J Forensic Sci, May 2004, Vol. 49, No. 3
Paper ID JFS2003366
Available online at: www.astm.org

Peter Gill,[1] *Ph.D. and Amanda Kirkham,*[1] *B.Sc.*

Development of a Simulation Model to Assess the Impact of Contamination in Casework Using STRs

ABSTRACT: Because contamination is usually tube-specific, negative controls cannot give assurance that an associated batch of extracted casework material is contaminant-free. However, it is possible to use them to predict the level of overall (undetected) contamination that is processed by an operational DNA unit. A MATLAB®-based program was used to combine results of negative controls with actual casework DNA profiles to assess the probability that laboratory contaminants will give rise to reportable profiles (along with their likelihood ratios). Using data from an operational DNA unit as an example, it was demonstrated that the risk is inextricably linked to guidelines used to interpret DNA profiles. We have demonstrated how computer-based models can predict the levels of contamination expected in the process and, in addition, how the process can be made more robust by changing reporting guidelines. There is a need to compare DNA profiles against staff and plasticware elimination databases in order to determine sources of contamination. The likeliest outcome of a contamination event is false exclusion.

KEYWORDS: forensic sciences, contamination, short tandem repeats, negative control, national DNA database, simulation

In this paper we use a computer simulation model to estimate potential levels of contamination that may be encountered in routine casework; we also demonstrate how the effect of contamination can be mitigated by reference to the simulation model.

First of all, it is important to make the distinction clear between the meaning of the terms "contamination" and "adventitious transfer" of DNA. Gill et al. (1) originally outlined a model where the definition of contamination was specifically restricted to describe the introduction of extraneous DNA into the process directly as a result of the intervention of a police investigator or scientist. The alternative kind of contamination, namely adventitious transfer, refers to the transfer of DNA from sources that may be unconnected to the case and the investigator, for example by secondary transfer from the perpetrator (2). The crucial point is that adventitious transfer occurs only before the crime scene is established, whereas contamination can occur only afterwards.

At the biochemical level, contamination may manifest itself in two nonexclusive ways. Either as allele "drop-in"—described by Gill et al. (3) where spurious alleles from multiple independent sources may find their way into a polymerase chain reaction (PCR)—or as a partial or complete DNA profile from a single (rather than from multiple) source. There are certain other generalizations to be made. First, no process is entirely free of contamination. Secondly, contaminants tend to be quantitatively at low level. This means that casework profiles that are also low level, partial, and consequently difficult to interpret, tend to be affected more than full profiles that are associated with substantial amounts of DNA. Thirdly, contamination can be characterized by reference to negative controls.

There are various sources of contamination that are known to occur. Each can be monitored by reference to staff elimination databases and negative controls in order to estimate or capture the effect:

1. Transfer from scientists and police investigators, e.g., skin flakes or saliva spray—prevented by good laboratory practice (gowns and face masks) and monitored by reference to staff databases that consist of relevant personnel.
2. Reagent contamination. Plasticware may be contaminated during the manufacturing process—sterile does not equate to DNA-free. Schmidt et al. (4) showed that reaction tubes used to analyze mtDNA were often contaminated. We have subsequently demonstrated several casework-related STR contamination events were derived from staff of a reaction tube manufacturer (after voluntary screening was carried out). Monitoring of consumables prior to introduction into casework (on a batch basis) is important to establish their quality and to ensure that manufacturers are not inadvertently contaminating their products. Ideally, a comparison of casework samples and negative controls against staff elimination databases from reagent manufacturers is needed.
3. The contaminant may arise from another sample that has been processed concurrently and cross contamination has resulted (e.g., lane-to-lane leakage on a flatbed gel or between adjacent wells of microtitre plates on automated systems). This can be monitored by software that compares samples within a batch for potential duplicate profiles—interpretation of mixtures may be necessary.

Multiplexes such as Applied Biosystems AmpF*l*STR SGM plus (5) are very sensitive using manufacturer recommended methods, down to 250 pg, or less than 100 pg if increased PCR cycles (3) are utilized—this means that just several cells may be required to give a signal. Low level contamination tends to be sporadic (i.e., tube specific). The logical consequence of this is that a negative control does not give confidence that the associated batch of extractions is contaminant free. The only exception to this is when the contaminant is observed both in the negative control and in an associated extraction tube(s), in which case the contamination is considered gross, i.e., multiple events that are possibly reagent based.

[1] Forensic Science Service, Trident Court, Birmingham Business Park, 2960 Solihull Parkway, Solihull, B37 7YN, UK.

Received 1 Nov. 2003; and in revised form 19 Jan. 2004; accepted 19 Jan. 2004; published 7 April 2004.

The purpose of this paper is to illustrate how negative controls can be used to predict the performance of a DNA unit in terms of an estimate of the expected number of false positive results along with their associated likelihood ratios. In addition we show how the use of reporting guidelines affects the results. Because contamination tends to be low level, this means that the lower the reporting threshold, the greater the chance that a contaminant will be reported.

For example, by analyzing negative control data, and casework data from a Forensic Science Service (FSS) operational laboratory, we demonstrated that under a reporting guideline utilizing a 50 rfu peak height lower threshold there was a chance of approximately 1 in 1000 that a sample would be a false positive with a likelihood ratio greater than 10^7. However, simply by raising the reporting lower threshold from 50 rfu to 80 rfu was sufficient to negate the effect.

This does not mean that low-level profiles cannot or should not be reported, rather the relevance of low-level DNA evidence is less certain. In particular, interpretation at the source level (i.e., association of the DNA profile with a particular body fluid) is uncertain. The most probable outcome of a contamination event is false exclusion. A framework to report low level DNA samples under hierarchy of propositions principles using Bayesian networks is given by Evett et al. (6)

Method

Negative Controls

To estimate the level of false positive reporting, a suite of MATLAB® (The Mathworks Inc., Natick, MA) programs (SIMLAB) was written to simulate the casework environment. The specific purpose was to predict the overall level of false positive results that may occur as a result of sporadic and undetected laboratory contamination. The term sporadic and undetected contamination specifically relates to:

1. Extraneous DNA in plasticware, solutions or other reagents that have been introduced during the manufacturing process (i.e., external to the laboratory).
2. Contamination from personnel who are not part of a staff elimination database and therefore undetected.

A contaminant may be detected if:

3. The contaminant has arisen from an operator who is on a staff elimination database—either the forensic lab or plasticware manufacturer.
4. The contaminant has come from another sample that has been processed concurrently and cross contamination has resulted (e.g., lane-to-lane leakage on flatbed gel or a microtitre plate used in automated systems).

In this paper only inexplicable contamination events (in 1,2 above) are simulated; i.e., profiles explained by staff contamination and cross contamination are not included in the analysis.

To determine overall levels of contamination and to estimate the associated effects, we examined negative extraction controls. These controls are a microcosm of casework samples that are assumed devoid of DNA. Nevertheless, reportable DNA profiles are sometimes observed. Because negative controls are processed in the same way as casework samples, they can be used to estimate the level of contamination in casework samples over the same period of time.

Most, if not all, contamination events seen in negative controls are sporadic single-tube events—i.e., the contaminant is specific to one tube only. Generally, this means that the contaminant has no relevance to the associated batch of extracted samples.

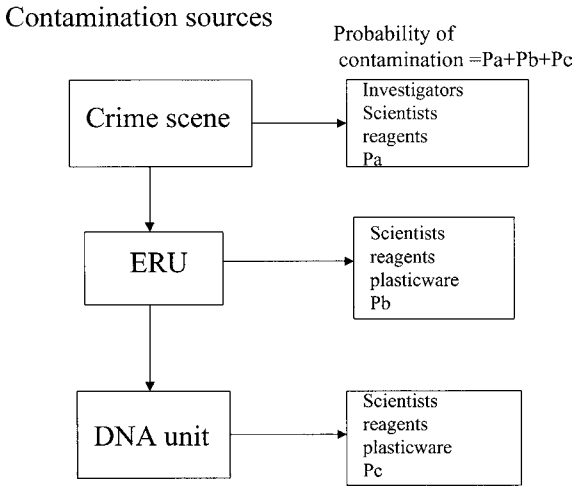

FIG. 1—*Flow diagram to outline potential sources of contamination.*

This does not mean to say that negative controls are useless, far from it; *however, we have to interpret them in a way that relates to the entire DNA process, not just in relation to a specific batch of samples with which the negative controls have been co-processed.* It may be believed that quality is assured simply by eliminating a batch of samples where the negative control gives a result—but this ultimately leads to a false sense of security. To date, the effect of sporadic contamination on the DNA process has not been considered and is therefore poorly understood.

How Contamination Arises and How the Process Can be Monitored to Produce Predictive Models of the Consequences

To be completely effective, any model must take account of the entire process. In the Forensic Science Service (FSS) the case is submitted to an evidence recovery unit (ERU) where it is evaluated and stains are then submitted to the DNA unit for processing. We can break down the origin of contaminants into three discrete consecutive categories (Fig. 1):

1. At the crime scene—contamination by investigating officers or reagents used to collect evidence (P_a).
2. Similarly, when the case is transferred to the ERU, where evidence is evaluated, contamination may result from scientists, reagents, or plasticware (P_b).
3. In the DNA unit the same as category 2 applies (P_c).

Each category comprises several subcategories; for example, P_b could comprise separate probabilities for swabs, plastic tubes, scientists, other reagents, etc. Over time, contaminants will pass from the crime scene to the ERU and finally to the DNA unit for analysis; at each stage there is additional opportunity for contamination to occur so that the process is additive.

To summarize, the chance of contamination (C) is subdivided into several non-exclusive categories such that $C = P_a + P_b + P_c$ (Fig. 1). If a control is to reflect the entire process it purports to control, ideally it should be prepared at the crime scene in order to capture P_a. Consideration also needs to be made about the kind of negative control that is employed.

For example, if a moistened swab is used to collect evidence, then the ideal negative control would be an additional blank swab also

moistened with water and concurrently prepared. Similarly, within the ERU, negative controls to measure P_b could also be prepared and cascaded to the DNA unit. If the only part of the process that produces negative controls is the DNA unit itself, then only P_c is estimated and this is therefore an underestimate of the total risk of contamination.

Method to Estimate P_c

The method illustrates how to estimate P_c measured from negative controls that are prepared and generated by the DNA analysis unit itself. In principle, P_a and P_b can also be estimated provided the appropriatenegative controls are collected and processed. The standard AmpF*l*STR SGM plus® (Applied Biosystems) system was used as described by Cotton et al. (5) using the standard 28 PCR cycles.

Collection of Data

Two subsets of data were collected from a DNA unit:

1. All 295 negative controls over the period of time for which an assessment was made (a period of five months—10 May 02 to 14 Dec. 02), including samples for which no signal was obtained.
2. A random collection of 50 casework DNA profiles over the same period of time, including samples that failed to give any signal.

Details of SIMLAB Computer Program

Each profile is represented in a single row of a spreadsheet format where each locus is defined by three parameters—allele designation, peak height, and peak area. Only true allelic peaks that are within their expected size ranges are used in the analysis.

Profiles from negative controls are similarly prepared.

All blank negatives and casework samples are included in the simulation regardless of whether they generate profiles. Thus in a sample of 295 negative controls 26 showed evidence of contamination. Conversely, in the sample of casework profiles, five failed to show a result.

Pairwise comparison was used to combine casework with contaminants observed in negative controls. For 50 casework samples and 295 negative controls there are 50×295 simulated casework/contaminant profiles. There are four possible outcomes: casework sample only, contaminant only, casework/contaminant mixture, or a blank result with no profile apparent.

The simulated profiles are analyzed further. For each sample, the mixture proportion (M_x) is approximately estimated by summing peak areas of contaminants and casework samples and calculating:

$$M_x = \text{peak heights negative controls/peak heights casework samples}$$

(Note: this could be improved by calculating M_x per locus and treating each locus separately.)

Once M_x has been calculated, then reporting guidelines are applied for each locus in turn as follows:

(*a*) Is the allele peak height greater than the threshold reporting level (e.g., 50 rfu)? Below this level the allele is deemed inconclusive and not reportable (falls into the low copy number (LCN) category not considered further by the program).

(*b*) Is the locus a mixture?

Further reporting guidelines are applied as follows: If the locus is a mixture, then is the major component a contaminant and is it reportable? To fulfill this requirement the peak height must be greater than the threshold reporting level and must also be distinct from the minor component of the mixture. To measure whether the contaminant profile is distinct from the casework sample a rule is applied to the effect that M_x must be >0.5; i.e., on average, contaminant peaks must be 50% higher than casework peak heights in order to be accepted. Peaks that fail the rule are deemed inconclusive.

(*c*) The surviving peaks are deemed reportable and then the major profile is converted into a likelihood ratio. Note that an important feature of the guideline is the ability to run "what-if" scenarios simply by changing the reporting parameters to any desired level. Changing reporting parameters alters the number of reportable peaks and the outputs are given in terms of reportable profiles that can be converted into likelihood ratios (or match probabilities).

Results and Discussion

Pre-assessment of the Data

(*a*) *Casework:* Out of the 50 samples analyzed, five failed to give a result; i.e., the probability of a sample failing to give any profile (P_F) $= 5/50 = 0.1$.

(*b*) *Negatives:* Out of 295 samples analyzed, a total of 26 samples gave a signal. This means that the probability of negative control giving a profile (P_N) of one or more alleles is $26/295 = 0.088$.

A contaminant is only detected or known to have occurred if it is found in a negative control tube purported to be free of DNA. The difficulty is that it is not possible to assess directly whether a casework sample is affected by sporadic contamination, as there is no supporting information. However, even though we cannot know which particular casework tube is contaminated, unless it matches a profile on a staff elimination database, we can assess the probability (P_c) of any given tube being affected because negative samples are simply a subset of casework samples. It is the same as the probability that a negative control is contaminated (where the contamination event may be one or more alleles)

$$P_N = P_c = 0.088$$

If a casework sample is contaminated then this will result in one of two different outcomes:

(*a*) If the casework sample is devoid of DNA, then only the contaminant will be visible and the profile appears unmixed. The chance of this occurrence (P_S) is the probability of contamination multiplied by the probability of a casework sample failing to give a profile:

$$P_S = P_N \times P_F$$

Specifically, in the DNA unit assessed $P_S = 0.009$ (or approximately 0.9% of samples will be contaminated and profile does not appear admixed).

(*b*) In addition, P_N–P_S, i.e., $0.088 - 0.0088 = 0.079$ (or 7.9%) of casework samples will contain sporadic contamination in admixture with a casework profile.

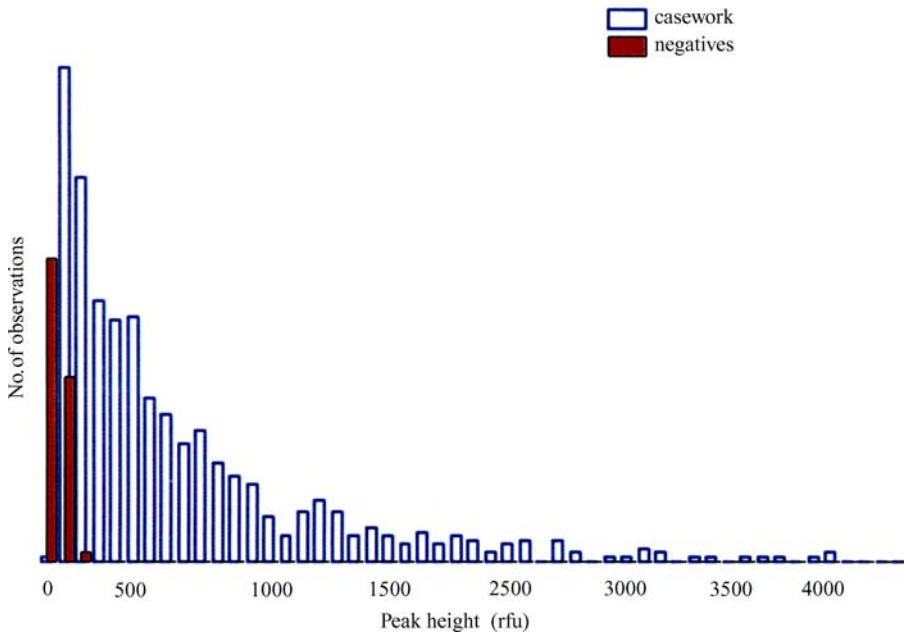

FIG. 2—*Comparison of peak heights >0 across all loci, plotted for negative controls (a population of 26 samples) and for casework samples (a population of 50 samples).*

Negative Controls Define When the Low Copy Number Framework Should Be Used for Reporting Purposes

The low copy number (LCN) probabilistic framework of reporting was originally developed by Gill et al. (3,7) in order to enable reporting of sub c. 100 pg quantities of DNA . The method was specifically designed to take account of sporadic contamination ("drop-in") and stochastic effects that led to allelic "dropout"; probabilities of both were factored in into likelihood ratio calculations. Duplicate analysis and derivation of a consensus profile are used. Furthermore, we recognized that the relevance of the evidence (origin of the sample tested and circumstances of deposition) was less certain compared with substantial profiles from large quantities of DNA derived from an obvious body fluid stain. Consequently, statements are specially tailored to reflect this uncertainty. In the United Kingdom, LCN is routinely reported to the courts. Although it is widely perceived that LCN refers to the use of DNA profiling under conditions that increase the sensitivity of detection by increasing the number of PCR cycles or increasing the injection time with CE instrumentation, in fact this is not the case. LCN has nothing to do with the ability to detect minute quantities of DNA; rather the decision to use the low copy number analysis and reporting framework is dominated by "drop-in" or contamination and this also occurs with conventional DNA profiling (note all data analyzed in this paper were derived by conventional methods). This is why it is useful to collate information from negative controls, recognizing that this is a direct mimic of the casework position. For example, if peaks c. 80 rfu are observed in 1 in 1000 negatives, then the same frequency of contamination is expected in casework samples. The 80 rfu observation in this example defines the "drop-in" threshold below which any alleles would be reported using the LCN framework, to take account of the fact that some may be present due to contamination. This example is somewhat oversimplified but serves to make the point. Whereas sporadic contamination will affect samples on a regular basis, it does not give an indication of the actual impact on casework reporting. To do this assessment properly, it is necessary to evaluate negative control and casework data in much greater detail with special emphasis on their relative peak areas or heights. Peak heights of negative control and casework data were combined and plotted across loci (Fig. 2). Note that the current data set used in the current analysis was rather limited. More data are preferable because a more comprehensive analysis could ensue; e.g., separate characterization of individual loci would be advantageous. The purpose of this paper is to provide a demonstration, in principle, of the proposed methodology, rather than a definitive analysis.

The majority (58%) of contaminant peak heights were <50 rfu. However, 42% were greater and consequently overlapped with case stain peak heights up to c. 150 rfu. We can summarize that in casework c. 17% of alleles were >50 <150 and c. 70% of data were <250 peak height (Fig. 2).

Can a Mixture of a Contaminant with an Evidential Sample Result in Mistyping?

The question arises, as a consequent of the foregoing, whether a mixture of a contaminant with a case sample can give a misleading result.

A MATLAB® program was used to rank the sum of peak heights of strongest weakest negative controls (Table 1) and weakest–strongest casework samples (Table 2), respectively. The worst scenario occurs when a strong contaminant combines with a weak or absent casework sample.

Characterization of Mixtures

Mixtures were simulated by using pairwise comparisons of casework versus negative control data (including all examples where profiles were absent). This means that from 50 casework samples and 295 negative controls, pairwise comparisons generated $50 \times 295 = 14750$ mixtures.

The simulated results comprised:

- Unmixed case samples only (82%)
- Unmixed contamination (0.9%)
- A mixture of case sample and contaminant (7.9%)
- No DNA profile detected (9.1%)

TABLE 1—*Case samples ranked in order of increasing summed peak height with numbers of alleles scored above a given peak height.*

No. allele>0 rfu	No. alleles>25 rfu	No. alleles>50 rfu	No. alleles>100 rfu	No. alleles>150 rfu	Sample No.
0	0	0	0	0	5
0	0	0	0	0	10
0	0	0	0	0	38
0	0	0	0	0	44
0	0	0	0	0	49
1	1	1	0	0	23
2	2	2	0	0	30
4	4	4	3	1	29
5	5	5	3	3	47
7	7	7	3	3	32
13	13	13	7	2	31
13	13	13	8	4	14
16	16	15	6	3	9
13	13	13	10	5	42
13	13	13	10	5	50
14	14	14	9	7	15
10	10	10	9	9	39
17	17	17	15	14	4
19	19	19	16	14	8
19	19	19	15	12	35
22	22	22	21	19	12
19	19	19	19	19	22
12	12	12	9	8	17
14	14	14	14	14	34
16	16	16	16	16	13
21	21	21	21	21	45
18	18	18	18	18	41
18	18	18	16	16	43
22	22	22	20	19	7
15	15	15	15	15	26
20	20	20	20	20	2
20	20	20	20	20	40
21	21	21	21	21	46
19	19	19	19	19	27
20	20	20	20	20	16
20	20	20	20	20	11
18	18	18	18	18	24
20	20	20	20	20	37
20	20	20	20	19	18
22	22	22	22	22	33
22	22	22	22	22	48
20	20	20	20	20	3
19	19	19	19	19	25
21	21	21	21	21	6
18	18	18	18	18	19
19	19	19	19	19	20
20	20	20	20	20	28
21	21	21	21	21	1
19	19	19	19	19	21
21	21	21	21	21	36

The mixture proportion (M_x) was calculated as $M_x =$ sum peak heights contaminant/sum peak heights casework samples (8)—the distribution of M_x is given in Table 3. Most mixtures gave $M_x < 1$, which means that the casework sample was usually the major component. In approximately 1 in 500 cases the major component was the laboratory contaminant; in the most extreme example $M_x = 25$.

Likelihood Ratios of Reportable Contaminant Profiles

Finally, we ask the question, What does this mean in practice in terms of casework reporting and in terms of the national DNA database (NDNADB)? Cases are reported in terms of likelihood ratios. We calculated the likelihood ratio of DNA profiles that originated from contamination. Following laboratory guidelines, alleles were not incorporated into calculations unless above the LCN threshold (currently 50 rfu). Only unmixed contaminants were collated in Table 4; inclusion of major contaminant profiles in mixture data made little difference because of their relative rarity. From Table 4 and Fig. 3, the chance of a laboratory contaminant resulting in a reportable profile LR$>10^7$ was approximately 1 in 1000. We then assessed the effect of changing the current 50 rfu reporting guideline. We demonstrated that the risk reduced if the guideline was increased—for example, if the LCN guideline was rfu = 80, then the maximum LR observed was 10^2.

Data analysis showed that 82% samples were case samples only, and 0.9% of samples will give an unmixed profile from a laboratory contaminant. The overwhelming majority of mixtures will be $M_x < 1$, which means that the casework sample was the major component in most cases. In approximately 1 in 500 cases the major component was the laboratory contaminant. Consequently, this means that by far the greatest problem is when the laboratory

TABLE 2—*Negative controls ranked in descending order of intensity, taken from a population of 295 negative controls—only 26 controls that gave a signal are listed (i.e., 275 controls were blank).*

Sum Peak Height	No. alleles > 0	>25	>50	>100	>150	>200	>250	Sample No.
1217	16	16	16	5	3	0	0	26
1109	17	17	9	4	2	2	1	15
695	11	10	6	1	1	1	1	3
481	11	11	3	0	0	0	0	16
413	8	6	2	2	1	0	0	1
334	10	9	0	0	0	0	0	8
290	7	6	2	0	0	0	0	14
242	7	5	1	0	0	0	0	4
234	5	4	3	0	0	0	0	2
226	4	4	2	0	0	0	0	9
140	3	3	1	0	0	0	0	20
104	4	2	0	0	0	0	0	17
67	2	2	0	0	0	0	0	6
64	1	1	1	0	0	0	0	18
60	1	1	1	0	0	0	0	5
51	2	1	0	0	0	0	0	24
50	1	1	0	0	0	0	0	7
50	2	1	0	0	0	0	0	10
50	1	1	0	0	0	0	0	19
46	2	1	0	0	0	0	0	13
36	1	1	0	0	0	0	0	12
36	1	1	0	0	0	0	0	25
33	1	1	0	0	0	0	0	21
33	1	1	0	0	0	0	0	23
31	1	1	0	0	0	0	0	11
16	1	0	0	0	0	0	0	22

TABLE 3—*Analysis of observations relative to M_x = sum peak heights negatives/sum peak heights casework samples. When in admixture, laboratory contaminants give profiles that are greater in size than casework samples when $M_x > 1$ (marked in bold type).*

Mixture (M_x)	No. Observations	Probability
Case sample only	12105	0.8207
Negative sample only	130	0.0088
No sample	1345	0.0912
<= 0.1 > 0	973	0.0660
<= 0.2 > 0.1	73	0.0049
<= 0.3 > 0.2	30	0.0020
<= 0.4 > 0.3	19	0.0013
<= 0.5 > 0.4	9	0.0006
<= 0.6 > 0.5	9	0.0006
<= 0.7 > 0.6	9	0.0006
<= 0.8 > 0.7	4	0.0003
<= 0.9 > 0.8	6	0.0004
<= 1 > 0.9	5	0.0003
<= 2 > 1	**11**	**0.0007**
<= 10 > 2	**19**	**0.0013**
<= 25 > 10	**3**	**0.0002**
Total	14750	

TABLE 4—*Probability estimates for achieving a given likelihood ratio where a laboratory contaminant is responsible for the major (unmixed) profile.*

Log 10 LR	Guideline (rfu)					
	rfu = 50	rfu = 60	rfu = 70	rfu = 80	rfu = 90	rfu = 100
1	0.00746	0.00502	0.00319	0.00251	0.00339	0.00339
2	0.00217	0.00095	0.00014	0.00088	0	0
3	0.00027	0.00041	0.00210	0	0	0
4	0.00095	0.00183	0	0	0	0
5	0.00014	0	0	0	0	0
6	0.00007	0	0	0	0	0
7	0.00020	0	0	0	0	0
8	0.00088	0	0	0	0	0
9	0	0	0	0	0	0

contaminant appears as a nonmixed sample. Simply reducing the failure rate will reduce the risks of reporting a contaminant.

Thompson et al. (9) suggest that any level of contamination reduces likelihood ratios. Whereas we would not disagree with the principle, each case needs separate consideration. In particular, we have shown that there is no one simple error rate that can be universally applied regardless of the circumstances—it is dependent upon several factors, in particular the quantity of DNA analyzed and the associated reporting lower threshold limit. In addition, replicate analyses of different stains or areas of the same stain will also significantly reduce the impact of potential error. Good quality management system (QMS) feedback results in continually updated procedures and in turn reduces potential error rates. The key to understanding errors, their impact, and enabling continuous improvement of existing analytical processes is dependent upon instigation of monitoring systems such as those outlined here. Once the data are collected they can be used to inform Bayesian networks (6) that incorporate the contamination rates into probabilistic formulae.

The primary risk of contamination is wrongful exclusion, particularly if the contaminant masks the perpetrator's profile. However, it is important to consider that most contamination incidents will result in partial DNA profiles for which random searches of DNA database will produce adventitious matches with low match probabilities. (Given that there are c. 2 m samples on the UK national DNA database a sample with match probability c. 10^{-6} would often result in one or more adventitious matches.) In itself this should not be problematic provided the non-DNA evidence is always carefully considered within the context of an "intelligence database." The primary purpose is to supply a list of potential suspects for further investigation. This follows principles

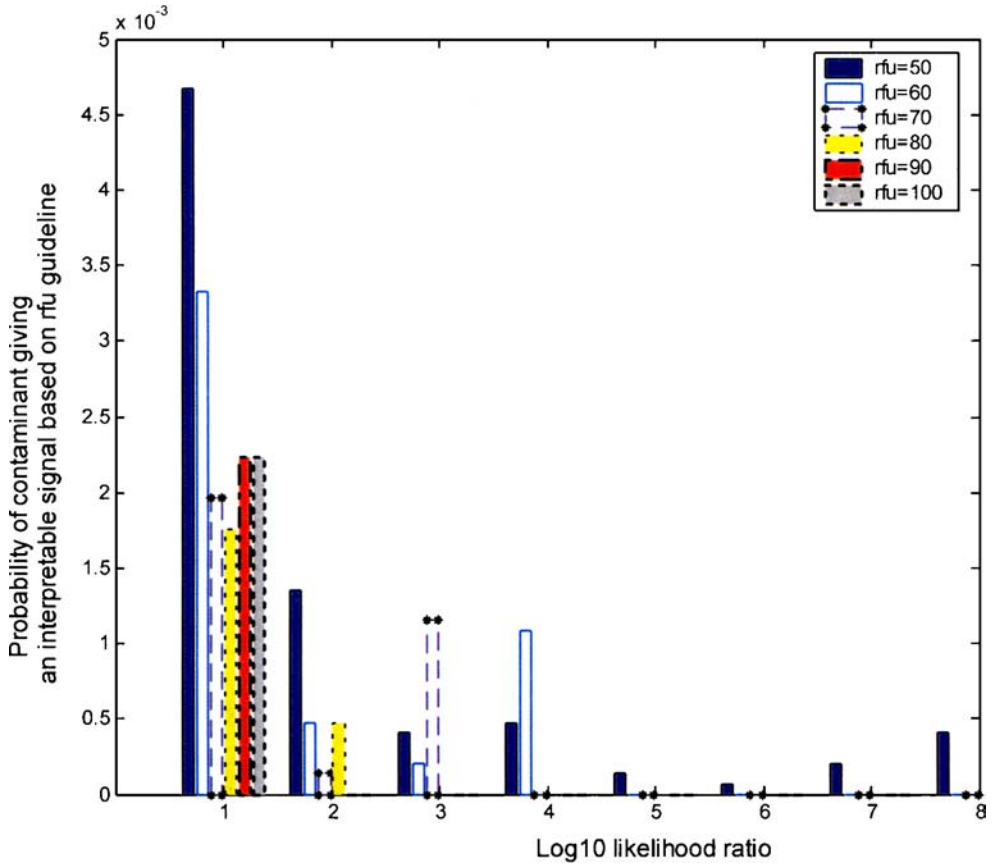

FIG. 3—*Histogram showing probability of a contaminant giving a reportable result (measured as log_{10} LR) relative to the reporting guideline (currently 50 rfu).*

of "hierarchy of propositions" developed by Cooke et al. (10,11) and Evett et al. (12) which has led to a much deeper understanding of the interpretation process. The potential of a result arising from a contamination incident is considered on a case-by-case basis. The approach of combining different kinds of evidence and incorporating contamination as part of the framework can be formalized by utilizing a probabilistic Bayesian network approach (6), and we consider this to be the best way forward.

Finally, it should be recognized that laboratory contamination is impossible to avoid completely but its extent is generally unknown unless proactively assessed—the probability of contamination must always be greater than zero. The effect can be mitigated by comparing all casework profiles against staff databases (including manufacturers of plasticware) but this can capture only a proportion of the events. It is possible to assess the prevalence of contamination and its effect using expert systems that analyze negative controls.

References

1. Gill P. Role of short tandem repeat DNA in forensic casework in the UK—past, present and future perspectives. Biotechniques 2002;32:366–85. [PubMed]
2. Lowe A, Murray C, Whitaker J, Tully G, Gill P. The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces. Forensic Sci Int 2000;129:25–34.
3. Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Forensic Sci Int 2000;112:17–40. [PubMed]
4. Schmidt T, Hummel S, Herrmann B. Evidence of contamination in PCR laboratory disposables. Naturwissenschaften 1995;82:423–31. [PubMed]
5. Cotton EA, Allsop JL, Guest JL, Frazier RRE, Koumi P, Callow IP, et al. Validation of the Ampf/STR SGM plus system for use in forensic casework. Forensic Sci Int 2000;112:151–61. [PubMed]
6. Evett IW, Gill PD, Jackson G, Whitaker J, Champod C. Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks. J Forensic Sci 2002;47:520–30. [PubMed]
7. Gill P. Application of low copy number DNA profiling. Croatian Med J 2001;42:229–32.
8. Gill P, Sparkes R, Pinchin R, Clayton T, Whitaker J, Buckleton J. Interpreting simple STR mixtures using allele peak areas. Forensic Sci Int 1998;91:41–53. [PubMed]
9. Thompson WC, Taroni F, Aitken CG. How the probability of a false positive affects the value of DNA evidence. J Forensic Sci 2003;48:47–54. [PubMed]
10. Cooke R, Evett IW, Jackson G, Jones PJ, Lambert JA. A hierarchy of propositions: Deciding which level to address in casework. Sci Just 1998; 38:231–9.
11. Cooke R, Evett IW, Jackson G, Jones PJ, Lambert JA. Case preassessment and review in a two-way transfer case. Sci Just 1999;39:103–11.
12. Evett IW, Jackson G, Lambert JA. More on the hierarchy of propositions: Exploring the distinction between explanations and propositions. Sci Just 2000;40:3–10.

Additional information and reprint requests:
Peter Gill, Ph.D.
Forensic Science Service
Trident Court
Birmingham Business Park
2960 Solihull Parkway
Solihull, B37 7YN
United Kingdom