, by just the properties of the DNA profile obtained, no definitive conclusions could be made by an analyst as to whether it had resulted from primary or secondary transfer, the order in which individuals had touched an item or whether the DNA detected had been deposited by regular use or a one-off contact (amongst other similar findings). This finding has then been interpreted by many analysts as saying that there is no evaluative information within the DNA trace, implying that, given the findings at hand, any explanation is possible. In response to this interpretation of the Meakin *et al.* findings, Casey *et al.* [16] called for the evaluation of DNA profiling results, in light of questions of activity, to be strived for regardless of the difficulties involved (see also a response to this response from the original authors in [17]).

This is a sentiment to which we agree and it has been influential in our decision to write this paper.

With increasing regularity Bayesian networks (*BN*) are being used to bring together various kind of datasets using Bayesian probability theory in order to help address questions at the source [18, 19] or activity level [20-23]. The graphical ability for *BN* to represent complex underlying calculations makes them ideal for addressing the issue of this paper i.e., the combination of biological results with the framework of circumstances that surrounds an activity in order to help address questions regarding the mechanism by which some DNA came to be on an item. We attempt to construct a *BN* in a manner that makes it adaptable to a wide range of situations. We do so by breaking apart considerations of transfer, persistence and recovery of DNA. In this research paper we have adopted a model that details many variables that are at play. We acknowledge that different models are possible, including ones with a less detailed account of the variables.
In the data collection section we review the findings of relevant literature to determine, firstly, which factors have been found to have an effect on trace DNA transfer and persistence and secondly what those effects are. In the Bayesian Network section we propose a structure for a *BN* that captures understanding and domain knowledge derived from published data, and then inform conditional probability tables with data wherever possible. In the application to different case examples section, we demonstrate how the *BN* developed and parameterised in the previous sections can be applied to several examples that the authors have encountered during testimony. Finally, in the Application to Controlled Case section, we study the performance of our BN on real results generated from known deposition events.

Preliminary considerations

A great advantage of Bayesian networks is that they help advance thinking. A crucial step will be the definition of the variables: in forensic science this typically involves formulating the propositions and the results to be assessed. This may seem obvious, but it is not [24]. This is because results need to be communicated, therefore summarized to some extent. But, on the other hand, they have to be considered in sufficient detail, so that differences may be observed depending on the proposition. We know, for example, that one can observe transfer of DNA in different scenarios and that the quantity of DNA (or the relative quantity of each contributors in presence of mixtures) varies. This is also true of other types of trace material, such as glass or fibres that have been used for many decades to help discriminate between

138 activity level propositions [25]. One key element that has been highlighted for such traces is
139 that it is not transfer per se that is of interest to forensic scientists, but how different the
140 results are given the alleged activities. If we take an example in glass, it is not the transfer of
141 glass that is key, but the recovery of *only one*[1] large group of fragments on a given garment
142 (i.e., what is called the extrinsic characteristics). Indeed, the probability of recovering any
143 glass in general may be very similar given the two activities, but the recovery of a large group
144 will not. With the increase of sensitivity, absence and presence of DNA is not sufficient to
145 discriminate between primary and secondary transfer. DNA results need therefore to be
146 described more precisely in terms of quantity and/or quality, to show which extrinsic
147 characteristics help discriminate alleged activities. This approach has been used, for example,
148 in the Weller case [26], where scientists have considered the probability of their results (in
149 terms of quantity of DNA and positioning) given the propositions describing competing
150 activities.
151
152 Before describing the studies that are available to inform our knowledge on transfer, we
153 would like to mention two last important points: first, answers to questions regarding activity
154 level propositions are probabilistic in nature and that no experiments will tell us whether
155 transfer was primary or secondary. Following this, whether transfer was primary or secondary
156 is, ultimately, for the Court to consider given all the available information. Notwithstanding,
157 because of their specialised knowledge, forensic DNA scientists can help the Court by
158 evaluating their results (e.g., a recovered quantity of DNA leading to a profile of this quality)
159 *given* the alleged activities (or given primary versus secondary transfer).
160
161 The above idea is often misunderstood [58] and it is worth expanding upon it. Any opinions
162 provided on the more probable mechanism of transfer, given the DNA amounts, is a comment
163 on the posterior probability, and as such, requires one or more prior probabilities. This is then
164 must be the remit of the Court, who has access to non-scientific information that will be used
165 to develop their prior belief. The BNs developed in this work always possess a parental node
166 that has states used to signify the positions of prosecution and defence. These states require
167 prior probabilities and for these we use a uniform distribution (i.e. all states are *a priori*
168 equally probably). This of course will not be the position of the Court of the jury, but allows
169 them to use the numerical value of the *LR* obtained from the BN as a 'belief updater'.

[1] By indicating 'only one', we also consider the absence of other glass that can be as important as the presence of material.

170

171 There are other parental nodes that require prior information (such as the DNA on hands or
172 Extraction Efficiencies), however these are the remit of the analyst as they have a scientific
173 basis and could not be expected to be informed by a lay person during a trial.

174

175

176 **DATA COLLECTION:**

177 We present here a series of findings from investigations regarding trace DNA transfer,
178 persistence and collection. For each publication we attempt to provide the raw data that we
179 use to inform the *BN* developed in the next section. Not all data will be able to be used as
180 many of the findings combine aspects of transfer, persistence and recovery in their ultimate
181 results. Factors such as instrumentation used in laboratory for processing DNA also
182 influenced experimental data. We will attempt to tease apart such interwoven aspects
183 wherever possible, but concede that for experimentation that targets trace DNA this can be
184 difficult to achieve.

185

186 We make a note here that the order in which we introduce these topics may not at first seem
187 intuitive as it does not follow the order of the elements of the BN, nor the order of laboratory
188 processes. We present the topics in the order in which they are required to model data, i.e.
189 initially extraction/sampling efficiency are presented as they are required in the model for
190 DNA on hands, all of which are then required for modelling transfer of DNA from hands to
191 objects, and so on.

192

193 Presence of DNA already on object:

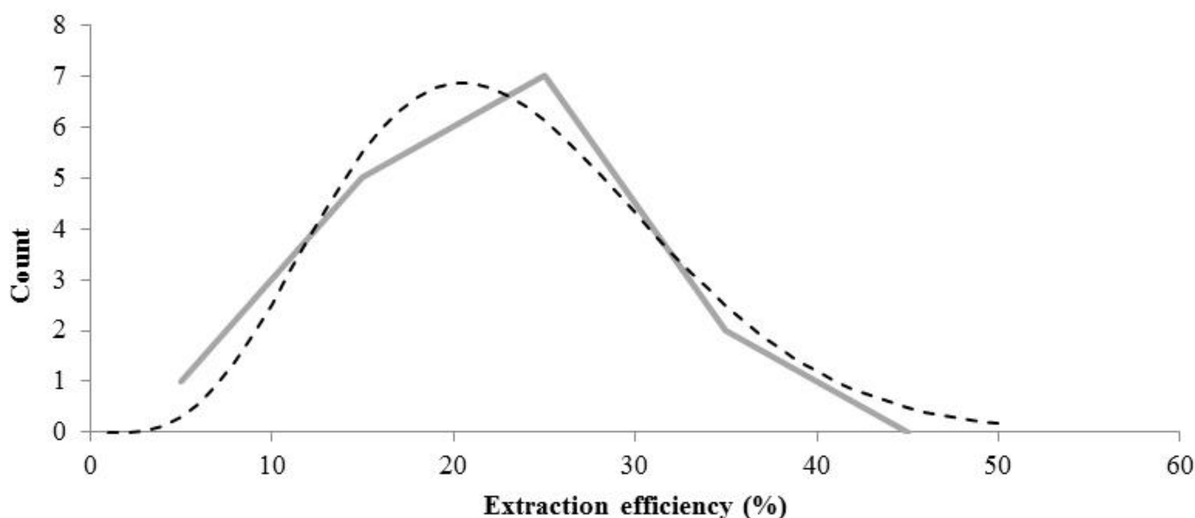
194 Lehmann *et al.* [27] found that, generally[2], the presence of trace DNA on an object did not
195 affect the deposition of further trace DNA and so there was no need to account for that in our
196 *BN*. It was found, however, that the presence of other body fluids could affect the deposition,
197 recovery and detection of trace DNA. However, we restrict our attention to situations where
198 trace DNA only is assumed.

199

200 Extraction efficiency:

[2] If there is DNA from numerous persons, then this will affect our ability to detect the profile of interest, as noted by Lehmann *et al.* 'The presence of several different sources of background DNA created mixed profiles and had major negative influences on the detection of the target source of DNA'.

201 Extraction efficiency is defined as the amount of DNA on a sampling device that is released
202 into a DNA extract and made available for PCR. This is known as 'absolute extraction
203 efficiency'. We use the information presented by Butts [28] who tested two extraction
204 procedures (a salting out method and the Qiagen EZ1 Advanced XL extraction robot) on
205 DNA, epithelial cells and blood. They found that the DNA loss was "*independent of*
206 *extraction method or source of DNA*" and so we show the distribution of combined extraction
207 efficiencies (from the summary graph given in [28]) from their work below in Figure 1. We
208 have modelled the observed extraction efficiencies using a $Beta(5,17)$ distribution. As seen
209 in Figure 1, this fits the observed data reasonably well.
210



211
212 *Figure 1: Distribution of extraction efficiencies determined by Butts [28] (grey) and fitted*
213 *Beta distribution (dashed black line)*
214
215 Differing extraction methodologies (not examined in [28]) may have varying efficiencies and
216 we recognise that the findings we provide may need to be recreated for other extraction
217 methodologies if the analyst wished the *BN* to reflect the properties of their laboratory
218 processes.
219
220 Further improvements to the extraction efficiency node would include:
221 - Trialling extraction efficiencies at low DNA levels. The study of Butts [28] trialled
222 DNA amounts from 24ng to 4800ng

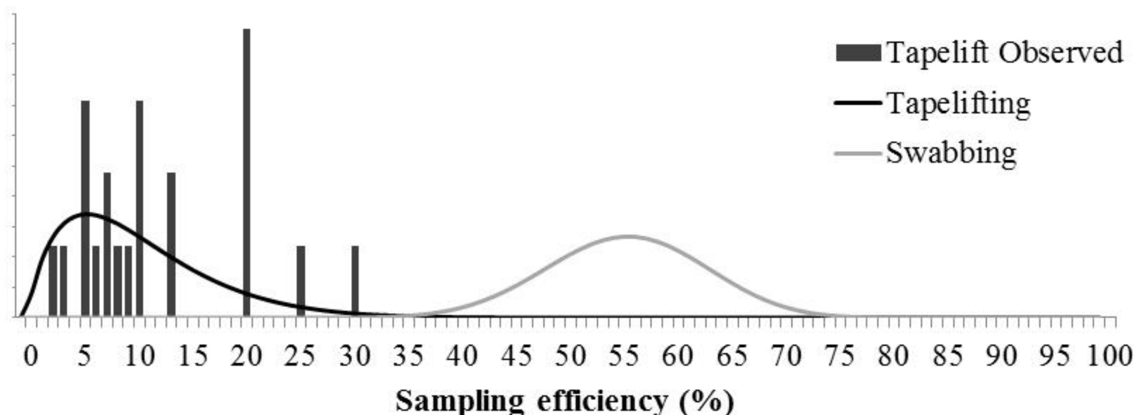
- Trialling of different extraction techniques not tried in the Butts [28] study. Note that some information to this effect can be found in [29] and [30]

DNA sampling/collection efficiency:

We investigate two broad categories of sampling device, tapelifts and swabs. We recognise that both categories could be refined into a number of sub-categories that take into account the type of swab or the type of tape used. It is possible that the variation in sample to sample efficiency (or analyst to analyst differences) may overwhelm the difference in distributions of sampling efficiency from such fine-scale considerations. However, anyone using the BN given in Figure 8 could carry out sampling efficiency validation work to produce findings that are specifically suited to their laboratories process and performance. For the sampling efficiency we used the results of Verdon *et al.* [31]. We define the sampling efficiency as the amount of DNA present on an item that is recovered by the sampling device as detailed further below. Note that the sampling device then goes on to a DNA extraction (typically) and there is a secondary process we consider, the extraction efficiency, which we defined in the previous section. Within the Verdon study tapelifts (using Scotch® Magic™ and Scenesafe FAST™) and swabs (FABswab, Puratin, USA) were used to collect DNA from swatches of cotton, flannelette, Poly/cotton blend and polyester strapping that had contact DNA transferred through vigorous rubbing. The swatches were sampled (either using swabs or tapelifts) and some DNA extracted. The swatch (post collection) was then extracted directly (i.e. not tape-lifted or swabbed further, but instead placed directly into an extraction reaction) and again DNA extracted. The sampling efficiency could then be calculated as the amount of DNA obtained from the DNA extraction of the device to the total DNA extracted from the device plus the swatch. By representing the results as a ratio the effects of the initial amount of DNA deposited and the extraction efficiency are removed from consideration.

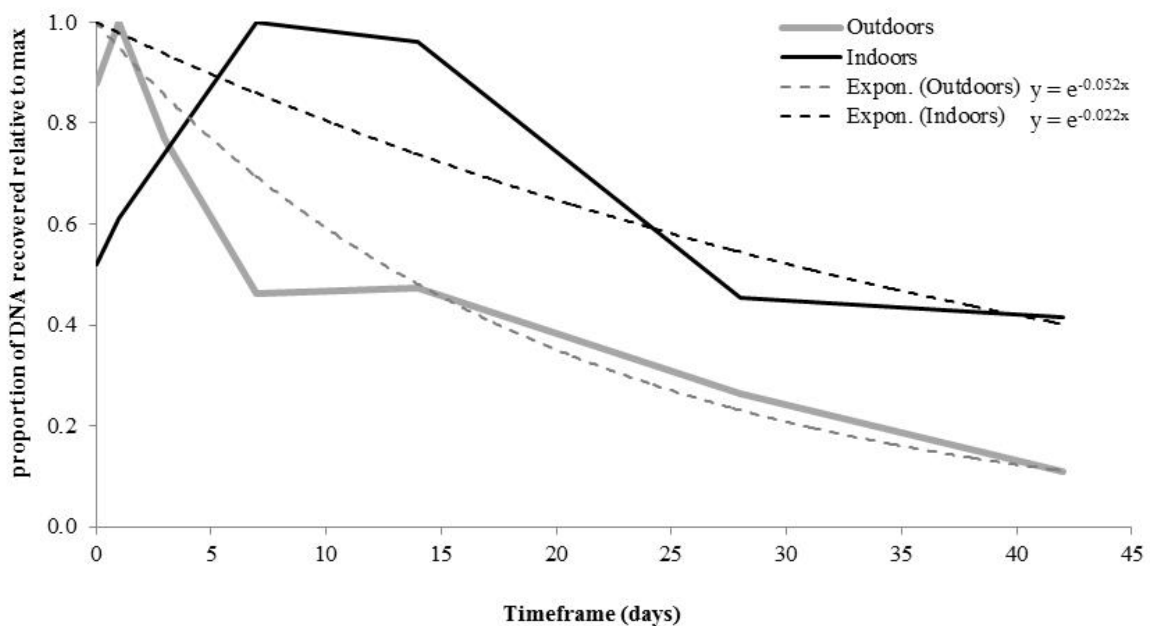
Verdon *et al.* [31] trialled swabs and tapelifts on both smooth (polyester strapping) and standard woven material (cotton, flannelette and Poly/cotton blend). In our BN we assume that swabs have been used on smooth surfaces and tapelifts on rougher, fabric, surfaces and so do not consider the cross-over of collection in our use of the Verdon *et al.* [31] findings. Verdon *et al.* [31] also found a significant difference between the two tapes trialled, and we choose to use the results from the Scotch® Magic™ tape as a more commonly used forensic tape. The findings of the tapelifting of fabrics (combining the findings of cotton, flannelette and Poly/cotton blend) in the Verdon *et al.* [31] study Figure 2, and fitting a Beta distribution

257 by least squares yields an efficiency of $Beta(1.9,16.6)$. For swabbing we use the results of
258 Verdon *et al.* [31] Figure 2 for the swabbing of strapping for which we use a $Beta(25,20)$
259 distribution. These two efficiencies are shown in Figure 2 and in a similar manner as for the
260 extraction efficiency modelling will be incorporated into the *BN*.
261



262
263 *Figure 2: Sampling efficiency of tapelifting and swabbing from results of Verdon et al. [31]*
264 *for tapelifting (black) and swabbing (grey). The histogram shows observed tapelift*
265 *efficiencies. For swabbing there was only one average value given.*
266
267 Persistence:
268 There is little data available on the persistence of trace DNA. There are a number of factors
269 that are likely to affect persistence, such as the surface type, the length of time and the
270 conditions the item is exposed to during the time. The best example of a trace DNA
271 persistence study for contact DNA is the work by Raymond *et al.* [32]. In [32] known
272 amounts of cellular (using buffy coat) and free DNA (using positive control DNA 9947A)
273 was deposited on:
274 - An outdoor window frame
275 - A vinyl bag kept outdoors
276 - Glass slides kept in controlled laboratory conditions
277 The outdoor samples were in partly shaded areas over average temperature and humidity
278 conditions of 24.1^{O}C, 63% humidity (day) and 18°C, 71% humidity (night).
279
280 While the collection and extraction methodology will mean that absolute DNA amounts
281 cannot be used, they are expected to remain an approximately constant factor throughout the

282 experiment of Raymond *et al.* [32]. This allows relative amounts of recovery to be used for
283 persistence. We recreate the results of Raymond *et al.* [32] Figure 1. However, we combine
284 the results of the outdoor bag and outdoor window frame experiments (by averaging) as well
285 as averaging the trends across cellular and neat DNA. The reason behind this is two-fold:
286 1) The data from Raymond *et al.* [32] do not show a noticeable difference between these
287 experiments
288 2) Trace DNA deposited onto an item is likely to consist of both cellular and free DNA
289 [33]
290 For the same reason we average the cellular and neat DNA findings for the laboratory
291 experiment. Finally, we display the results as a ratio relative to the maximum DNA amount
292 observed (because clearly there must have been at least this amount of DNA available at time
293 zero). All of this is shown in Figure 3.
294



295
296 *Figure 3: Data from Raymond et al. grouped into two categories, outdoors (grey) and*
297 *indoors (black) with the trends modelled with an exponential curve (dashed lines)*
298
299 Using the information from Figure 3 we then implement an exponential decay curve in the
300 *BN* for the DNA reduction in samples that are kept in 'poor' or 'favourable' conditions over a
301 number of days, '*t*'.

302 $$Decay = DNA \times e^{-\alpha \times t} \qquad \text{where} \begin{cases} \alpha = 0.022 & favourable \\ \alpha = 0.052 & poor \end{cases}$$

Further improvements to the persistence node would include:

- Taking into account the nature of the surface type the DNA has been placed on when considering persistence
- Further investigation into different environmental conditions (e.g. rain, washed, full sun, etc.) on DNA persistence
- More data to confirm the DNA persistence rates found in Raymond *et al.* [32].
- Data on DNA persistence of DNA on objects after extended handling by other individuals (of which some work has been done in [34]), or from physical movements after initial deposition.

<u>DNA on hands (shedder status):</u>

Initially the *BN* will need to contain information regarding the amount of DNA available for transfer to an object, which is present on an individual's hand. This node encompasses the idea of good and poor shedding of DNA. The idea that individuals may deposit variable amounts of DNA is described in [35] and there has been some debate as to whether the level of shedding that has been attributed to 'good shedders' or 'bad shedders' is a transient property depending on day to day variation, the closeness to last cleaning [36] or that there are simply too many factors to consistently label someone as a 'good' or 'bad' shedder [33]. A recent work [37] suggests that the DNA available for deposition through contact is a mixture of skin cells, free DNA in sweat and sebum and a combination of other bodily fluids present on the individual's hands. Van Den Berge *et al.* [38] show that sebum and sweat contribute to an increase of the quantity of DNA on hands with a lower effect of the sweat compared to sebum. In their work Lacerenza *et al.* [37] swabbed 120 individuals' hands and submitted those swabs for DNA extraction and profiling and RNA extraction for use in body fluid identification. Investigation by Lacerenza *et al.* [37] of a number of factors found that the only significant factor was gender, where males had typically more DNA on their hands than females. The authors attributed this to a difference in general levels of hygiene between the genders. A work by Bontadelli [39] swabbed the hands of 50 individuals and found no difference between males and females.

By analysis of all these findings it seems likely that in reality there are not two distinct groups of individuals, some of which are prone to shedding and others that are not, but rather a

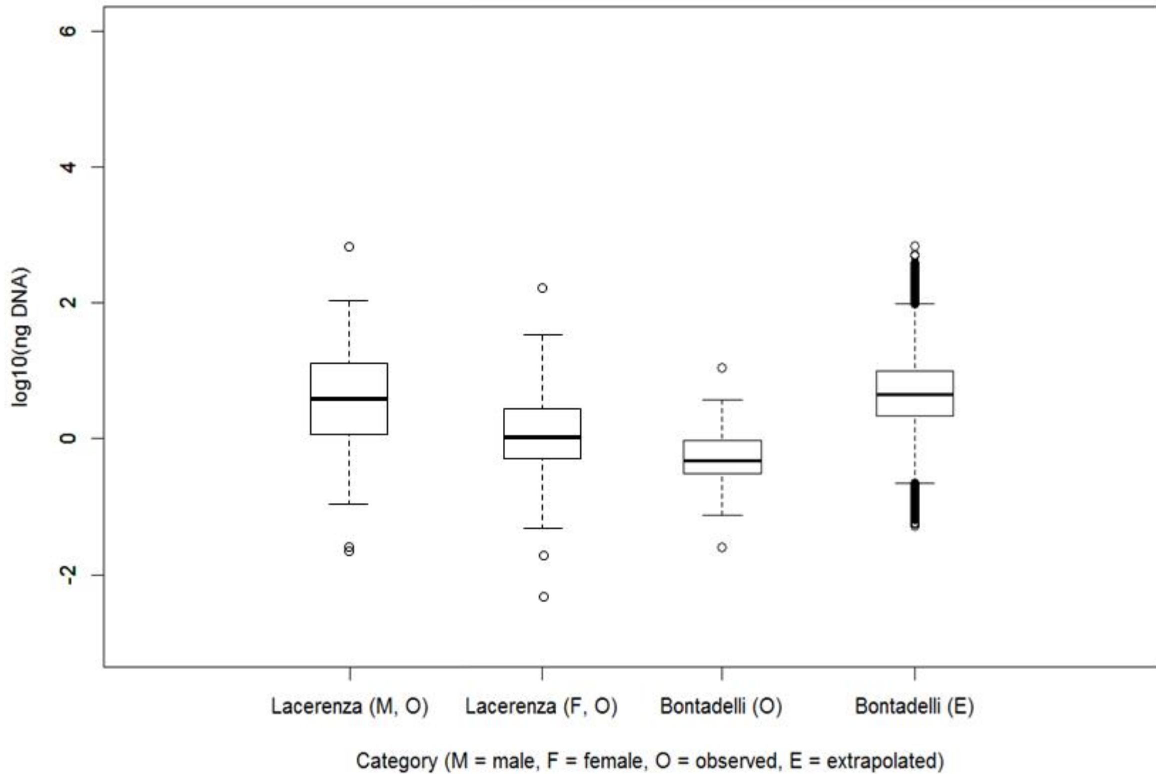
distribution of shedding propensity, on which people will exist at different points. Certainly there are individuals who consistently shed (or perhaps better put transfer) more DNA than others. Like many acts of grouping data in a binary fashion, the designation of 'good' and 'bad' shedder groups has, over time, lead to the belief in two distinct groups of people, when in reality it is simply that the binary classification is an oversimplification of an underlying continuous distribution. The DNA on hands node represents our uncertainty in the amount of DNA on individual's hands, and within this uncertainty exists the propensity of that individual to shed DNA. The node represents a distribution of DNA amounts for a 'random' individual meaning that if the case circumstances indicates some reason that the POI did not behave in a manner similar to a random individual (e.g. had a skin condition, which made them more prone to shedding DNA, such as the well-known case of R v David Butler in Liverpool 2012) then some modification of the distribution would need to be made for them. This could be accounted for in the BN by the use of a parent node to the DNA on hands node (e.g. a 'skin condition' node that would specify one of two possible distributions in the 'DNA on hands' node when instantiated).

The results of Bontadelli [39] and Lacerenza *et al.* [37] have both a sampling efficiency effect and a DNA extraction effect present in the data and the actual amount of DNA available on hands is likely to be higher than the quantification results they obtained. To account for these effects we carry out the following process using the Bontadelli [39] data using the software R [40] as follows:

a) Randomly draw a DNA amount from a normal distribution that describes the distribution of log(ng of DNA) found by Bontadelli [39]
b) Randomly draw an extraction efficiency from the Beta distribution described in the 'Extraction Efficiency' section and use this to adjust the DNA amount in a) to a DNA amount that was present on the swab head
c) Randomly draw a sampling efficiency from the Beta distribution described for swabbing in the 'Sampling Efficiency' section and use this to adjust the DNA amount on the swab head in b) to a DNA amount that was present on the hand of the individual

Carrying out such a simulation for 100, 000 iterations produces the distribution shown in Figure 4, labelled 'Bontadelli (E)', which we model as normal distributions for use in the *BN*. Also shown in Figure 4 are the observed (O) distributions of DNA from the Bontadelli [39] and Lacerenza *et al.* [37] studies. All show a similar distribution.

370



371

372 *Figure 4: Observed (O) DNA amounts on hands from Bontadelli [39] and Lacerenza et al.*
373 *[37] studies and adjusted values obtained by simulation (O) for DNA on hands, based on*
374 *Bontadelli [39] data.*

375

376 The log(ng of DNA) for DNA on hands from the 'Bontadelli (E)' distribution is modelled by
377 $N(0.66,0.49)$.

378

379 Further work in this area could include:

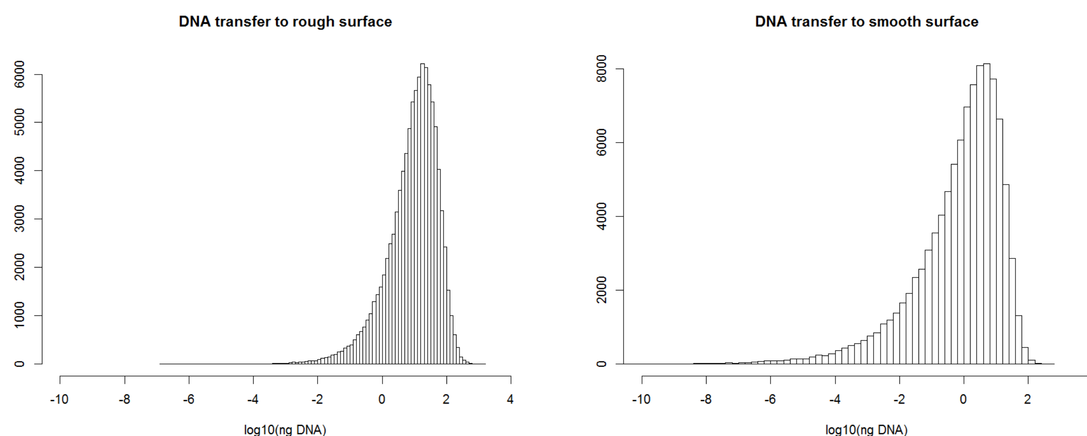
380 - Shedder consistency studies, i.e. whether an individual sheds DNA consistently in the
381 upper or lower quantiles of the population shedding distribution. Studies could extend
382 to a standard method for determining the approximate shedding propensity of an
383 individual for use in primary vs secondary transfer considerations. Some work in this
384 area has been done in [41].

385

386

387 <u>Transfer from hand to object (Primary transfer):</u>

388 There is much literature that presents findings of transfer from hand to object as the results of
389 obtaining full, partial or no DNA profile [10, 42, 43]. While this information is indeed useful,
390 for the current study (and in particular the primary transfer events depicted by the '*DNA*
391 *transferred to object*' nodes) what is required is absolute DNA amounts. For this, we use the
392 data obtained from Daly *et al.* [44]. In their work they asked 300 random volunteers to grasp
393 an 8mL glass vial, a 7×7cm cotton cloth or an 8.5×1.7×3cm piece of wood for 60 seconds.
394 They then tapelifted with Minitape (WA Products Ltd., UK) and extracted using Qiagen®
395 QIAamp DNA mini kit.
396
397 The proportion of hand surface area that contacted the items is not known and so we assume
398 this proportion to be one. We make this assumption so that the transfer data can be directly
399 compared to the amount of DNA on direct hand swabs (which swab 100% of the hand). We
400 also combine the results from the wood and cotton samples and combine them under the
401 surface type category 'rough' and then use the glass results in the surface type category of
402 'smooth'. We fit gamma distributions to the observed data from Daly *et al.* [44] using least
403 squares. These were $\Gamma(0.64, 3.87)$ and $\Gamma(0.33, 1.75)$ for rough and smooth surfaces
404 respectively (graphs not shown). We then adjust the gamma distributions of DNA amounts
405 observed to model the amounts that were present on the item taking into account sampling
406 and extraction efficiency, in the same manner as we did in the 'DNA on hands' section of this
407 paper, to obtain a distribution of DNA amounts transferred by the 300 volunteers to either
408 smooth or rough surfaces as seen below in Figure 5.
409



410
411 *Figure 5: Simulated values for DNA amounts transferred to rough (left) or smooth (right)*
412 *surfaces, based on the results of Daly et al. [44].*

Having produced the two distributions, we are interested in the distribution that describes the decrease from the total amount of DNA present on an individual's hand (seen in the 'Bontadelli (E)' category in Figure 4) to the amount of DNA transferred (Figure 5).

In the Daly *et al.* [44] study there is no account of the type of contact that has been made with the object e.g. a glancing touch, pressure for a short period of time, friction, etc. However studies such as that conducted by Goray *et al.* [45] show that the type of contact is an important factor. In their study Goray *et al.* [45] trialled three different contact types:
Passive – described as a placing two substrates together for 60s
Pressure – described as the same as passive, but applying 1kg of weight
Friction – described as the same as pressure but moving the weight around for the 60s

We believe that the category of friction from the Goray *et al.* [45] study is best aligned with the experimental setup of Daly *et al.* [44]. The results given in the Goray *et al.* [45] study are given as mean percentage of DNA transfer. We take the results from Table 3 of that study and scale all findings between smooth (plastic) and rough (cotton) so that the friction category has a value of 1. We show the results of this data manipulation in Figure 6.

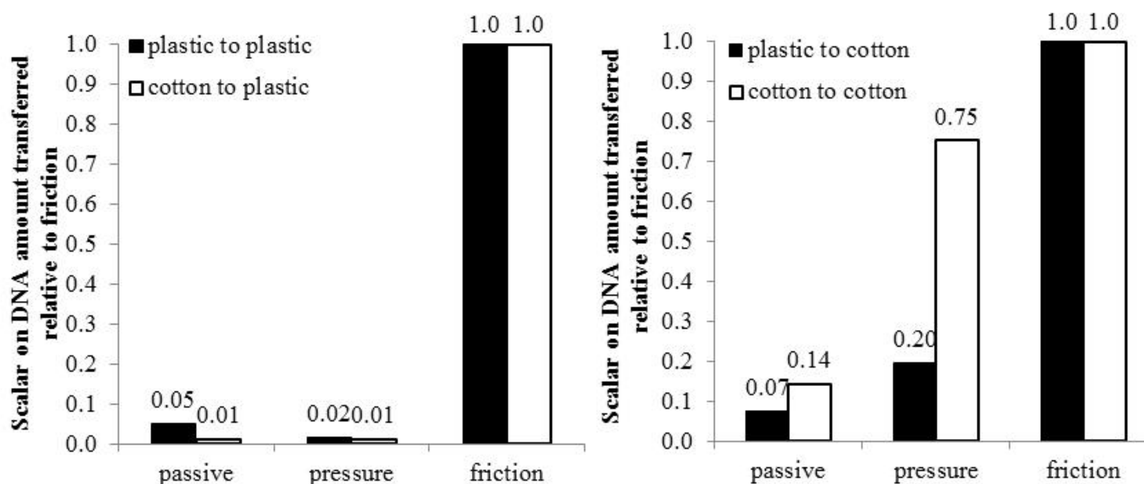


Figure 6: Effect of pressure type on amount of DNA transferred. We break the data into two groups; DNA transferred onto smooth objects (left) and DNA transferred onto rough object (right).

437 We model the proportion of DNA deposited by a simplified comparison of distributions.
438 Using software R we go through the following steps:

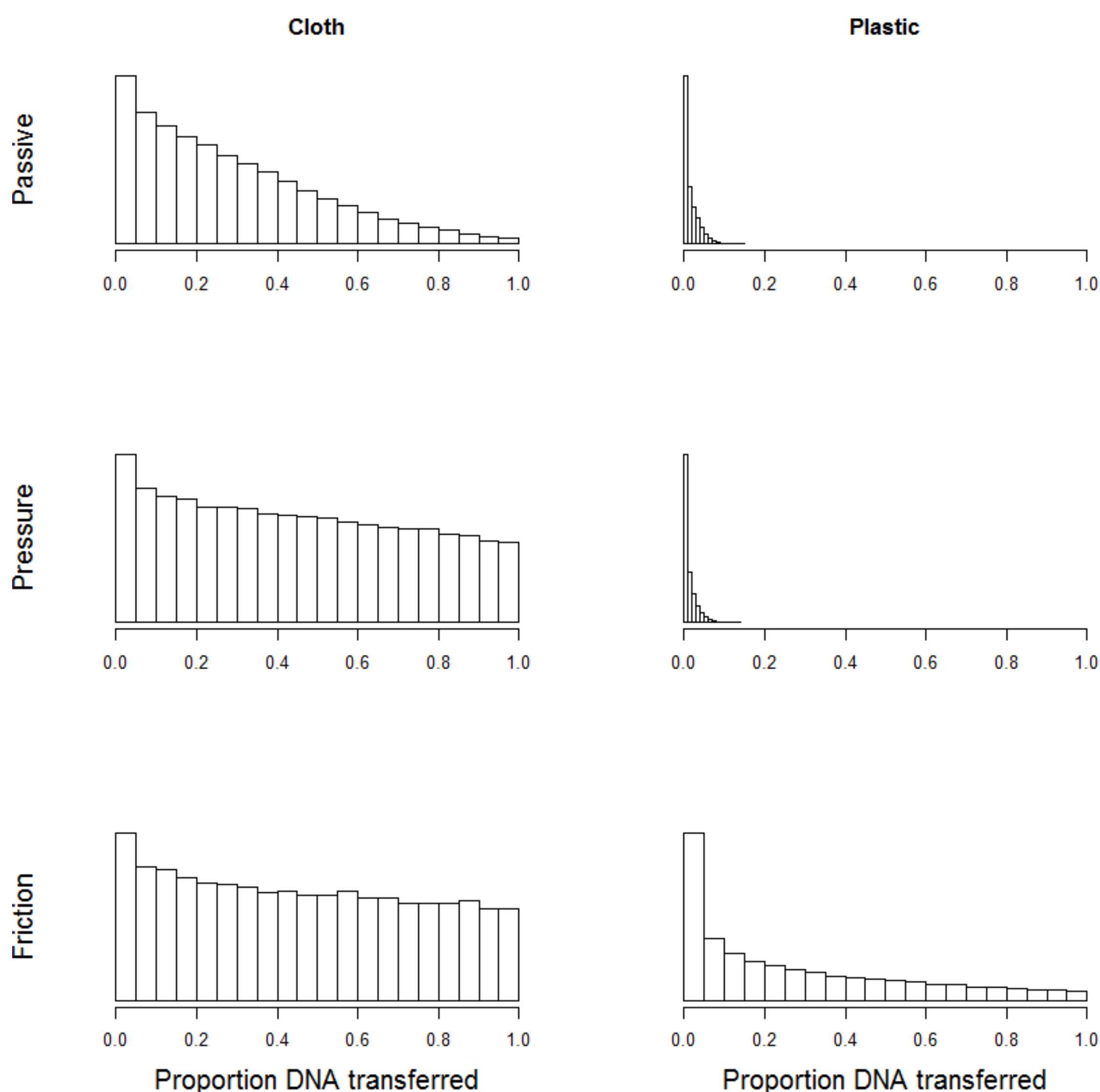
439 a) Generate an array of 100,000 variables drawn from the extrapolated DNA amount on
440 the presence on hands (seen in the 'Bontadelli (E)' category in Figure 4).
441 b) Generate two arrays of 100,000 variable drawn at random from the distribution of
442 values seen for deposition onto smooth and rough surfaces as shown in Figure 5.
443 c) Order the arrays generated in a) and b) and generate two arrays of the proportion of
444 DNA transferred from hands to object by dividing the values in the DNA of rough or
445 smooth object array by the corresponding entries in the DNA on hands array. The
446 result is two arrays of values between 0 and 1 which represent the proportion of DNA
447 transferred from hands to smooth and rough objects[3].
448 d) Each of the values seen in Figure 6, has a level of data variability in the Goray *et al.*
449 [45] study given as a standard deviation. We use these standard deviations (scaled
450 down to align with the values seen in Figure 6) and draw values from the distributions
451 of the reduction factor for passive, pressure and smooth from both rough and smooth
452 surfaces. These values are then multiplied by the reduction values from c) to produce
453 distributions for the reduction in DNA from hand to rough or smooth object for either
454 a passive, pressured or frictional contact type.
455 e) The resulting distributions for transfer type are seen in Figure 7. Beta distributions
456 were fit using MLE to these distributions to represent the proportion of DNA
457 transferred. For the distribution to cloth with a pressure or friction contact we fit a
458 mixed beta distribution, which is given in Table 2.

459

460 With these distributions available, we have all the information required for the '*DNA*
461 *transferred to object*' node of the *BN*.

462

[3] Due to the stochastic nature of the data and simulation there were a number of values within the 'proportion of DNA transferred' arrays that had values greater than 1. This is equivalent to greater than 100% of DNA from hands being transferred to an item and hence is nonsensical. In these instances the data was truncated at 1 to obtain the sensible values seen in Figure 7.



463

464 *Figure 7: Proportion of DNA transferred from hands to smooth (plastic) or rough (cotton)*
465 *objects when contact is passive, pressure or friction.*

466

467 The age of the donor may also have an influence. The study of Poetsch [46] shows that there
468 could be a correlation between the quantity of transfer DNA, the quality of the profiles
469 coming from fingerprints and the age of the donor. On 209 child and adults, a full DNA
470 profiles is obtained in 75% of cases with children under 11 (47 children), 9% with teenager
471 between 12 and 20 (32 teenagers), 25% with adults between 21 and 60 (81 adults) and 8%
472 with persons older than 60 (49 seniors). The factor is not taken into account in the BN if the
473 relevant population is an individual between 21 and 60.

474
475 Further work that could be investigated in this area includes:

476 - A study of how the amount of time an object is held affects the amount of DNA
477 transferred.
478 - To date we have found such information in [47] for hand to hand transfers mostly by
479 length of handshake. Van Oorschot *et al.* [48] studied this question using
480 polypropylene tubes held for varying lengths of time (5 s, 30 s, 3 min, 10 min) and
481 found that the length of the contact did not influence the amount of DNA transferred.
482 In contrast Saravo *et al* [49] showed that the quality of the profile is influenced by the
483 length of contact (using steel cable).
484 - A study of the absolute amount of DNA transferred from hands to objects for different
485 contact types e.g. light touch, pressure, friction (i.e. so that data from multiple studies
486 does not need to be combined and extrapolated as we have done here).

487
488 <u>Secondary transfer from object to object:</u>
489 For this final section we again turn to the work of Daly *et al.* [45]. Again we use the results of
490 Table 3 where dried contact DNA is transferred from object to object with varying primary
491 and secondary substrates and different contact types. Table 3 from the Goray *et al.* [45] study
492 gives percentage transfer, mean and standard deviations for all considered transfer scenarios.
493 We apply these transfer distributions with the obvious restrictions that the transfer percentage
494 is bound by 0 and 100%.
495
496 Further work that could be done in this area:

497 - Consideration of the amount of DNA transferred to object from a habitual use e.g.
498 items in the home. Some work has been done in this area, such as [50].
499 - Transfer DNA amount for varying length of time of contact between primary and
500 secondary substrates and for different types of activities.

501
502 **BAYESIAN NETWORK:**
503 From the literature we have found the following factors to be important to:
504 <u>The amount of DNA available for deposition</u>

505 - Propensity of an individual to shed DNA [37]. We consider this node as describing
506 the amount of DNA on an individual's hands available for transfer. It therefore

encompasses the idea of shedding ability of the individual, and we could consider aspects such as cleanliness, sweating, skin disease at this point.

- Amount of individual’s hand that was in contact with object (no reference, this is based on common sense).

The transfer of contact DNA:

- Surface of object being touched [44, 45].
- Surface of object that DNA is currently on [45].
- Vigour and length of contact [45].

The persistence of DNA on an object:

- Time between deposition and sampling (or further deposition) [32].
- Condition the item is kept in between deposition and sampling (or further deposition) [32].
- The type of surface of the object.

The recovery of DNA:

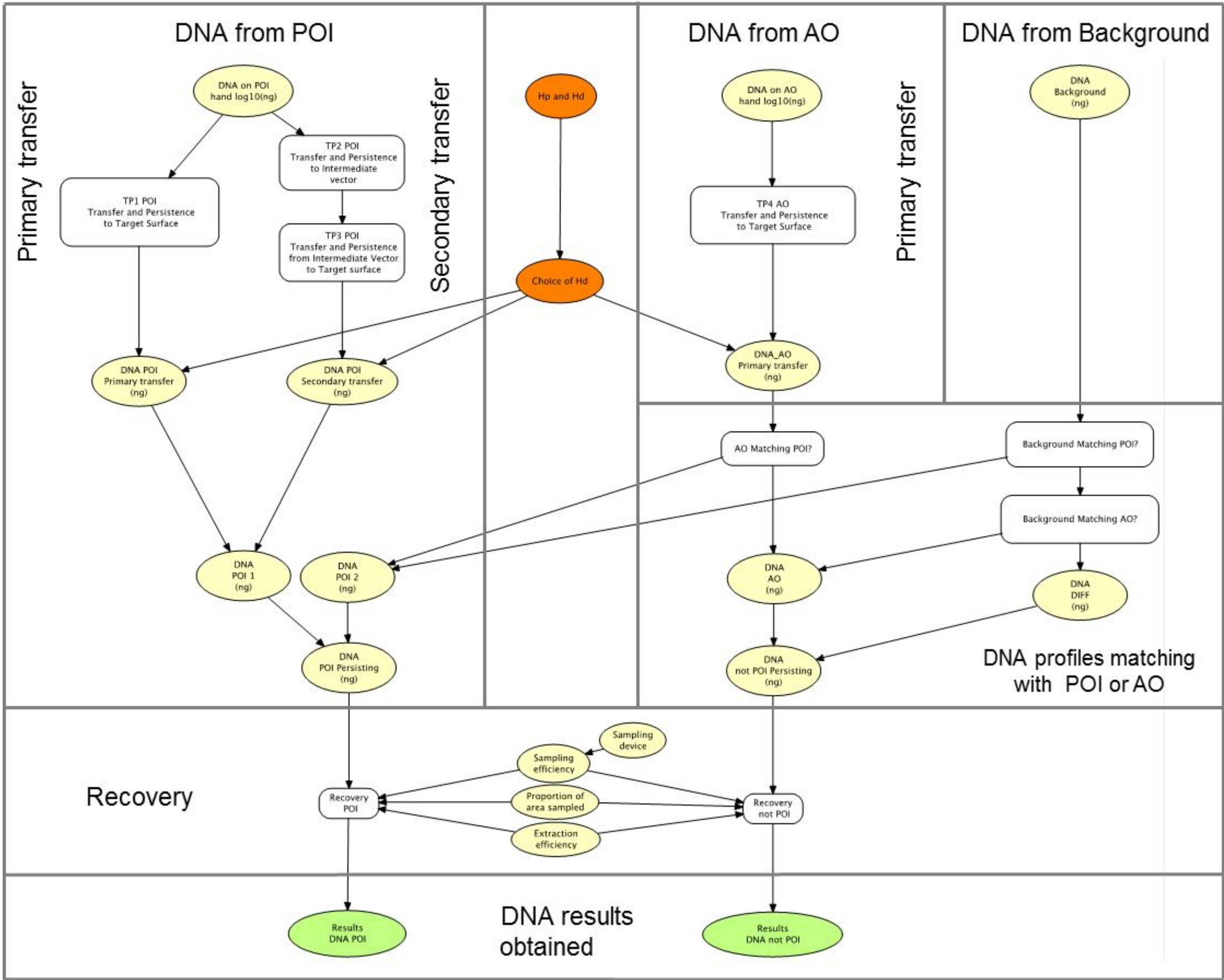
- Sampling device used [31].

Note that we do not consider specific laboratory aspects such as the profiling system used, the number of PCR cycles or models or settings of laboratory hardware. We have made a deliberate choice to model DNA amount, which precedes these laboratory considerations and simplifies the BN. If a laboratory wished they could add nodes onto the BN that translate DNA amount to peak height.

By combination of these factors we formulate the *BN* shown in Figures 8 to 11. This *BN* is constructed as an object-oriented BN (OOBN) with the same sub-networks that are used at multiple points. This model has the advantage that it can be easily expanded to consider a range of transfer scenarios (something we demonstrate in this paper). Overall in Figure 8 we model from top to down the amount of transferred DNA from contact to recovery. In each column, we distinguish the quantity of DNA from the person of interest (POI), the quantity of DNA from the alternative offender (AO), if any, and the quantity of DNA present as background. The proposition nodes at the centre allow activating or not the various transfer

540 options to be considered as a function of the choice of the prosecution or defence allegations.
541 Typically, under the prosecution view, it will be alleged that the POI invoked a primary
542 transfer with the item under examination. Under the defence account, the POI may invoke a
543 secondary transfer with the item alone (Hd1) or, that an alternative offender had a primary
544 transfer (Hd2). On each column (POI or AO), the DNA can take the routes denoted as
545 primary transfer or secondary transfer. The background DNA (on the right side) is not
546 conditioned on the chosen propositions as it pertains to the item regardless of their states.
547
548 Part of the BN deals with the issue as to whether or not the obtained DNA profiles will match
549 with either POI or AO. Before dealing with the results, we deal with the recovery of the DNA
550 from the item. This structure represents the obtained results (at the bottom) in the form of the
551 quantities of DNA arising from transfers of various types (primary or secondary) from POI,
552 or not.
553
554 The use of object-oriented structures is shown by the use of the white blocks called TP
555 (transfer and persistence), R (recovery) and M (matching DNA profiles).
556
557 The block TP is shown in Figure 9. It takes a given amount of DNA as input (DNA IN) and
558 progresses it through a transfer and persistency model to a resulting amount of DNA (DNA
559 OUT). Within the block is highlighted the various factors that impact the transfer and
560 persistency.
561
562 The block R (for recovery) shown in Figure 10 uses a DNA quantity as input and passes it
563 through the steps of sampling (depending on the technique used) and extraction, leading to
564 the final amount of resulting DNA.
565
566 The block M in Figure 11 is assigning the DNA as matching POI (or AO) versus different
567 profiles (called DNA DIFF) as a function of the match probabilities (themselves depending
568 on the quantity of DNA as input).
569
570 When considering DNA in term of a primary transfer, it will go through one TP block. When
571 considering DNA in terms of a secondary transfer, two TP blocks are applied. The first deals
572 with the transfer on the intermediate object and the second deals with the transfer from the

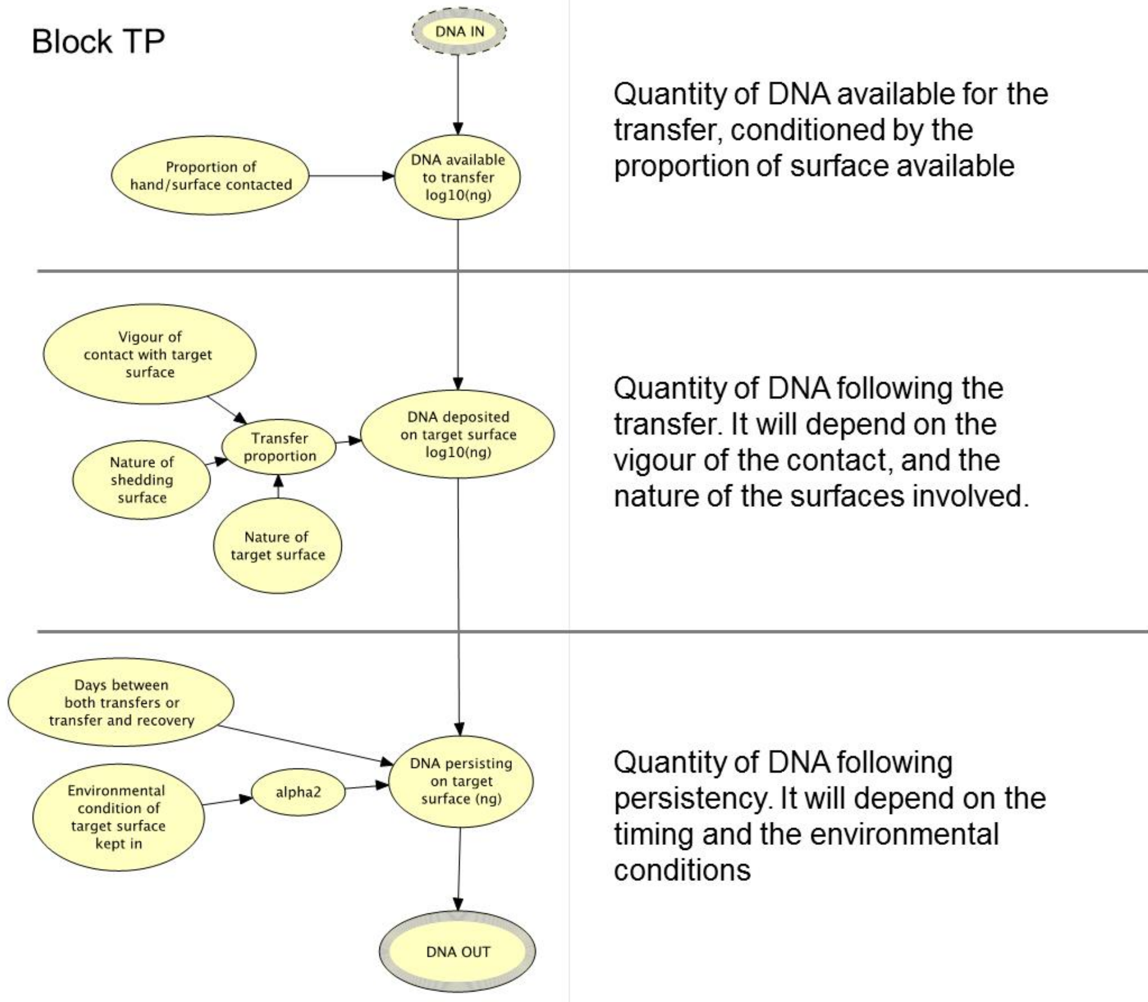
573 intermediate to the item under examination. This flexible construction also allows us to
574 model more complex scenarios (e.g. tertiary transfer), if necessary.
575
576



577

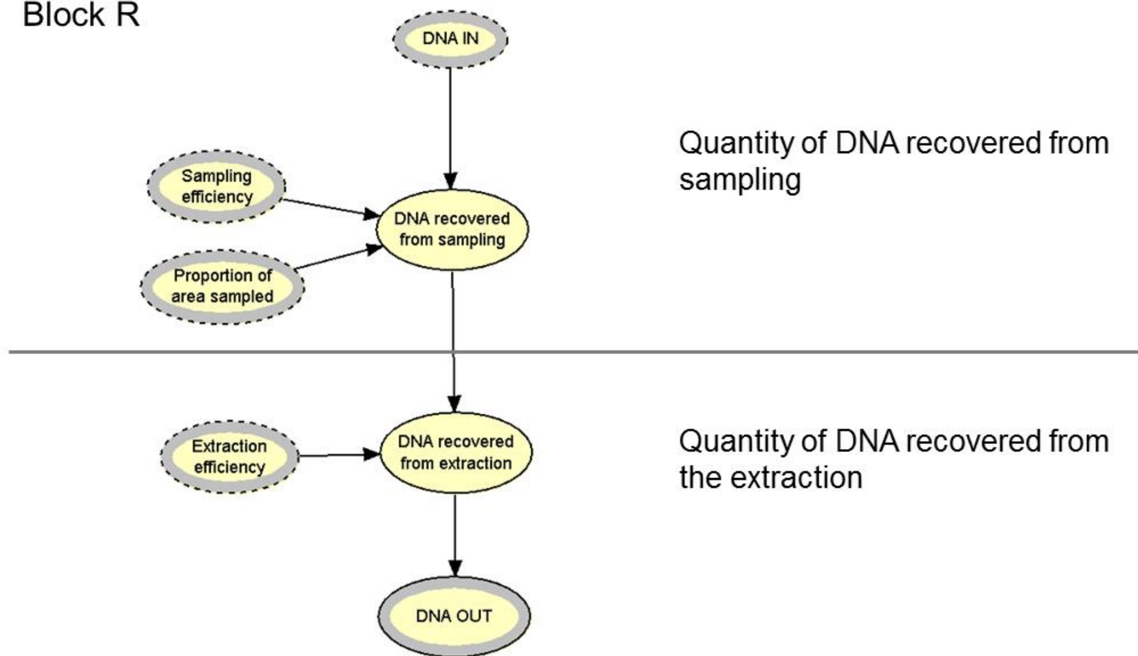
578 *Figure 8: Bayesian network used to evaluate the findings with given activity level propositions involving primary vs secondary transfer event*

579



580 *Figure 9: Bayesian network in block TP*

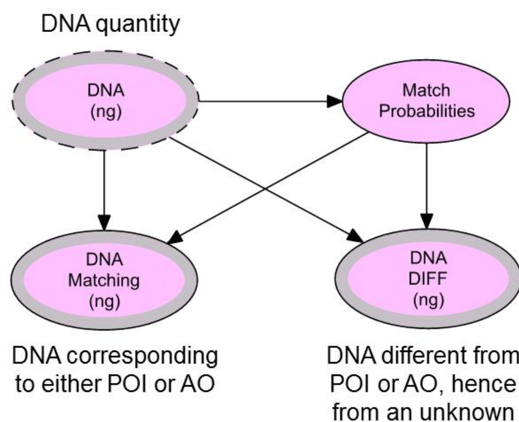
581



582

583 *Figure 10: Bayesian network in block R*

584



585

586 *Figure 11: Bayesian network in block M*

587

588 A combination of conditional probability tables and expressions have been used in the *BN*
589 shown in Figures 8 to 11. Tables 1-4 (in appendix 1) summarise the node definitions, the
590 node states and the manner in which probabilities are provided or generated. Some nodes
591 repeat for each transfer step. The general terminology is given in Tables 1 to Table 4.

592

593 **APPLICATION TO DIFFERENT CASE EXAMPLES:**

Having constructed and examined the BN that can help address secondary vs primary transfer we now apply these data to real casework examples. We provide three examples that have been encountered during testimony, giving a brief description of the alleged offence and the competing propositions. Details have been altered slightly from the real case so that we can demonstrate a range of situations.

Case example 1:

A bus driver (the POI) is charged with indecent assault where Prosecution alleges that he touched the breasts of the victim over the top of her T-shirt. The victim's T-shirt was seized and sampled the following day and a tapelift from the outer front of the victim's shirt revealed the sole presence (0.15 ng) of the POI's DNA[4]. The POI claims that he put the seatbelt on the victim and his DNA transferred from his hands, to the seatbelt and then secondarily to the victim's T-shirt. The propositions are therefore:

Hp: The POI touched the breasts of the victim on the outside of her T-shirt

Hd1: The POI put the seatbelt on the victim and did not touch her breasts

Given the alternative proposition in this scenario, there is no indication of an alternative offender (AO), hence only Hd1 will be considered. The node "choice of Hd" has been set accordingly.

Any state within any node of the BN can be set as being true (with all other states within that node therefore being false). Information provided to a BN in this manner is called 'instantiation' (i.e. the user is instantiating the states of nodes) and once done the laws of probability can be used to propagate the information throughout the BN and update the posterior probabilities for states in non-instantiated nodes. Our instantiations (and rationale) of the nodes is given in Table 5.

Node	Instantiated	Reason/Explanation

[4] The findings were initially expressed as a likelihood ratio considering the probability of obtaining the DNA profiling results if the POI was the source of DNA rather than if an unknown male was the source of DNA. The *LR* in this instance was strongly in support of the first proposition over the second and it was conceded by both parties that the POIs DNA was present on the shirt of the victim. In subsequent scenarios when we talk about an individual's DNA being found on an item a similar course of events has taken place to come to that statement. We are not simply assigning identity as the sub-source level *LR* reaches some threshold.

	state	
Proportion of hand /surface contacted	0.9 – 1 0.9 – 1 0.6 – 0.7	TP1: It is assumed that most of the hand of the POI would have contacted the shirt in the described assault. TP2: Under secondary transfer the same value will assuming that most of the hand of the POI would have contacted the seatbelt when he it on the victim. TP3: Typically only inner surface of the seatbelt would be contacting the shirt, which is 0.5, however there may be slight opportunity for the outer surface to contact the shirt so we choose 0.6 – 0.7
Nature of target surface	rough	The surface of the victim's woven shirt is rough
Vigour of contact with target surface	friction	TP1: The alleged assault describes a rubbing motion over the top of the victim's shirt. TP2 and TP3: There is typically gripping and pulling motion when putting on a seatbelt, which can be assimilated to friction.
Nature of shedding surface	hand rough	TP1 and TP2: The hand is the primary source of DNA TP3: The surface of the seatbelt is rough
Days between both transfers or transfer and recovery	1 0	TP1 and TP3: 1 day, the item was examined the following day. TP2: The seatbelt was immediately in contact with the victim's shirt
Environmental conditions	favourable	TP1 and TP3: Kept indoors in a paper bag TP2: Inside bus (note that as the time node associated with this condition node is set to 0, either condition would give the same result when instantiated)
Proportion of area sampled	0.9 – 1	The entire front of the shirt was sampled

Sampling device	tapelift	A tapelift was used to sample the shirt
Results DNA POI	0.1-0.2	The DNA amount obtained from laboratory analysis
Results DNA not POI	0-0.01	The DNA amount obtained from laboratory analysis

Table 5: Choice for node instantiations as seen in Figure 9

The result was 0.15ng of the DNA for the POI without any other DNA contribution (hence DNA not POI is set to 0). It gives a *LR* of 4. The *LR* for different observed quantities of the POIs DNA in scenario 1 shown in Figure 12.

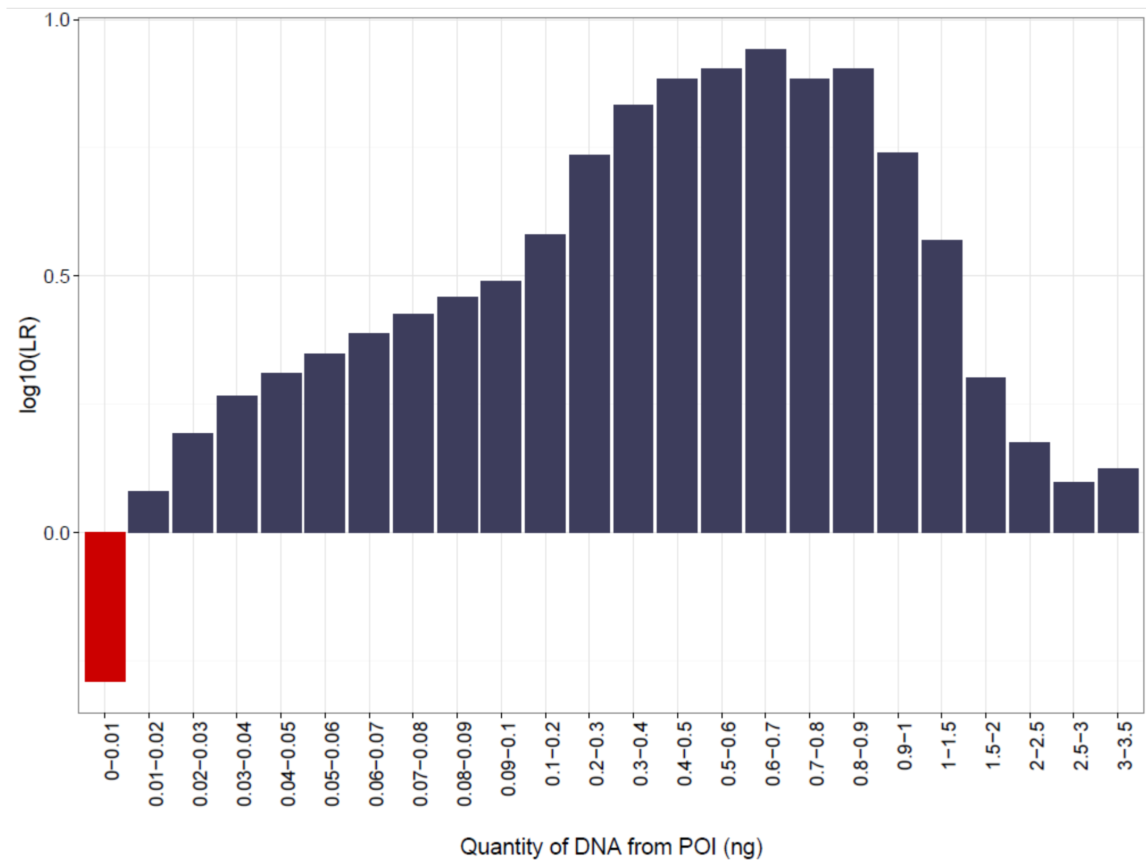
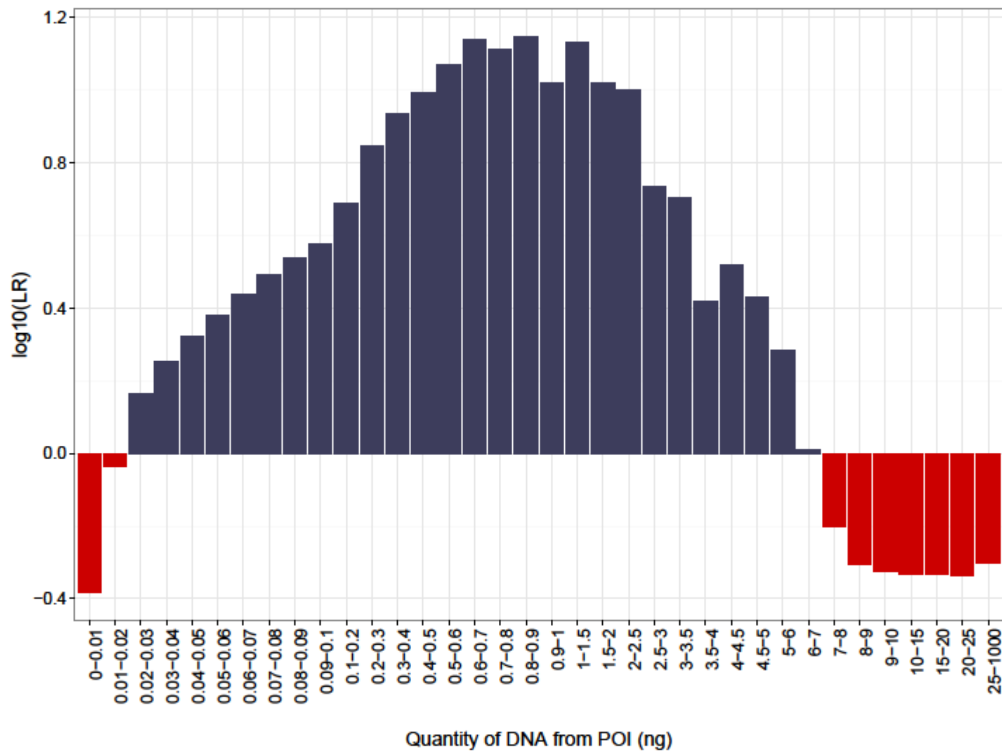


Figure 12: Ratio of probability for primary vs secondary transfer obtained for the instantiated network considering different levels of detected POI DNA.

What is interesting is that as the observed amount of the POI's DNA is above 0.01ng, the support for a primary transfer over a secondary transfer is now above 1 and increases, until it reaches approximately 10 when DNA amounts are 0.6-0.7 ng. As DNA amount increases

635 beyond 0.7 the level of support for a primary transfer decreases compared to a secondary
636 transfer. This initially appeared counterintuitive, but an examination of the '*DNA on hands*'
637 node reveals that such a finding means that the individual is a very good DNA shedder. This
638 then obscures the difference between a primary and secondary transfer event. If the amount
639 of DNA on an individual's hands is instantiated as low, then the effect is increasing support
640 for primary over secondary transfer as the amount of POI's DNA detected increases.
641
642 The other point to notice is that in this case example, the maximum value the *LR* reaches is
643 approximately 10. This demonstrates – in this case example – the support we can assign to
644 propositions of transfer type with our current knowledge. It should also be noted that
645 increasing our knowledge may not necessarily yield higher levels of support. If the variability
646 of transfer events is high (even after taking into account additional factors) then further
647 knowledge and experimentation will only serve to reinforce that fact.
648
649 Also, there are other factors that could be taken into account with further modelling and
650 information. For example, it may be that the suspect is the regular driver of the bus and it
651 could therefore be expected that a level of his DNA is present on the seatbelt prior to the time
652 of the alleged crime. This too could be modelled through a node that considers the amount of
653 DNA present on an item through habitual use. Or a more direct study could test seatbelts in
654 buses and compare them to the reference DNA of the regular driver, although this would be a
655 more difficult task for many laboratories to perform.
656
657 Case example 2:
658 A POI was accused of stabbing the victim with a plastic handled knife. A swab of the handle
659 of the knife produced a single sourced DNA profile that was conceded as originating from the
660 POI. Prosecution claim that that it was the POI who used the knife to stab the victim. The
661 POI claims that he was at a party shortly before the incident, where he shook hands with a
662 male friend. His DNA could have been on the knife handle because he transferred his DNA
663 to the hands of his friend, who then used the knife to stab the victim. In doing so the POI's
664 friend transferred the POI's DNA onto the knife handle. The propositions are therefore:
665 Hp The POI stabbed the victim with the knife
666 Hd2 The POI shook hands with a friend who stabbed the victim with the knife
667

668 In this case, the alternative proposition assumes an alternative offender (AO), hence Hd2 will
669 be considered. The node "choice of Hd" has been set accordingly.
670
671 The same structure of the BN as shown in Figure 8 can be used, but instantiated differently.
672 We do not provide the same explanation regarding node instantiations as we did for scenario
673 1 in Table 5. The choices themselves are not as important in these demonstrations as the
674 construction and function of the *BN* itself. One can imagine what assumptions may be made
675 regarding surface types and types of contact in regular casework, and if this information is
676 unknown then it is always possible to either leave the node(s) uninstantiated or to trial the
677 effect of instantiating with different values. In this way the scientist can determine how
678 important that piece of information is to the robustness of the *LR* provided. In this scenario
679 we have a hand to hand transfer and have not obtained data for such situations. We consider
680 that a surface type 'hand' is the same that a surface type 'rough'. However, DNA transfer and
681 persistence on hands may not act the same as the data for 'rough' or 'smooth' surface types
682 and ideally further considerations for the surface type 'hands' should be considered within
683 the *BN*. We note that the DNA persistence on hands may be different due to bacterial
684 degradation or normal wear and tear on hands from everyday activities.
685
686 What is interesting is that as the observed amount of the POI's DNA increases above 0.02ng,
687 the support for a primary transfer over a secondary transfer is now above 1 and increases,
688 until it reaches approximately 14 when DNA amounts are 0.8-0.9 ng. As DNA amount
689 increases beyond 0.9 the level of support for a primary transfer decreases compared to a
690 secondary transfer. Above 7ng, the findings start to support a secondary transfer over a
691 primary transfer can be observed. (see Figure 13). This is a combination of the BN moving to
692 a position where the POI is considered a high shedder and the AO a low shedder and the level
693 of background DNA being higher with a coincidental matching alleles. Instantiating these
694 factors so that they cannot be the case i.e. restricting the shedder status of the POI to an
695 average value, removing the consideration of background DNA and specifying that the AO
696 DNA profile is not matching with the POI sees the *LR* with propositions as stated above
697 steadily increase with increasing POI DNA amounts.
698



699

700 *Figure 13: Ratio of probability for primary vs secondary transfer obtained for the*
701 *instantiated network, but considering different levels of detected POI DNA, but always*
702 *keeping the detected levels of the unknown individual's DNA as 0 – 0.01ng.*

703

704 We could also create a similar *BN* for situations like case example 2, but where there are two
705 POIs, one of which stabbed the victim, both of their DNA is detected on the item, and both
706 having submitted reference DNA samples. This could be achieved by the addition of a
707 secondary transfer route on for the AO (who would be considered the second POI) into the
708 BN in Figure 8 with the 'Choice of Hd' node specifying either a primary transfer for POI1
709 and a secondary transfer for POI2 or vice versa. We do not provide a Figure showing such a
710 BN.

711

712 Scenario 3:

713 An unregistered firearm was found on a couch in the house owned by the POI. Police seized
714 the firearm and a swab of the firearm stock revealed a mixed DNA profile originating from
715 two individuals of which the POI was conceded as being the minor source. The prosecution
716 claim that the POI handled the firearm. The defence claim that the POI had not known about
717 the firearm and someone must have put it on his bed and then moved it to the couch very
718 recently. The POI also claims that he saw something sitting on his bed earlier that day and

719 thinks it may have been the firearm, but he didn't investigate or touch it. The POI therefore
720 states that the presence of his DNA on the firearm is due to it transferring from the POI to the
721 couch, or bed, and then to the firearm. The propositions are therefore:

722 Hp The POI recently handled the firearm

723 Hd2 The POI has never handled the firearm, but someone must have moved it

724

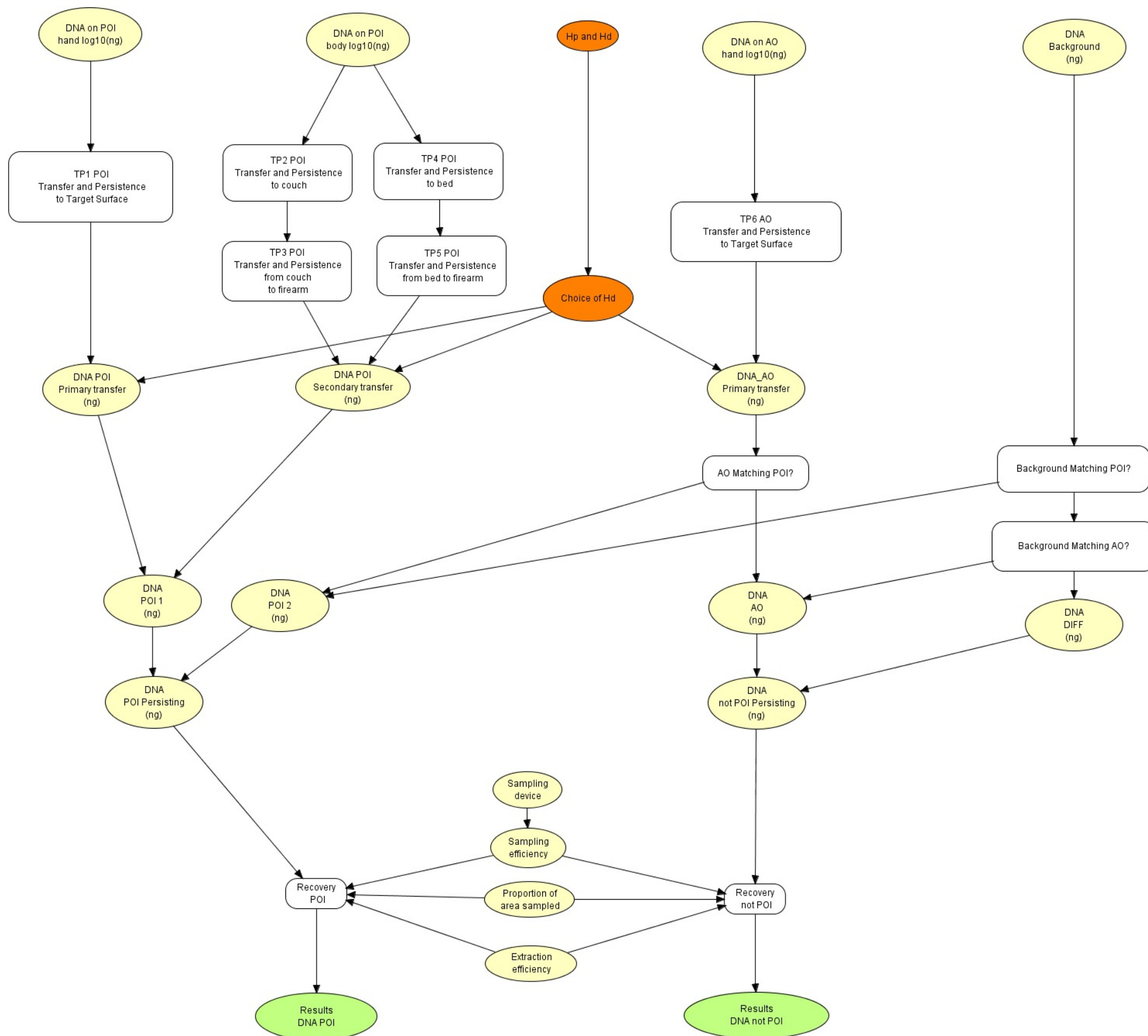
725 Here we have the deficiency of knowledge regarding the amount of DNA on a regularly used
726 couch and bed. There is also the possibility of an accumulative effect, i.e. if someone rubs
727 their hands on the same item multiple times on different occasions, does the DNA keep
728 accumulating or does it reach a saturation point? At present this information is not known and
729 so we make a number of assumptions that we explain below. These are further areas of
730 research that would be beneficial for the forensic community to carry out.

731

732 Again we do not specify the reasoning behind each instantiation in the *BN*. The intention here
733 is to demonstrate the adaptability and power of the *BN* to handle a variety of situations.

734

735 The extended *BN* for scenario 3 can be seen in Figure 14. We can see now that this is similar
736 to the initial *BN* as seen in Figure 8, however we have the two routes for secondary transfer
737 from the POI, one in relation to the bed, the other in relation to the couch. They meet in a
738 node that adds the DNA amounts from secondary transfer together. In doing this we assume
739 that if DNA has been transferred from both secondary pathways then the effect is pure
740 accumulation and none of loss of DNA. This is another area which requires some research to
741 be conducted.



742

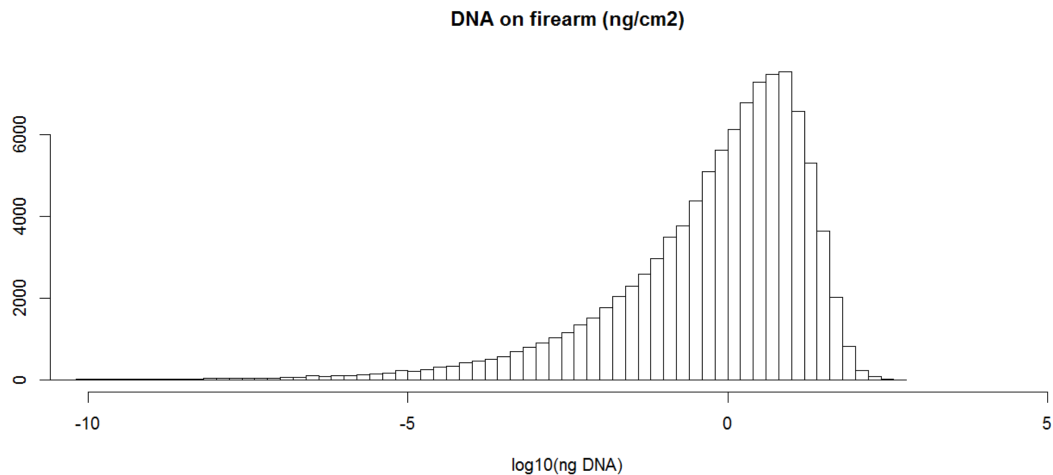
743 *Figure 14: Bayesian network used to evaluate the findings with given activity level propositions involving primary vs secondary transfer event*
744 *for scenario 3*

745

746

747 For the background node we require the amounts of DNA found on firearm from individuals
748 other than the primary user(s). The reason for this is that the primary user(s) will have
749 transferred their DNA via a primary contact and we are interested in background levels
750 (deposited through some unknown mechanism, and not by a known contact). This
751 information is not readily provided by any literature sources we could find. The study of
752 McKenna [51] found that a DNA profile was not observed in 26% of firearms swabs. A study
753 by FSSA [52] examined DNA amounts from 300 firearm swabs obtained in casework. While
754 ideally for our purpose this data would be compared to the owner or regular user of the
755 firearm so they could be screened out of the DNA obtained, we use the distribution of DNA
756 amounts observed from [52]. We then extrapolate back to DNA amounts found on the
757 firearms in the same manner as previously described to obtain the distribution of DNA
758 originally on the firearm, which we show in Figure 15. We apply the distribution values seen
759 in Figure 15 directly into the Background DNA node.

760



761

762 *Figure 15: DNA amount (grey) obtained from swabs of firearms and fitted distribution*
763 *(black)*

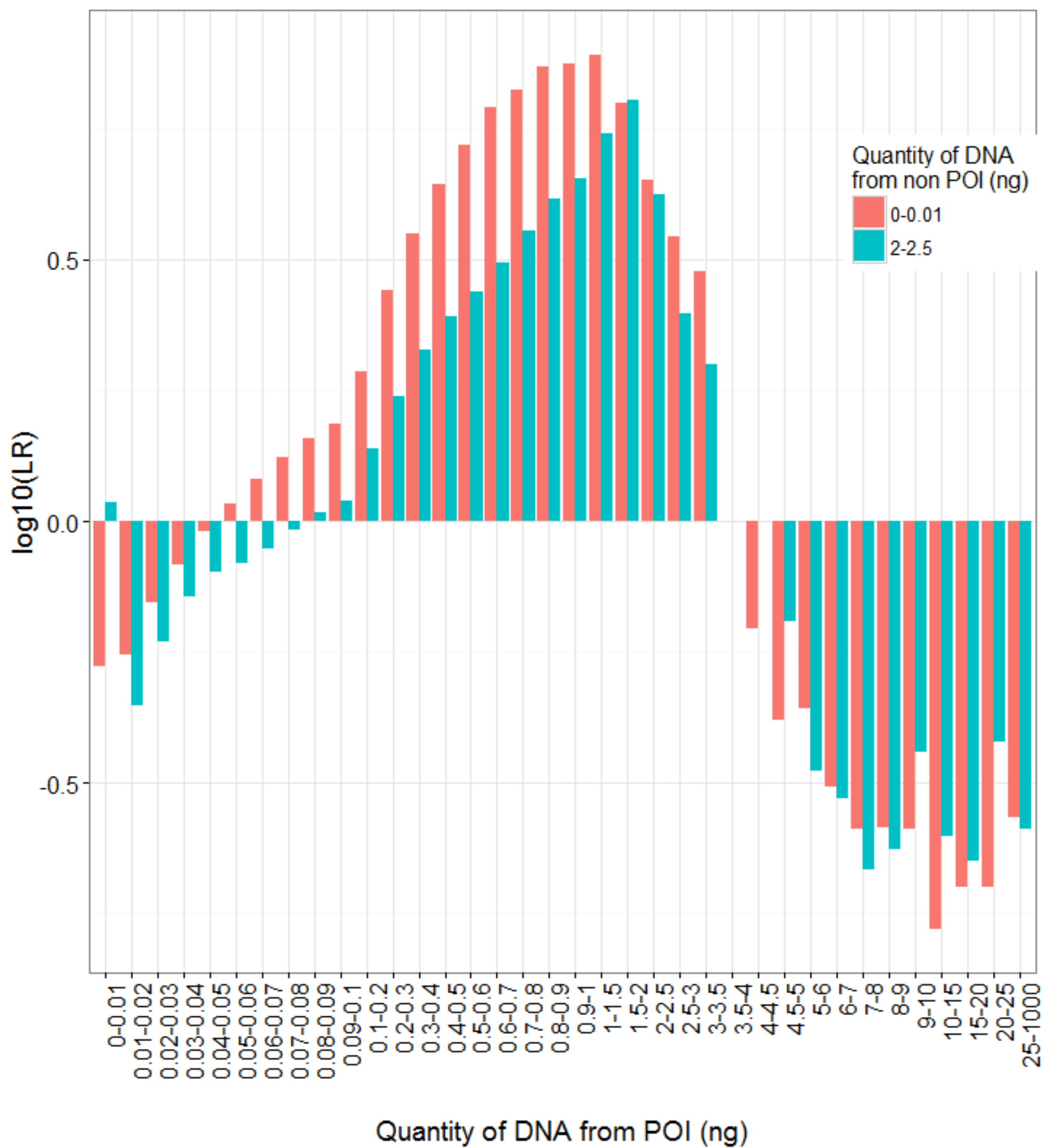
764

765 In this case example we assume that under the prosecution proposition that the contact by the
766 POI to the firearm is a brief one-off contact (perhaps as an allegation of a recent purchase). If
767 we were to assume an extended habitual use then different data would be required.

768

769

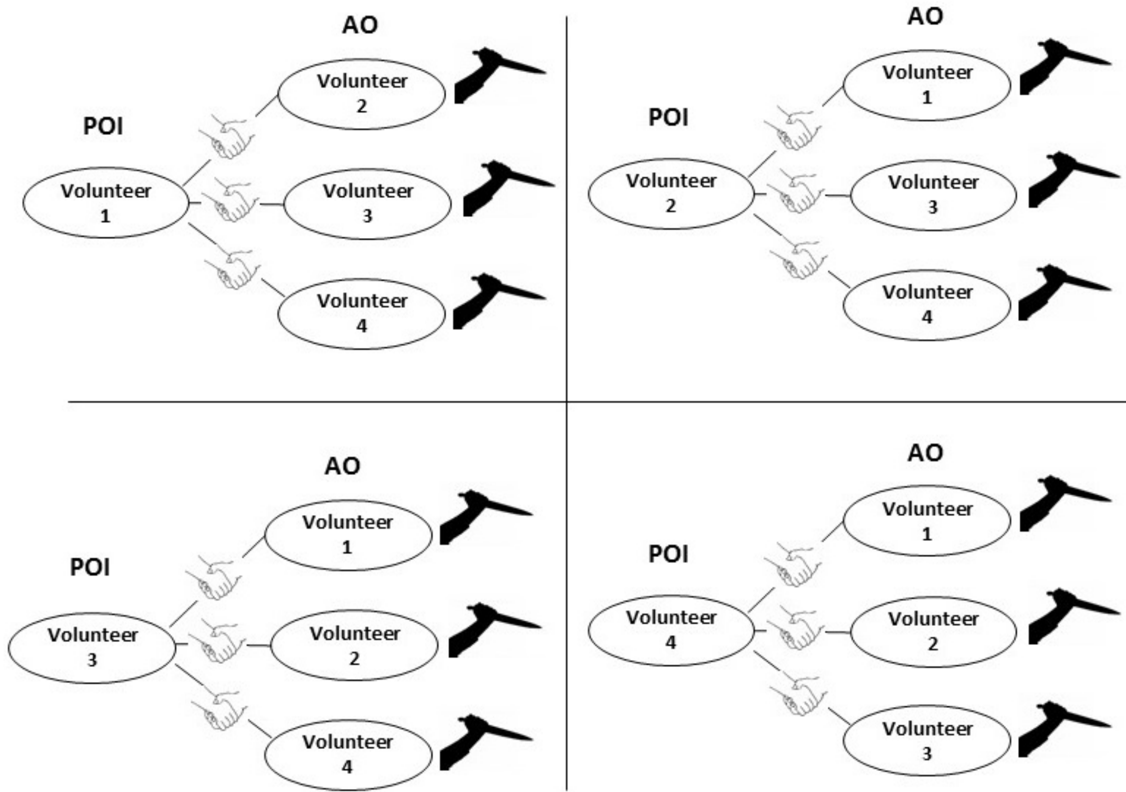
770 Using the *BN* seen in Figure 14 yields *LR* = 8 when the amount of the POI's DNA detected is
771 above approximately 1-1.5ng. The POI's DNA has come to be on the item either from a
772 direct contact with his hands or from a secondary transfer from the POI's bed or couch
773 transferred onto the firearm when it was placed there. The amount of DNA expected from
774 such a transfer is less that what is possible from primary transfer. Figure 17 shows the LR
775 over a range of DNA amount for POI and either relatively high or low DNA amounts from
776 the non-POI.
777



778
779 *Figure 17: LR over range of detected DNA for POI and relatively high and low DNA*
780 *amounts for non-POI.*

781
782 Again, as the DNA amount for the POI increases from zero to 0.5ng the support for primary
783 over secondary transfer increases, then as DNA increase further the support for primary
784 transfer decreases until approximately 4ng, when the findings start to support secondary over
785 primary transfer. At higher DNA levels the support for secondary over primary transfer is a
786 product of the fact that the modelling of background DNA, coupled with the accumulation
787 from multiple sources (couch and bed) means that higher amounts of DNA are more
788 indicative of secondary transfer in this scenario.
789
790 This example shows the importance of a good set of data for modelling background DNA and
791 levels of DNA expected on items from habitual use. The number of samples and methods in
792 which they have been collected for our example, suggest that further work would need to be
793 required in order to address the findings in consideration of the propositions in scenario 3.
794 The closeness of the primary transfer probability to the background DNA probability in
795 Figure 15 suggests that much of the data captured in the FSSA study are likely to be resulting
796 from primary contacts.
797
798 As a point of interest, if the scenario were changed to one which stated that the unknown
799 male rubbed the gun on the bed in a deliberate attempt to transfer DNA, then the support for
800 secondary transfer over a primary transfer persists for quite high levels of observed POI
801 DNA. Note that in the evaluation of all the evidence the court is likely to have quite different
802 prior beliefs on whether someone else brought a firearm into the house, compared to that
803 person then wishing to deliberately 'frame' the suspect. This shift in prior beliefs may well
804 outweigh the differences obtained from the activity level considerations of the DNA findings,
805 but of course it is not up to the scientist to base their decisions on such considerations.
806
807 **APPLICATION TO CONTROLLED CASES**
808
809 We have explained the construction of an OOBN that can consider competing transfer
810 mechanisms and demonstrated its use in several scenarios. We wish now to test the ability of
811 the BN to distinguish known primary and secondary transfer events as described in scenario 2
812 of the stabbing case. We have used the work of Samie et al. [43], which importantly, was not
813 used in any of the modeling to assign conditional probabilities in the BN. This allows the
814 results of Samie et al. [43] - where the authors study primary DNA transfer to knife handles

815 during the action of stabbing- to act as a test set. For the purpose of this research, a series of
816 secondary transfer experiments was this time performed. 12 experiments were carried out. In
817 order to mimic stabbing under the second scenario (secondary transfer), the same four
818 individuals (two males and two females) and same type of knives were used. The persons'
819 hands were washed at 8 am. Around 11 am or 3 pm, one volunteer (POI, acting as the
820 innocent suspect) was asked to shake hands with the volunteer who would act as the stabber
821 (alternative offender, AO). The person then carried out normal activities in their office
822 environment (i.e., having lunch or coffee with their colleagues, speaking with them etc.).
823 Thirty minutes later, the AO was asked to 'stab' a cardboard box with a knife. Right
824 afterwards, traces were collected using the double swab method. The following day, the
825 experiment was repeated with another volunteer acting as the innocent POI (Figure 18). Each
826 volunteer took the role of POI associated with each of the three stabbers. In order to limit
827 background DNA (i.e., DNA present for unknown reasons), the knife was cleaned between
828 each experiment (by using bleach, ethanol and leaving the items under UV light for 30 min).
829 To monitor background, a negative control was taken from the knife after cleaning and before
830 the experiment. Results were all negative (no DNA profile). To collect, extract and amplify
831 the DNA, we have used the same method as described in Samie et al. [41]. However, here,
832 DNA was quantified using the Investigator® Quantiplex (Qiagen) kit following standard
833 protocols and the amplicons were analysed using a 3500 Genetic Analyser ABI (Applied
834 Biosystem) and GeneMapper®IDX Software.
835 Then, the DNA profiles were interpreted using STRmix™ v2.3.05 [1, 53, 54, 55]. This
836 forensic software has been developed to resolve mixed DNA profiles based on a continuous
837 approach. The programme uses peak height information and statistically accounts for the
838 possibility of degraded DNA and stochastic variation such as stutter, allelic drop-in and drop-
839 out. It provides information regarding the mixture proportion and the weight that is given to
840 the possible genotypes of the contributors. The weight is used to express how well a proposed
841 genotype explains the profile. These information, combined with the total of DNA quantity
842 obtained, allowed to inform the node "Results DNA POI" and "Results DNA non POI". The
843 number of contributors was determined based on the number of the peaks detected at each
844 locus, peak height balance information and how the experiments were planned (i.e., we
845 expected the DNA of two persons in different proportions).
846
847



848

849 *Figure 18: Figure showing the experimental design adopted for experiments on primary and*
850 *secondary transfer.*

851
852
853
854

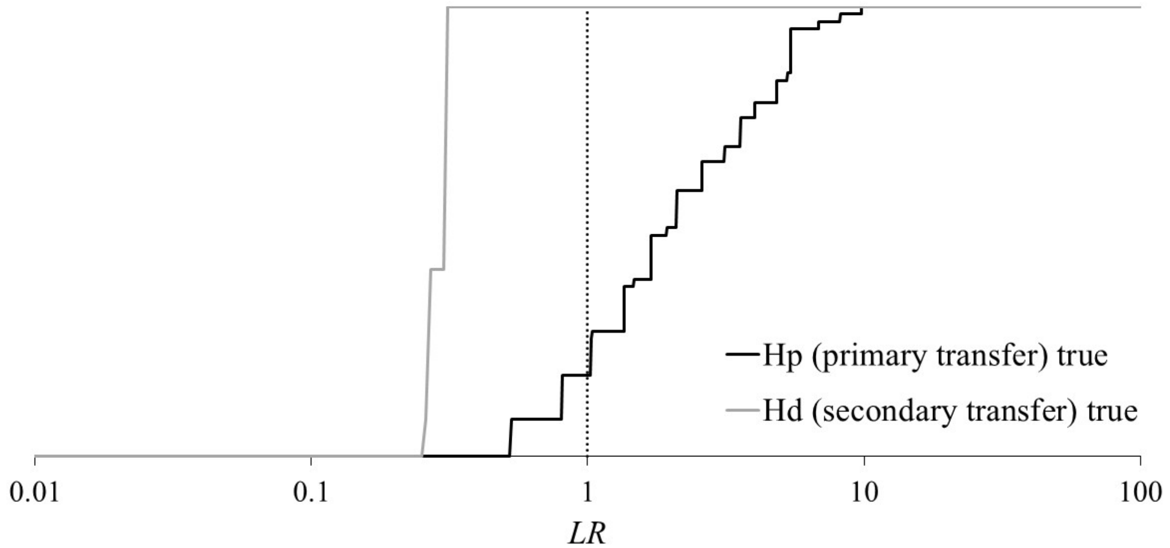


Figure 19: Tippett curve showing the ability of the developed BN to distinguish primary and secondary transfer events.

Figure 19 shows the results of considering each of the transfer events as primary. Given the limited dataset available, there appears that there is some ability for the system to distinguish primary from secondary transfer events, even without the knowledge of the shedder status of those involved.

CONCLUSION:

We show here the construction of *BN*s built up from building blocks that address different features of contact DNA transfer, persistence and recovery. By splitting the transfer, persistence and recovery into separate nodes we make the BN high configurable to a number of situations, which we demonstrate through three quite different scenarios. The scenarios also highlight the importance of having the applicable data to inform conditional probabilities that underlie each node.

In theory the complexity of the scenarios that *BN* can consider are endless but in reality there will be a diminishing return as more complexity is added. There are likely to be key information that has the greatest effect on the posterior probabilities of the propositional node and these can be explored with sensitivity analyses (something which we have not

877 demonstrated in this paper, but intend to pursue to determine which factors the LR is most
878 sensitive to, and also to help direct further research areas [56]).
879
880 There are a number of ways in which *BN* such as those we have developed here could be used
881 other than to evaluate findings given competing transfer mechanisms. If the mechanism is
882 known, then the *BN* could be used to investigate whether an individual is a high or low
883 shedder. Alternatively, the *BN* could be used in a case assessment and interpretation
884 framework (see [57] for an explanation of case assessment and interpretation), where items
885 are triaged depending on the level of power that will exist to support one proposition over the
886 other when transfer mechanism is disputed.
887
888 We show, in the 'application to controlled cases' section that the *BN* appears to have some
889 ability to distinguish primary and secondary transfer under the conditions tested. This is
890 despite not having information regarding the shedder status of the individual, which has been
891 classically thought to be a limiting factor in the ability to evaluate evidence given competing
892 transfer scenarios. Additional ground truth tests would be beneficial and would likely
893 highlight situations where additional complexity, or additional modelling is required to
894 inform the *BN*. In appendix 2 we suggest a number of studies we have already identified that
895 would provide beneficial data to inform us on the important factors in transfer and persistence
896 of trace DNA. As further information regarding the factors affecting DNA findings is
897 obtained it is the authors' hope that they can be incorporated into evaluations in a logical and
898 helpful manner.
899
900 The work here also brings up an interesting point regarding the modelling of 'background
901 DNA'. This was particularly important in the assignment of probabilities in scenario 3. All
902 DNA must have come to be on an item through some mechanism. Background DNA is
903 defined as DNA that is expected to be present on an item, but is not related to the activity in
904 question. Typically we consider that background DNA has not come from one of the POIs,
905 but (as with any DNA source) may adventitiously possess the same alleles. When background
906 DNA is modelled it is common practise to consider background DNA on an item as that
907 which has not come from the primary handler. However, in a BN that it used to model multi-
908 step transfer mechanisms the more correct modelling of background DNA would be DNA
909 that has come from one more transfer step than the most complex mechanism being modelled
910 i.e. if one scenario being modelled is a tertiary transfer then background DNA would be DNA

911 that has come from a quaternary transfer event. All other sources of DNA that have come
912 from primary, secondary or tertiary transfers should be modelled in full. This is not an easy
913 task to achieve, firstly due to the complexity of the BN that would arise, but secondly in the
914 modelling of background DNA amounts at differing levels of transfer complexity.
915
916 There is always a difficulty when modelling data from the literature to find experimental
917 designs that match the case scenario exactly. Indeed, some case scenarios simply cannot be
918 exactly replicated in experimental work due to ethical or legal reasons. A common question
919 arising from this is therefore, whether *BN* such as those constructed here can be applied to
920 casework at all (see [58] for a discussion of this point and others in the same vein). Such a
921 line of argument fails to recognise how information and casework circumstances are
922 evaluated by the court. If the primary dispute when considering DNA evidence is one of a
923 transfer mechanism, then the presence or absence of DNA is not in dispute. We then must
924 ask, who is best placed to answer questions of transfer mechanism, which will inevitably
925 require knowledge of transfer, persistence and recovery of DNA and the levels of background
926 DNA in the environment. We suggest that it would be unrealistic to expect the average juror,
927 judge or lawyer to possess such knowledge and that the highly specialised considerations are
928 best explored by the scientist. Having made this decision, the question needs to be asked, how
929 will the scientist take into account such a wide range of considerations and where will they
930 draw their knowledge from. We would argue that the most logical and transparent manner in
931 which this can be done is by setting out all the factors requiring consideration within a *BN*
932 and populating probabilities using the most applicable data available. This will inform the
933 scientist, which in turn can educate the court in the most robust manner possible. Having said
934 all of this, we do not advocate the use of data which are clearly not suitable for assigning
935 probabilities; but more often the case when some specific piece of information is not known,
936 a reasonably close substitute can be used. As long as the conditions to which the data pertain
937 (i.e. experimental settings) are judged appropriately close to be considered as an acceptable
938 substitute for the case at hand then the resulting *BN* still represents the best evaluation of
939 findings available. When this practice is adopted the use of the substitute data should be
940 clearly pointed out in the report. One example of how this can be applied in the BN shown in
941 our work is for the proportion of the hand that has contacted an item. Clearly, this
942 information will not be available to the analyst. However, depending on the item an educated
943 value can be used, e.g. if a swab of the trigger of a firearm was taken then only the upper part
944 of one finger is likely to have contacted the item and this can be portrayed in the values

945 instantiated in the "Proportion of Hand contacted" node. If there is a high degree of
946 uncertainty regarding the proportion of area contacted, then a distribution of prior
947 probabilities across the states in the node can be applied to reflect this.
948
949 No-one is served by the scientist simply refusing to provide an opinion due to an absence of
950 some small area of data that perfectly aligns with case circumstances. When the absence of
951 knowledge is great (such as is scenario 3) then this will become apparent to the scientist and
952 under these circumstances it is appropriate for them to state that the results cannot be
953 evaluated robustly given the two competing propositions (see [56] for a discussion on this).
954
955 **ACKNOWLEDGEMENTS:**

956 Points of view in this document are those of the authors and do not necessarily represent the
957 official position or policies of their organisations.
958 Alex Biedermann gratefully acknowledges the support of the Swiss National
Science Foundation through grant No. BSSGI0_155809.

959 **REFERENCES:**

960 [1] Taylor D, Bright J-A, Buckleton J. The interpretation of single source and mixed DNA profiles.
961 Forensic Science International: Genetics. 2013;7:516-28.
962 [2] Haned H. Forensim: An open-source initiative for the evaluation of statistical methods in forensic
963 genetics. Forensic Science International: Genetics. 2011;5:265-8.
964 [3] Lohmueller K, Rudin N. Calculating the weight of evidence in low-template forensic DNA
965 casework. Journal of Forensic Sciences. 2013;58:s234-59.
966 [4] Steele CD, Balding DJ. Statistical Evaluation of Forensic DNA Profile Evidence. Annual Review of
967 Statistics and Its Application. 2014;1:361-84.
968 [5] Puch-Solis R, Rodgers L, Mazumder A, Pope S, Evett I, Curran J, et al. Evaluating forensic DNA
969 profiles using peak heights, allowing for multiple donors, allelic dropout and stutters. Forensic
970 Science International: Genetics. 2013;7:555-63.
971 [6] Balding DJ, Buckleton J. Interpreting low template DNA profiles. Forensic Science International:
972 Genetics. 2009;4:1-10.
973 [7] Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, et al. Validating TrueAllele®
974 DNA mixture interpretation. Journal of Forensic Sciences. 2011;56:1430-47.
975 [8] Cook R, Evett IW, Jackson G, Jones PJ, Lambert JA. A hierarchy of propositions: Deciding which
976 level to address in casework. Science and Justice. 1998;38:231-40.
977 [9] Fonneløp AE, Egeland T, Gill P. Secondary and subsequent DNA transfer during criminal
978 investigations. Forensic Science International: Genetics. 2015;17:155-62.
979 [10] Goray M, Oorschot RAHv. The complexities of DNA transfer during a social setting. Legal
980 Medicine. 2015;17:82-91.
981 [11] Meakin G, Jamieson A. DNA transfer: Review and implications for casework. Forensic Science
982 International: Genetics. 2013;7:434-43.
983 [12] Van Oorschot RA, Jones M. Retrieval of DNA from touched objects. The 14th International
984 Australia and New Zealand Forensic Science Society Symposium for Forensic Sciences. Adelaide,
985 Australia1998.
986 [13] van Oorschot RAH, Ballantyne KN, Mitchell RJ. Forensic trace DNA: A review. Investigative
987 Genetics 2010;1.
988 [14] van Oorschot RAH, Goray M, Eken E, Mitchell RJ. Impact of relevant variables on the transfer of
989 biological substances. Forensic Science International: Genetics Supplement Series.In Press, Corrected
990 Proof.
991 [15] Farmen RK, Jaghø R, Cortez P, Frøyland ES. Assessment of individual shedder status and
992 implication for secondary transfer. Forensic Science International: Genetics Supplementary Series.
993 2008;1:415-7.
994 [16] Casey D, Clayson N, Jones S, Leisj J, Boyce M, Fraser I, et al. A response to Meakin and Jamieson
995 DNA transfer: Review and implication for casework. Forensic Science International: Genetics.
996 2016;21:117-8.
997 [17] Meakin G, Jamieson A. A response to a response to Meakin and Jamieson DNA transfer: Review
998 and implications for casework. Forensic Science International: Genetics. 2016;22:e5-e6.
999 [18] Taylor D, Abarno D, Champod C, Hicks T. Evaluating forensic biology results given source level
1000 propositions. Forensic Science International: Genetics. 2016;21:54-67.
1001 [19] Wolff TRD, Kal AJ, Berger CEH, Kokshoorn B. A probabilistic approach to body fluid typing
1002 interpretation: an exploratory study on forensic saliva testing. Law, Probability and Risk.
1003 2015;14:323-39.
1004 [20] Evett IW, Gill PD, Jackson G, Whitaker J, Champod C. Interpreting small quantities of DNA: the
1005 hierarchy of propositions and the use of Bayesian networks. Journal of Forensic Sciences.
1006 2002;47:520 - 30.
1007 [21] Gittelson S, Biedermann A, Bozza S, F. T. Bayesian Networks and the Value of the Evidence for
1008 the Forensic Two-Trace Transfer Problem. Journal of Forensic Sciences 2012;57:1199-216.

1009 [22] Biedermann A, Bozza S, Taroni F. Probabilistic evidential assessment of gunshot residue particle
1010 evidence (Part I): Likelihood ratio calculation and case pre-assessment using Bayesian networks.
1011 Forensic Science International. 2009;191:24-35.
1012 [23] Wieten R, Zoete JD, Blankers B, KokShoorn B. The interpretation of traces found on adhesive
1013 tapes. Law, Probability and Risk. 2015;14:305-22.
1014 [24] Hicks T, Biedermann A, Koeijer JAd, Taroni F, Champod C, Evett IW. The importance of
1015 distinguishing information from evidence/observations when formulating propositions. Science &
1016 Justice. 2015;55:520-5.
1017 [25] Curran JM, Hicks TN, Buckleton J. Forensic Interpretation of Glass Evidence. London: CRC Press;
1018 2000.
1019 [26] R v Peter Weller [2010] EWCA Crim 1085.
1020 [27] Lehmann VJ, Mitchell RJ, Ballantyne KN, Oorchot RAHv. Following the transfer of DNA: How
1021 does the presence of background DNA affect the transfer and detection of a target source of DNA?
1022 Forensic Science International: Genetics. 2015;19:68-75.
1023 [28] Butts E. Exploring DNA Extraction Efficiency Forensics@NIST 2012 Meeting, Gaithersburg.
1024 2012;http://www.nist.gov/oles/3_Butts-DNA-extraction-2.pdf.
1025 [29] van Oorschot RAH, Phelan DG, Furlong S, Scarfo GM, Holding NL, Cummins MJ. Are you
1026 collecting all the available DNA from touched objects? International Congress Series. 2003;1239:803-
1027 7.
1028 [30] de Jong B, Antonie M, Hardon A, van den Bergen D, Kneppers A. Successful validation of a fully
1029 automated sample lysis workstation. Forensic Science International: Genetics Supplement Series.
1030 2013;4:e93-e4.
1031 [31] Verdon TJ, Mitchell RJ, Oorschot RAHv. Evaluation of tapelifting as a collection method for
1032 contact DNA. Forensic Science International: Genetics. 2014;8:179-86.
1033 [32] Raymond JJ, Oorchot RAHv, Gunn PR, Walsh SJ, Roux C. Trace evidence characteristics of DNA: A
1034 preliminary investigation of the persistence of DNA at crime scenes. Forensic Science International:
1035 Genetics. 2009;4:26-33.
1036 [33] Quinones I, Daniel B. Cell free DNA as a component of forensic evidence recovered from
1037 touched surfaces. Forensic Science International: Genetics. 2012;6:26-30.
1038 [34] Buckingham AK, Harvey ML, van Oorschot RAH. The origin of unknown source DNA from
1039 touched objects. Forensic Science International: Genetics. 2016;25:26-33.
1040 [35] Lowe A, Murray C, Whitaker J, Tully G, Gill P. The propensity of individuals to deposit DNA and
1041 secondary transfer of low level DNA from individuals to inert surfaces. Forensic Science
1042 International. 2002;129:25-34.
1043 [36] Phipps M, Petricevic S. The tendency of individuals to transfer DNA to handled items. forensic
1044 Science International. 2007;168:162-8.
1045 [37] Lacerenza D, Aneli S, Omedei M, Gino S, Pasino S, Berchialla P, et al. A molecular exploration of
1046 human DNA/RNA co-extracted from the palmar surface of the hands and fingers. Forensic Science
1047 International: Genetics. 2016.
1048 [38] Berge Mvd, Ozcanhan G, Zijlstra S, Lindenbergh A, Sijen T. Prevalence of human cell material:
1049 DNA and RNA profiling of public and private objects after activity sceanrios. Forensic Science
1050 International: Genetics. 2016;21:81-9.
1051 [39] Bontadelli L. Study of DNA shedder quality. School of Criminal Justice of Lausanne, University of
1052 Lausanne. Lausanne; 2009.
1053 [40] Plummer M. Bayesian graphical models using MCMC. RJAGS. 2012.
1054 [41] Goray M, Fowler S, Szkuta B, Oorschot RAHV. An analysis of self and non-self DNA in multiple
1055 handprints deposited by the same individuals over time. Forensic Science International: Genetics.
1056 2016;23:190-6.
1057 [42] Meakin G, Butcher EV, Van Oorschot RAH, Morgan RM. The deposition and persistence of
1058 indirectly-transferred DNA on regularly-used knives. Forensic Science International: Genetics
1059 Supplement Series 5. 2015:e498–e500.

1060 [43] Samie L, Hicks T, Castella V, Taroni F. Stabbing simulations and DNA transfer. Forensic Science
1061 International: Genetics. 2016;22:73–80.
1062 [44] Daly DJ, Murphy C, McDermott SD. The transfer of touch DNA from hands to glass, fabric and
1063 wood. Forensic Science International: Genetics. 2013;6:41-6.
1064 [45] Goray M, Mitchell RJ, Oorschot RAHv. Investigation of secondary transfer of skin cells under
1065 controlled conditions. Legal Medicine. 2010;12:117-20.
1066 [46] Poetsch M, Bajanowski T, Kamphausen T. Influence of an individual's age on the amount and
1067 interpretability of DNA left on touched items. International Journal of Legal Medicine.
1068 2013;127:1093–6.
1069 [47] Ladd C, Adamowicz MS, Bourke MT, Scherczinger CA, Lee HC. A systematic analysis of secondary
1070 DNA transfer. Journal of Forensic Sciences. 1999;44:1270 - 2.
1071 [48] Van Oorschot RA, Jones M. DNA fingerprints from fingerprints. Nature. 1997;387:767.
1072 [49] Saravo L, Spitaleri S, Piscitello D, Travali S. DNA typing from steel cable. International Congress
1073 Series. 2004;1261:473–5.
1074 [50] Breathnach M, Williams L, McKenna L, Moore E. Probability of detection of DNA deposited by
1075 habitual wearer and/or the second individual who touched the garment. Forensic Science
1076 International: Genetics. 2016;20:53-60.
1077 [51] McKenna L. Understanding DNA results within the case context: importance of the alternative
1078 proposition. Frontiers in Genetics. 2013;4:1-4.
1079 [52] Sly N, Fietz P, Windram K, Windram R. The success of trace DNA recovery using STRmix. The
1080 22nd International ANZFSS Symposium. Adelaide, South Australia2014.
1081 [53] Bright JA, Taylor D, Curran JM, Buckleton JS. Developing allelic and stutter peak height models
1082 for a continuous method of DNA interpretation, Forensic Science International: Genetics. 2013;
1083 7:296–304.
1084 [54] Bright JA, Taylor D, Curran JM, Buckleton JS. Degradation of forensic DNA profiles, Australian
1085 Journal of Forensic Sciences. 2013; 45(4):445–449.
1086 [55] Bright JA, Taylor D, Curran JM, Buckleton JS. Searching mixed DNA profiles directly against
1087 profile databases, Forensic Science International: Genetics, 2014; 9:102–110.
1088 [56] Taylor D, Hicks T, Champod C. Using sensitivity analyses in Bayesian Networks to highlight the
1089 impact of data paucity and direct future analyses: a contribution to the debate on measuring and
1090 reporting the precision of likelihood ratios. Science & Justice. 2016;56:402-10.
1091 [57] Cook R, Evett IW, Jackson G, Jones PP, Lambert JA. A model for case assessment and
1092 interpretation. Science and Justice. 1998;38:151-6.
1093 [58] Biedermann A, Champod C, Jackson G, Gill P, Taylor D, Butler J, et al. Evaluation of forensic DNA
1094 traces when propositions of interest relate to activities: analysis and discussion of recurrent
1095 concerns. Frontiers in Genetics. 2016;(accepted).

1096

1097

1098

1099 ## Appendix 1: Tables that describe elements of the BN in Figure 8 to 11

1100

Node	Categories	Values
Hp and Hd	Hp Hd	uniform prior
Choice of Hd	Hp: Primary transfer Hd1: Secondary transfer without AO Hd2: Secondary transfer with AO	uniform prior or adapted as a function of the alleged activities put forward by the defence.
DNA on POI hand log10(ng)	interval node; From -inf to -1.5 From -1.5 to 3 in steps of 0.1 3 to 4 in steps of 0.5	N(0.66, 0.49)
DNA on AO hand log10(ng)	interval node; From -inf to -1.5 From -1.5 to 3 in steps of 0.1 3 to 4 in steps of 0.5	N(0.66, 0.49)
DNA POI Primary transfer (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	$10^{\text{DNA OUT}}$ from TP if Hp is true 0 otherwise
DNA POI Secondary transfer (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	$10^{\text{DNA OUT}}$ from TP if Hd1 or Hd2 is true, 0 otherwise
DNA AO Primary transfer (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	$10^{\text{DNA OUT}}$ from TP if Hd2 is true 0 otherwise
DNA Background (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5	uniform prior Note the probabilities provided in this node will need to be tailored to the item and circumstances of the case. A uniform distribution is not a realistic description of background DNA on most items.

	25 to 1000 1000 to inf	
DNA POI 1 (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA POI primary transfer + DNA POI secondary transfer
DNA POI 2 (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	AO matching POI + Background matching POI
DNA AO (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	AO DIFF POI + Background matching AO
DNA DIFF (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	Background DIFF AO
DNA POI persisting (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA POI 1+ DNA POI 2
DNA not POI Persisting (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1	DNA AO + DNA DIFF

	10 to 25 in steps of 5 25 to 1000 1000 to inf	
Results DNA POI	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA OUT from R
Results DNA not POI	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA OUT from R
Proportion of area sampled	interval node; values from 0 to 1 in steps of 0.1	Uniform Prior
Sampling device	Tapelift Swab	Uniform Prior
Sampling efficiency	interval node; from 0 to 1 in steps of 0.05	B(1.9,16.6) If Sampling device is Tapelift B(25,20) If Sampling device is Swab
Extraction efficiency	interval node; from 0 to 1 in steps of 0.05	Beta(5,17)

1101 *Tables 1: Expressions and probabilities for nodes that underlie the main BN in Figure 8*

1102

Node	Categories	Values
DNA IN	interval node; from -inf to -1.5 then from-1.5 to 3 in steps of 0.1 then from 3 to 4 in steps of 0.5	Input node
Proportion of hand/surface contacted	interval node; from 0 to 1 in steps of 0.1	uniform prior
DNA available to transfer log10(ng)	interval node; From -inf to -1.5 From -1.5 to 3 in steps of 0.1 3 to 4 in steps of 0.5	DNA IN + $\log_{10}$ (Proportion)
Vigour of contact with target surface	Passive Pressure Friction	Uniform Prior
Nature of shedding surface	Hand Smooth Rough	Uniform Prior
Nature of	Hand	Uniform Prior

target surface	Smooth Rough	
Transfer proportion	interval node; from 0 to 1 in steps of 0.05	B(0.89,2.25) if Surface=Rough/Hand & Vigour= Passive B(0.49,24.11) if Surface= Smooth & Vigour= Passive 0.33B(0.77,1.45) + 0.67B(1,1) if Surface= Rough/Hand & Vigour= Pressure B(0.47,30.85) if Surface= Smooth & Vigour= Pressure 0.2B(0.7,1.64) + 0.8B(1,1) if Surface= Rough/Hand &Vigour= Friction B(0.45,1.13) if Surface= Smooth & Vigour= Friction
DNA deposited on target surface log10(ng)	interval node; From -inf to -1.5 From -1.5 to 3 in steps of 0.1 3 to 4 in steps of 0.5	DNA available to transfer + $\log_{10}$(Transfer proportion)
Days between both transfers or transfer and recovery	interval node; 0, 0.5 then 1 to 31 in steps of 1	Uniform Prior
Environmental condition of target surface kept in	Favourable Poor	Uniform Prior
alpha 2	-0.052 -0.022	1 if Environmental is poor, 0 otherwise 1 if Environmental is favourable, 0 otherwise
DNA persisting on target surface (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	$10^{\text{DNA deposited}}\ e^{\text{alpha2 * Days}}$
DNA OUT	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	$\log_{10}$(DNA persisting)

1103 *Tables 2: Expressions and probabilities for nodes that underlie the main BN in Figure 9*

1104

Node	Categories	Values
DNA IN	interval node; 0 to 0.1 in steps of 0.01	Input node

	0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	
Proportion of area sampled	interval node; values from 0 to 1 in steps of 0.1	Input node
Sampling efficiency	interval node; from 0 to 1 in steps of 0.05	Input node
DNA recovered from sampling	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA IN * Proportion* Sampling efficiency
Extraction efficiency	interval node; from 0 to 1 in steps of 0.05	Input node
DNA recovered from extraction	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA recovered from sampling * Extraction efficiency
DNA OUT	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA recovered from extraction

1105 *Tables 3: Expressions and probabilities for nodes that underlie the main BN in Figure 10*

1106

Node	Categories	Values
DNA (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	Input node

Match Probabilities	True False	True with probability between 0.3 and 10E-9 depending on DNA profile
DNA Matching (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA (ng) if Match Probabilities true, 0 otherwise
DNA DIFF (ng)	interval node; 0 to 0.1 in steps of 0.01 0.1 to 1 in steps of 0.1 1 to 5 in steps of 0.5 5 to 10 in steps of 1 10 to 25 in steps of 5 25 to 1000 1000 to inf	DNA (ng) if Match Probabilities false, 0 otherwise

1107 *Tables 4: Expressions and probabilities for nodes that underlie the main BN in Figure 11*

1108

1109

1110

1111 **Appendix 2: Further studies which would assist in answering questions of transfer**
1112 **mechanism**

1113 It is important that the studies reflect as much casework circumstances as possible: studies
1114 that try and maximise transfer cannot be used for casework purposes. Data collection and
1115 data treatment should also reflect casework procedures. Here, we have used data where
1116 quantity of DNA was available. What would be ideal is to report the quantity of DNA relative
1117 to each contributor. This can be estimated by combining the quantity and the relative
1118 contribution of the persons to the mixture (such as in [41]). Through this work we have
1119 identified a number of studies which would assist in evaluation of evidence in light of
1120 propositions that suggest differing DNA transfer mechanisms. We collectively provide the
1121 list below:

1122 1. Trialling extraction efficiencies at low DNA levels. The study of Butts [28] trialled
1123 DNA amounts from 24ng to 4800ng.
1124 2. Trialling of different extraction techniques not tried in the Butts [28] study. Note that
1125 some information to this effect can be found in [29].
1126 3. Taking into account the nature of the surface type the DNA has been placed on when
1127 considering DNA persistence.
1128 4. Investigation into different environmental conditions (e.g. rain, washed, full sun, etc)
1129 on DNA persistence.
1130 5. More data to confirm the DNA persistence rates found in Raymond *et al.* [32].
1131 6. Data on DNA persistence of DNA on objects after extended handling by other
1132 individuals (of which some work has been done in [34]), or from physical movements
1133 after initial deposition.
1134 7. Data on DNA persistence in standard exhibit packaging
1135 8. Shedder consistency studies, i.e. whether an individual sheds DNA consistently in the
1136 upper or lower quantiles of the population shedding distribution and whether this is
1137 noticeable through events such as washing or sweating. Studies could extend to a
1138 standard method for determining the approximate shedding propensity of an
1139 individual for use in primary vs secondary transfer considerations.
1140 9. A study of how the amount of time an object is held affects the amount of DNA
1141 transferred. To date we have found such information in [47] for hand to hand transfers
1142 only by length of handshake.

1143 10. A study of the absolute amount of DNA transferred from hands to objects for different
1144 contact types e.g. light touch, pressure, friction (i.e. so that data from multiple studies
1145 does not need to be combined and extrapolated as we have done here).
1146 11. Consideration of the amount of DNA transferred to object from a habitual use e.g.
1147 items in the home. Some work has been done in this area, such as [50].
1148 12. Transfer DNA amount for varying length of time of contact between primary and
1149 secondary substrates and for different contact types.
1150 13. A study into the level of transfer from hand-to-hand for various contact types (e.g.
1151 handshake, high-five, clasping, struggling) for different times of contact and the
1152 persistence of DNA on hands through various timeframes and activities.
1153 14. The levels of background DNA (i.e., not from known users) on various items.
1154 15. The levels of primary user's DNA on regular use items such as furniture or objects
1155 around the home or office.
1156 16. A study into the accumulation of contact DNA on items resulting from multiple
1157 contacts from the same person.
1158 17. All studies used in the construction of the *BN* in this paper concentrate on DNA
1159 amount. The rate of degradation across the profile may also have some power to
1160 distinguish factors involved in transfer, persistence and recovery (particularly
1161 persistence). There are a number of experiments, that would be worthy of study in this
1162 area. For those experiments that have already been carried out (many of which we
1163 reference) the data already exists, it just requires analysis of a different measured
1164 variable.