

Principal Forensic Services Ltd.

Statistical Study: Report

September 2014 v1.3

Dr Gillian Tully & Dr Susan Pope

Contents

Summary of Findings.....	3
Introduction	6
Scope	6
Outputs.....	6
Methods & Data	7
Data Collection	7
Data Analysis.....	7
Evaluation of the Expected Scale of Adventitious Matches.....	8
Evaluation of the Expected Scale of Adventitious Matches: Bulk Exchange ...	9
Evaluation of the Expected Scale of Adventitious Matches: ongoing exchange	9
Evaluation of the expected scale of true matches.....	10
Results.....	11
Evaluation of the Expected Scale of Adventitious Matches: Bulk Exchange	11
Evaluation of the Expected Scale of Adventitious Matches: ongoing exchange	14
Evaluation of the expected scale of true matches.....	15
Discussion & Conclusions	19
Basis of Recommendations on what composition of profiles should be exchanged between UK and other EU MS, and Expected Scale of Adventitious Matches	19
8-locus matches and above: UK crime scene profiles vs. MS databases	19
6- and 7-locus matches: UK crime scene profiles vs. MS databases	20
UK Subject Profiles: Comparison versus other MS databases	21
Basis of anticipated match rate that would be produced when the UK initially engage in Prüm DNA and search their crime scene stains (as a bulk exchange) to other Member States	22
Match Validation Arrangements	22
Acknowledgements.....	23
Abbreviations (and definitions).....	23
References	24
Appendix 1: Data Returns from Member States	25



Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

Appendix 2: Assumptions & Simplifications	46
Appendix 3: Supporting information and data.....	47

Summary of Findings

1. **Recommendations** on composition of profiles that should be exchanged between UK and other EU MS; the basis for the recommendations is discussed in detail in the body of the report.
 - a. It is recommended that crime profiles with 8 or more loci (and which have not previously matched against a subject record) be compared against the databases of all other MS, to identify all 8+ locus matches.
 - b. Where matches of interest are obtained with 8 or 9 loci, it is recommended that, if sufficient sample remains, the analysis is repeated to increase the number of loci prior to any court proceedings. Although the level of adventitious matches is very low at 8 or 9 loci, upgrading matches to at least 10 loci (the number of loci that have historically been analysed in the UK from 1999 until 2014) is good practice.
 - i. It is not possible, ahead of comparison, to identify which UK crime scene profiles will result in 8-locus matches or more: it is inevitable that some 8-locus crime scene profiles from the UK will give matches with fewer than 8 corresponding loci with profiles from other MS. Such matches should be treated in the same way as 6- or 7-locus matches.
 - c. There is a chance that any match identified through a database search is adventitious for UK crime scene profiles with 6 or 7 loci, and for international matches with only 6 or 7 loci in common. The number of adventitious matches will depend on the size of the database searched. If the profiles matched share 8 loci fewer adventitious matches would be expected. There are therefore two approaches that could be taken:
 - i. For the UK not to share any crime scene profiles with fewer than 8 loci
 - ii. For the UK to share all crime scene profiles, and follow up potential matches only where these:
 1. have 8 or more matching loci (and of course no non-matching loci); or
 2. relate to the most serious crimes.
- If the first option is chosen, not to share any crime scene profiles with fewer than 8 loci, the risk is that real matches of interest to UK Policing will not be identified. In France and the Netherlands it has been found that [1,2]:
- 26-38% of 6-locus matches were true matches;
 - 82-94% of 7-locus matches were true matches.
- We can assume that approximately this range of true 6- and 7-locus matches would be seen in comparisons with UK profiles also.

Prum Feasibility Project HOME/2011/ISEC/AG/400002997

If the second option is chosen, the risk is one of perception: that the UK had in its possession the information necessary to identify an overseas offender, but did not follow up the lead. However, with this option, there is the potential to follow up leads in serious cases, should resources and priorities permit.

- d. Where any 6- or 7-locus matches are obtained and are of interest, it is recommended that reanalysis to increase the number of matching loci is *always* undertaken.
 - e. In the long term, it would be beneficial if database operators were furnished with software assistance in making decisions with regard to following up retrieved matches. It would be possible to design and implement software to provide the operator with a robust assessment of evidential weight in the form of a likelihood ratio. This measure of value could be combined with a prior probability, based on criminological factors – in particular the existing scale of cross-border crime. Coupled with a measure of utility based on the seriousness of the offence and policy considerations, this would provide an objective aid to decision making. Such software, once validated, would be useful to all MS participating in Prum exchange.
 - f. It is recommended that the UK shares its subject profiles, but routinely requires at least 10 matching loci prior to releasing demographic details to another country. We understand that only profiles from convicted offenders would be shared; this represents a very high percentage of the total number of subject profiles on the database. The analysis in this report is based on all subject profiles in the database; the level of adventitious matches expected for convicted offenders only would therefore be within a few percent of the totals presented herein. The subject profiles will be full results for the particular multiplex used in their analysis. So, with very few exceptions, these will have 6 (SGM), 10 (SGMPlus) or 16 (DNA17) fully designated loci. SGM profiles have insufficient loci to be included in a search. SGMPlus and DNA17 profiles are suitable for routine searching.
 - i. Any SGM profiles for subjects must be upgraded if a Prum search is required
 - ii. For exceptional cases, where a very serious crime is involved, consideration could be given to sharing demographic details where there are at least 8 matching loci.
2. **Expected number of true matches** that would be produced when the UK initially engage in Prüm DNA and search their crime scene stains (as a bulk exchange) to other Member States as is required by Prüm.

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

- a. The anticipated match rate in the bulk exchange is in the order of 14,000 true matches, with approximately 3000-4000 true matches annually thereafter.
 - b. Because France and Germany have the largest databases, these are the countries with which the majority of matches would be expected. However, patterns of cross-border crime may result in a different outcome.
 - c. The bulk searches do not have to be conducted simultaneously: the search against the database for each MS can be staged. The data provided in this report can be used to inform the order of searches, starting with a smaller MS database to test the protocol, gradually adding those with larger databases that would produce more matches, requiring more resources to follow up.
3. **Expected scale of adventitious matches** if the UK were to engage in Prüm (DNA) with each other MS.
- a. Figures 1 – 4 and Tables 3, 4, 5 and 7 illustrate the expected scale of adventitious matches during bulk exchange and subsequently.
 - b. Fewer adventitious matches will be expected for those with 8 loci than for those with 6 or 7 loci.
4. **Any recommended changes to match validation arrangements**
- a. It is recommended that all possible steps are taken to eliminate the potential that a match is due to contamination before it is reported. This will include checking all UK crime scene profiles against an effective elimination database prior to comparison with other MS, and as far as possible, checking any matching crime scene profiles from other MS against available elimination databases prior to reporting matches. Where any gaps exist in elimination databases, reports should be caveated to ensure that the possibility of contamination is considered.
 - b. It is therefore recommended that all matching profiles be searched against the UK elimination databases for manufacturers and unsourced profiles before any further action is taken on the match.

Introduction

The Peer Review Group defined the scope of and output from the project on 25/04/2014, as follows:

Scope

1. Developing a model to determine:
 - a. The likely impact of the composition of profiles being exchanged from UK to the other European Union (EU) countries, including the consideration of the exchange of incomplete crime scene profiles, the number of loci required for a valid match and the compatibility of the different data sets within the different EU member states (MS).
 - b. The likely DNA match rate(s) between the UK and other EU MS, depending on the composition of profiles being exchanged from UK to other EU member states.
2. Developing a model to evaluate the likely scale of adventitious matches if the UK were to engage in Prüm (DNA) with each EU Member State.
3. The work must also consider the partiality of profiles exchanged and the relative likelihood values of DNA matches with other EU MS (subject – subject, stain - stain, stain – subject and subject – stain) and in particular their value to UK law enforcement.
4. Advise the Home Office on other aspects of the Project as required.
5. Review UK procedures for validating matches

Outputs

5. Design of study (delivered)
6. Final Report September 2014 (the present document):
 - a. *Summary of findings*: single page list, including:
 - i. **Recommendations** to what composition of profiles should be exchanged between UK and other EU MS
 - ii. **Anticipated match rate** (e.g. the estimated scale of hits) that would be produced when the UK initially engage in Prüm DNA and search their crime scene stains (as a bulk exchange) to other Member States as is required by Prüm.
 - iii. **Expected scale of adventitious matches** if the UK were to engage in Prüm (DNA) with each other MS.
 - iv. **Any recommended changes to match validation arrangements**
 - b. *Main body*: Basis for recommendations, anticipated match rate and expected scale of adventitious matches
 - c. *Appendices*: Supporting information & data

Methods & Data

Data Collection

Questionnaires were designed and sent to Prüm contact points for each of the member states (MS) listed in Table 1. Responses were collated and are provided in full in Appendix 1.

Country	Abbreviation	Response Received
Austria	AT	Full
Cyprus	CY	Partial
Czech Republic	CZ	Full
Estonia	EE	Full
Finland	FI	Full
France	FR	Full
Germany	DE	Full
Hungary	HU	Full
Latvia	LV	None
Lithuania	LT	Full
Netherlands	NL	Full
Poland	PL	Full
Romania	RO	Full
Slovenia	SL	Partial
Spain	ES	
United Kingdom	UK	Full

Table 1: Countries to which requests for data were sent, and responses

Face to face discussions were held with National Database personnel from key MS to gather further detailed information on experiences to date and on processes in place:

1. Kees van der Beek, Custodian for National DNA Database, NL
2. Adam Shariff, DNA Technical Lead, UK National DNA Database (NDNAD)

Information from a French analysis of Prüm matches was obtained from Mathilde Huet, Ministry of the Interior, France [1].

Data Analysis

All assumptions and simplifications are collated in Appendix 2. The project brief was to estimate the “scale” of matches rather than precise numbers. Although we quote numbers (which are all rounded), these should be read as an approximate level (a “scale”), rather than precise numbers, since not all of the assumptions and simplifications can be tested in detail.

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

Evaluation of the Expected Scale of Adventitious Matches

The expectation for chance matches when databases are compared can be estimated using the formula:

$$\text{Expected adventitious matches} = nNPm$$

where n = the number of records in database 1

N = the number of records in database 2

Pm = the probability of a random match

The probability of a random match (match probability) for any number of DNA markers (loci) is calculated by multiplying together the match probabilities for the individual loci. This calculation makes an assumption that the loci are inherited independently from each other.

When profiles from crime scenes are analysed, not all loci will necessarily yield a result. This may be because the DNA is degraded, or because there is a mixture of DNA from two or more individuals, and not all loci are visible. When not all loci have yielded a result, a “partial” DNA profile is obtained.

To calculate Pm for partial profiles, the following method was used:

1. For each number of loci, a random selection from the loci in the multiplex was chosen (using the statistical programming software “R”);
2. This random selection of loci was repeated 100 times;
3. For each, the Pm was calculated.
4. The mean Pm for each number of loci was calculated as the mean of the Pm values for the 100 replicates.

The requirement for this work was to estimate the likely *scale* of adventitious matches rather than to provide an accurate point estimate. Therefore, any deviation from the assumption of independence between loci and the use of an average Pm rather than weighting the average to account for some loci being more likely to be missing from partial profiles than others, are unlikely to have a material impact.

Throughout the report, when we refer to an x -locus profile or an x -locus match, (where x can be between 6 and 16), each locus included is a fully designated locus, with no wild-cards. For example, a profile with 8 fully designated loci and one locus containing a wildcard (e.g. “R” for rare allele) would be counted as an 8-locus profile. If this profile were to match with one containing 7 overlapping and fully designated loci and a further locus in which a wildcard was assigned, the match would be a 7-locus match.

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prüm Feasibility Project HOME/2011/ISEC/AG/4000002997

Data from Cyprus and Slovenia were not included in the graphs and tables, since an accurate breakdown of partial profiles was not available. The c.1000 profiles from Cyprus analysed using Profiler Plus chemistry would not be suitable for comparison with UK data, as insufficient overlapping loci are present.

Evaluation of the Expected Scale of Adventitious Matches: Bulk Exchange

When a new MS begins Prüm comparisons, a “bulk exchange” is carried out of its entire Prüm database against the entire Prüm database of each other participating MS with which it is exchanging information.

It is in this bulk exchange that the largest number of adventitious matches will be encountered, as it is at this stage that the largest number of comparisons will be performed.

The UK data were compared against each MS for which data were available, as follows:

1. The number of total UK crime scene profiles for each number of loci in the database was decreased to 38% of the number provided, as only crime scene profiles that have not matched against a subject profiles are eligible for Prüm comparison. Currently, this represents 38% of UK crime scene profiles. We have made the simplifying assumption that profiles are equally likely to fulfil this criterion irrespective of the number of loci present.
2. Following equation 1, the comparisons in Table 2 were carried out, to estimate in each case, the number of adventitious matches.
3. Since results for both crime stain and subjects profiled using the DNA17 multiplex have only been accepted for loading onto the UK NDNAD since late July, it is assumed that the UK NDNAD profiles used for the bulk exchange will comprise SGMPPlus results with 6-10 loci.

	UK profiles compared	Profiles compared from each MS compared
1	Crime scene profiles with 6-10 loci	Entire MS database
2	Crime scene profiles with 6-10 loci	All MS crime scene profiles
3	All subject profiles	Entire MS database with 6-10+ loci

Table 2: Classes of estimate calculated

Evaluation of the Expected Scale of Adventitious Matches: ongoing exchange

After the bulk exchange has been carried out, the ongoing exchange of data will consist of:

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

1. Comparison of UK crime scene profiles not already matched against a subject profile against MS databases (Subjects and crime scene profiles). This will include two classes of UK crime scene profiles:
 - a. Historic, including those not matched during the bulk exchange or since
 - b. Recently added.
2. Comparison of *all* UK subject profiles against recently added MS profiles

We know the composition of the historic profiles, in terms of full and partial profiles, and have used this in the analysis (point 1a above). We cannot know accurately, however, what composition of partial profiles will be obtained in the future (for the analysis in point 1b). We have therefore made an assumption that a similar spread will be achieved as has been achieved historically.

For example, historically, c.78% of crime scene profiles in the UK database are full profiles: we have assumed that this will continue. However, because of the recent adoption of new multiplexes containing 16 rather than the previous 10 loci, a full profile for ongoing exchange will have 16 loci. Similarly, historically, c.6% of crime scene profiles have given 8/10 loci; for ongoing data exchange, under our assumption, this would equate to 6% of recent profiles being 13-locus partial profiles. This is likely to be a worst case scenario, since the new chemistries with 16 loci are substantially more sensitive than the old, 10-locus chemistry.

Evaluation of the expected scale of true matches

It is not possible to statistically evaluate the expected level of true matches, since this depends on criminological factors and not statistical factors. However, in order to provide an estimate of the likely order of magnitude of true matches, observations in countries which have actively been exchanging data over an extended period were studied.

Our analysis and previous work in the Netherlands [2] and France [1] are in close agreement that more adventitious matches occur with 6- and 7-locus matches. With 8 loci and above, c.98% or more of the matches observed will be true matches [1].

To estimate the number of true 6- and 7-locus matches, the ratio of true: false matches from the Netherlands [2] and France [1] were used to extrapolate an estimate of true matches from the expected levels of false matches calculated in this study.

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

For 8-locus matches and above, the numbers of expected adventitious matches were too small for any such extrapolation. Therefore, the matches observed in the Netherlands were used to extrapolate expectations for the UK:

1. International true matches as a proportion of the total number of international comparisons carried out; and
2. International true matches as a proportion of the number of reported National matches

However, it should be noted that an assumption of a similar pattern of cross-border crime would be required for this extrapolations to be valid. Criminology and patterns of cross-border crime fall outside the remit of this work, and the assumptions have not therefore been validated.

The bulk searches do not have to be conducted simultaneously: the search against the database for each MS can be staged. The data provided in this report can be used to inform the order of searches, starting with a smaller MS database to test the protocol, gradually adding those with larger databases that would produce more matches, requiring more resources to follow up.

Results

Evaluation of the Expected Scale of Adventitious Matches: Bulk Exchange

Tables 3, 4 & 5 shows the expected scale of adventitious matches as a result of bulk exchange between the UK and other MS in the categories listed in Table 2; these are shown graphically in Figures 1,2 & 3.

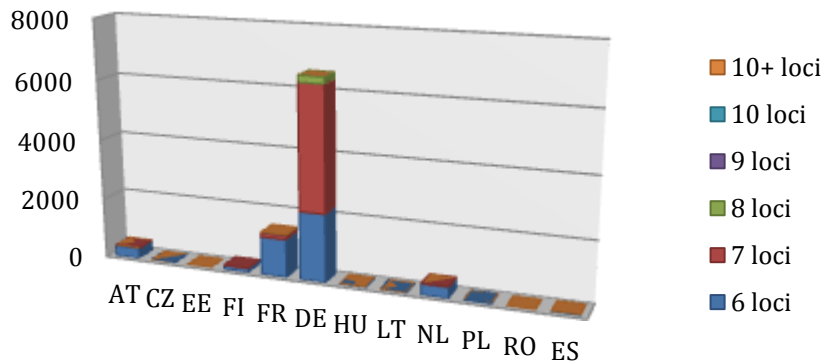
Results from Cyprus and Slovenia are not included in the tables, as a detailed breakdown of partial profiles was not available; any instances where expected results from Cyprus or Slovenia are non-zero are noted in the table legends.

Number of Loci	AT	CZ	EE	FI	FR	DE	HU	LT	NL	PL	RO	ES
6	340	10	0	130	1260	2240	20	10	360	10	0	23
7	20	0	0	10	130	4110	0	0	20	0	0	3
8	0	0	0	0	10	220	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
10+	0	0	0	0	10	0	0	0	0	0	0	1

Table 3: Comparison of all UK subject profiles against each MS database in a bulk exchange; rounded to nearest 10. If all of the Cypriot Powerplex 16 profiles were compared with the UK subject profile database, c.10 adventitious matches may be expected.

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Figure 1: Expected scale of Adventitious Matches between UK Subject Profiles & Full Databases of Member State Countries



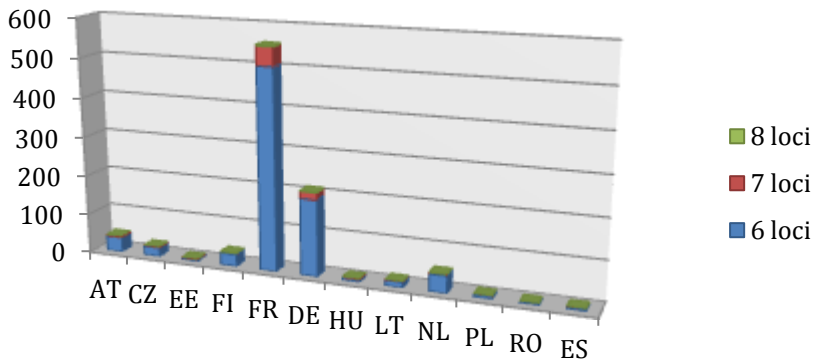
Number of loci	AT	CZ	EE	FI	FR	DE	HU	LT	NL	PL	RO	ES
6	38	23	5	30	504	192	7	14	43	7	4	6
7	4	2	0	3	47	18	1	1	4	1	0	0
8	0	0	0	0	3	1	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
10+	0	0	0	0	0	0	0	0	0	0	0	0

Table 4: Comparison of all UK crime scene profiles against each MS database in a bulk exchange; rounded to nearest integer



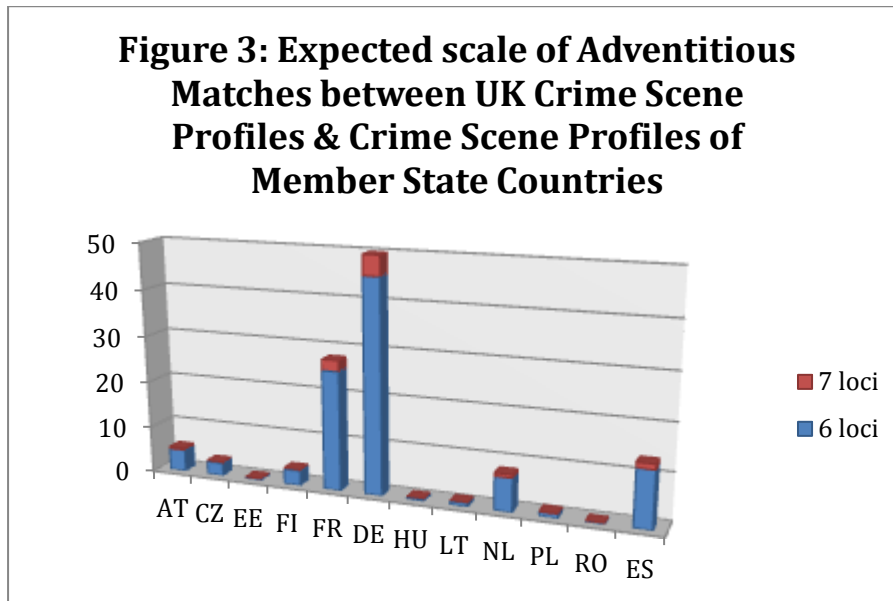
Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

Figure 2: Expected scale of Adventitious Matches between UK Crime Scene Profiles & Full Databases of Member State Countries



Number of loci	AT	CZ	EE	FI	FR	DE	HU	LT	NL	PL	RO	ES
6	5	3	0	3	25	45	0	1	7	1	0	12
7	0	0	0	0	2	4	0	0	1	0	0	1
8+	0	0	0	0	0	0	0	0	0	0	0	0

Table 5: Comparison of all UK crime scene profiles against each MS crime scene profiles in a bulk exchange; rounded to nearest integer



In order to calibrate the expectations and check for any deviations caused by our assumptions and simplifications, the method used to compare UK data against other MS data was applied to data from the Netherlands, France and Germany. Previous analyses [1] have evaluated the actual number of adventitious matches between these countries, thus enabling our expectations to be compared against reality. The results are shown in Table 6.

Countries Compared	Actual number of adventitious matches	Expected scale of adventitious matches using the methods in this report
FR crime stains vs DE database	211	259
FR crime stains vs NL database	51	57

Table 6: Expected versus observed adventitious matches

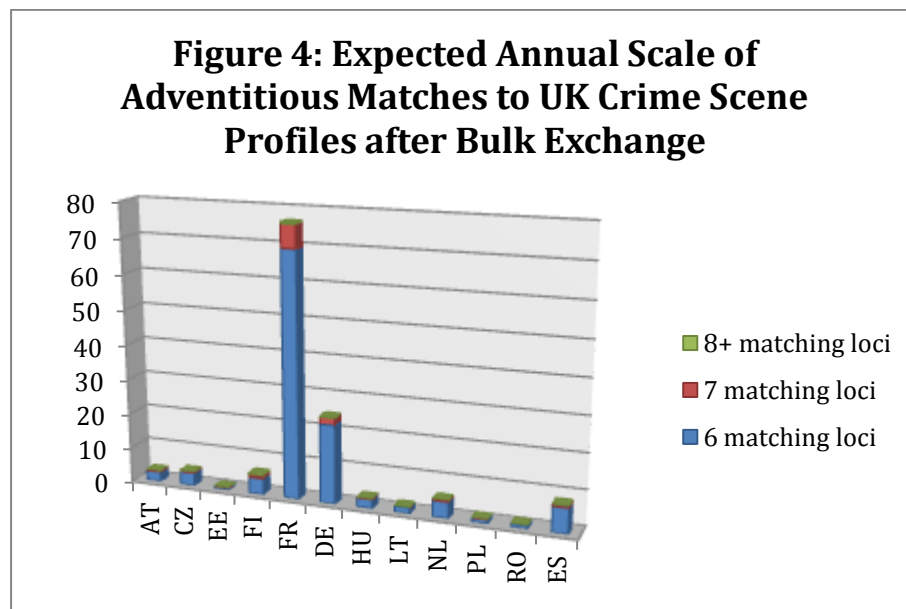
Evaluation of the Expected Scale of Adventitious Matches: ongoing exchange

Table 7 shows the expected annual scale of adventitious matches as a result of ongoing exchange between UK crime scene profiles and other MS; the data are shown graphically in Figure 4.

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

Number of loci	AT	CZ	EE	FI	FR	DE	HU	LT	NL	PL	RO	ES
6	3	4	1	5	69	22	3	2	5	1	1	7
7	0	0	0	0	6	2	0	0	0	0	0	1
8	0	0	0	1	0	0	0	0	0	0	0	0
9+	0	0	0	0	0	0	0	0	0	0	0	0

Table 7: Comparison of all UK subject profiles against each MS database on an annual basis; rounded to nearest integer



Evaluation of the expected scale of true matches

Estimates of the likely scale of true 6- and 7- locus matches, by extrapolation from French and Netherlands proportions of true: adventitious match proportions are shown in Figures 5 & 6 for bulk exchange, and in Figures 7 & 8 on an ongoing annual basis.

Figure 5: Estimate of the Scale of True 6-locus Matches to UK Crime Scene Profiles Expected during Bulk Exchange

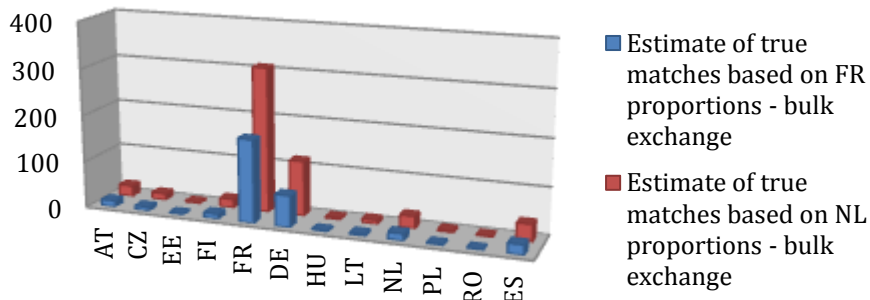


Figure 6: Estimate of the Scale of True 7-locus Matches to UK Crime Scene Profiles Expected during Bulk Exchange

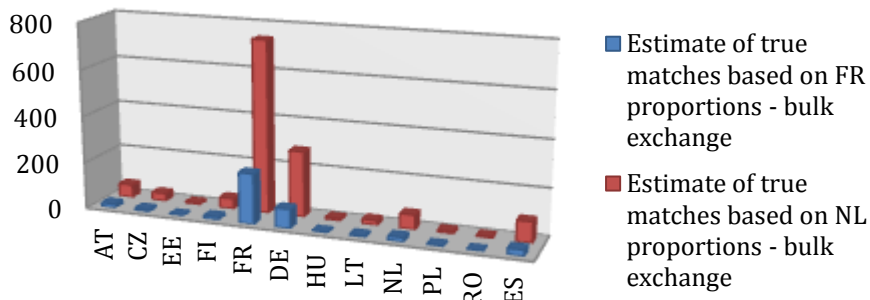


Figure 7: Estimate of the Scale of True 6-locus Matches to UK Crime Scene Profiles Expected Annually after Bulk Exchange

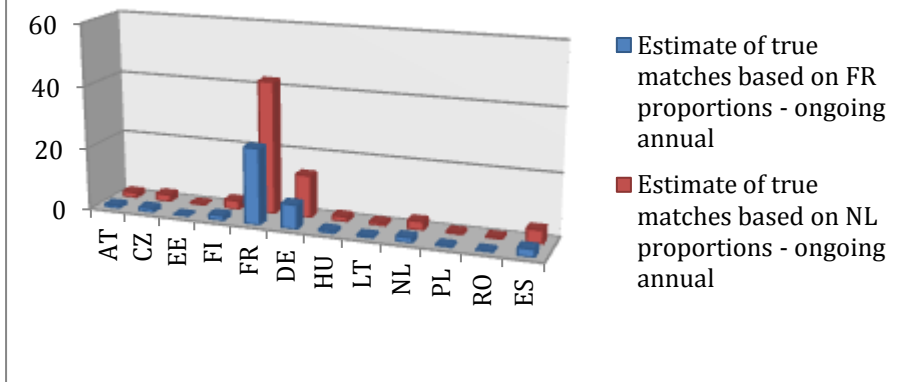
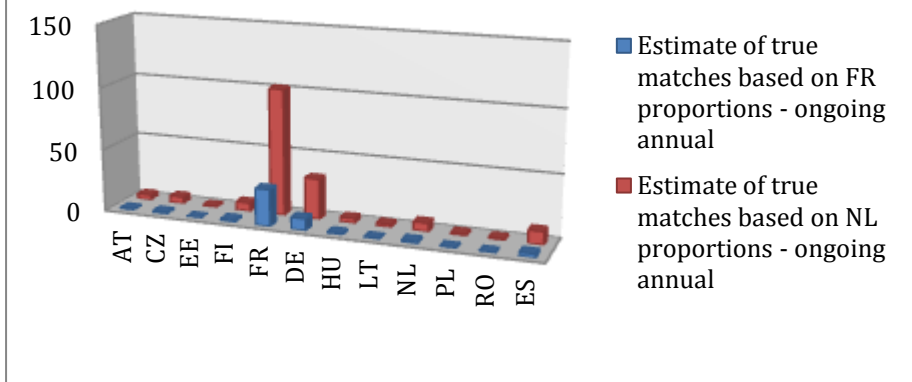


Figure 8: Estimate of the Scale of True 7-locus Matches to UK Crime Scene Profiles Expected Annually after Bulk Exchange



Using the data from the Netherlands on the proportion of comparisons yielding true matches, estimates of the scale of matches:

1. on bulk exchange; and
2. on an ongoing annual basis

by country are given in Figures 9 and 10 respectively.

Figure 9: Estimate of scale of true matches to UK crime scene samples in bulk exchange

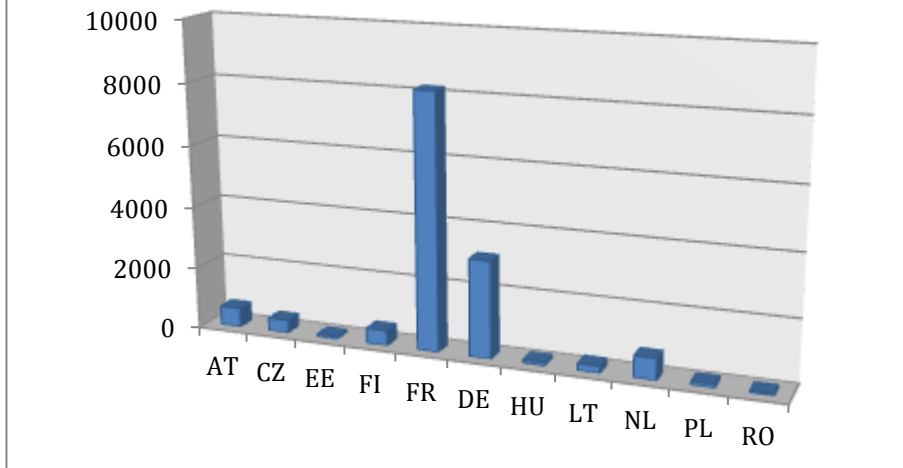
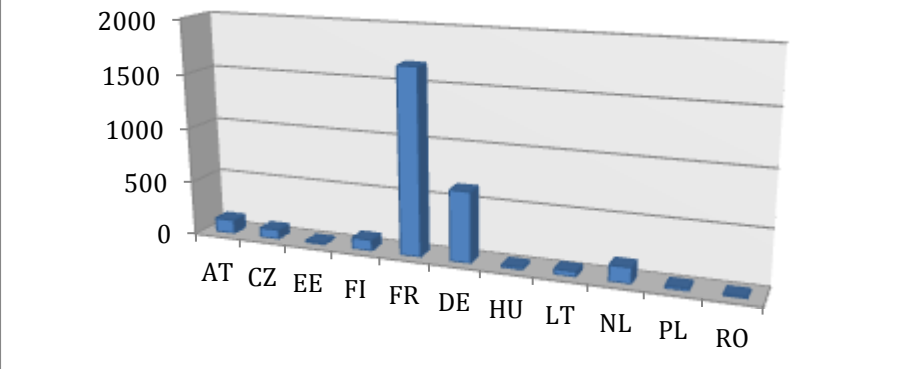


Figure 10: Estimate of scale of true matches to UK crime scene samples annually, after bulk exchange



A second estimate of the likely scale of true matches was provided by Kees van der Beek: in the Netherlands, for every 100 national matches seen, the international matches add a further 20.

For the UK, using data from the NDNAD Annual Report 2012/13, this would equate to approximately 4000 international true matches per year. This estimate is of the same order of magnitude as that shown in figure 10 (total from Fig 10 is approximately 3000).

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Discussion & Conclusions

The specification for the current work was to provide **recommendations** on what composition of profiles should be exchanged between UK and other EU MS, an **anticipated match rate** and **expected scale of adventitious matches** if the UK were to engage in Prüm (DNA) with each other MS and **any recommended changes to match validation arrangements**.

Basis of Recommendations on what composition of profiles should be exchanged between UK and other EU MS, and Expected Scale of Adventitious Matches

The work is based on data provided by other MS, and includes a number of assumptions and simplifications as detailed in Appendix 2. We have therefore, where possible, calibrated our results against those observed by MS which have been participating in Prüm data exchange for a number of years. The results presented in Table 3 show the outcome of this calibration, and give confidence that our estimates for the scale of adventitious matches are robust. Nonetheless, they should be seen as an approximation of the level (“scale”) of matches and not as precise numerical estimates.

Figure 1 & 2 demonstrate that the number of adventitious matches seen with 8 loci is much lower than for 6 and 7 loci, even where the number of comparisons performed is very large. It is clear, therefore, that the approach to 6- and 7-locus matches should be considered separately from the approach to 8-locus matches and above.

However, it is not possible, ahead of comparison, to identify which UK crime scene profiles will result in 8-locus matches or more: it is inevitable that some 8-locus crime scene profiles from the UK will give matches with fewer than 8 corresponding loci with profiles from other MS. Such matches should be treated in the same way as 6- or 7-locus matches.

Separate consideration will be given to UK crime scene profiles and UK subject profiles: it is likely that the matches to UK crime scene profiles will be of greater significance to UK law enforcement than matches to UK subject profiles; the latter will be of greater value to law enforcement agencies in other MS.

8-locus matches and above: UK crime scene profiles vs. MS databases

The number of adventitious 8-locus matches between UK crime scene profiles and the databases of other MS is expected to be very low (Table 3 and Figure 2). It is therefore recommended that crime profiles with 8 or more complete loci be compared against the databases of all other MS, to identify all 8+ locus matches.

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

Where matches of interest are obtained with 8 or 9 loci, it is recommended that, if sufficient sample remains, the analysis is repeated to increase the number of loci prior to any court proceedings. Although the level of adventitious matches is very low at 8 or 9 loci, upgrading matches to at least 10 loci (the number of loci that have historically been analysed in the UK from 1999 until 2014) is good practice.

6- and 7-locus matches: UK crime scene profiles vs. MS databases

There is a chance that any match identified through a database search is adventitious for UK crime scene profiles with 6 or 7 loci, and for international matches with only 6 or 7 loci in common. The number of adventitious matches will depend on the size of the database searched. If the profiles matched share 8 loci, fewer adventitious matches would be expected. There are therefore two approaches that could be taken:

1. For the UK not to share any crime scene profiles with fewer than 8 loci
2. For the UK to share all crime scene profiles, and follow up potential matches only where these:
 - a. have 8 or more matching loci (and of course no non-matching loci); or
 - b. relate to the most serious crimes.

If the first option is chosen, not to share any crime scene profiles with fewer than 8 loci, the risk is that real matches of interest to UK Policing will not be identified. In France and the Netherlands it has been found that [1,2]:

- 26-38% of 6-locus matches were true matches;
- 82-94% of 7-locus matches were true matches.

We can assume that approximately this range of true 6- and 7-locus matches would be seen in comparisons with UK profiles also.

If the second option is chosen, the risk is one of perception: that the UK had in its possession the information necessary to identify an overseas offender, but did not follow up the lead. However, with this option, there is the potential to follow up leads in serious cases, should resources and priorities permit.

Where any 6- or 7-locus matches are obtained and are of interest, it is recommended that reanalysis to increase the number of matching loci is *always* undertaken.

In the long term, it would be beneficial if database operators were furnished with software assistance in making decisions with regard to following up retrieved matches. It would be possible to design and implement software to provide the

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/400002997

operator with a robust assessment of evidential weight in the form of a likelihood ratio. This measure of value could be combined with a prior probability, based on criminological factors – in particular the existing scale of cross-border crime. Coupled with a measure of utility based on the seriousness of the offence and policy considerations, this would provide an objective aid to decision making. Such software, once validated, would be useful to all MS participating in Prum exchange.

UK Subject Profiles: Comparison versus other MS databases

Because the number of comparisons is greater than for crime scene profiles (there are more subject profiles to compare), the expected scale of adventitious matches to UK subject profiles is greater (Figure 1 & Table 3), with a small number of adventitious matches expected even with 10 loci.

Any matches obtained to UK subject profiles are likely to be of primary interest to the MS from which the relevant crime scene stain originated. It would be for this MS to conduct any follow-up analysis, and for the UK to set the standard for the number of loci required in a match before any demographic data from the UK subject would be released.

The bulk searches do not have to be conducted simultaneously: the search against the database for each MS can be staged. The data provided in this report can be used to inform the order of searches, starting with a smaller MS database to test the protocol, gradually adding those with larger databases that would produce more matches, requiring more resources to follow up.

It is recommended that the UK shares its subject profiles, but routinely requires at least 10 matching loci prior to releasing demographic details to another country. We understand that only profiles from convicted offenders would be shared; this represents a very high percentage of the total number of subject profiles on the database. The analysis in this report is based on all subject profiles in the database; the level of adventitious matches for convicted offenders only would therefore be expected to be approximately the same as the totals presented herein. The subject profiles will be full results for the particular multiplex used in their analysis. So, with very few exceptions, these will have 6 (SGM), 10 (SGMPlus) or 16 (DNA17) fully designated loci. SGM profiles have insufficient loci to be included in a search. SGMPlus and DNA17 profiles are suitable for routine searching.

- i. Any SGM profiles for subjects must be upgraded if a Prum search is required

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/400002997

- ii. For exceptional cases, where a very serious crime is involved, consideration could be given to sharing demographic details where there are at least 8 matching loci.

Basis of anticipated match rate that would be produced when the UK initially engage in Prüm DNA and search their crime scene stains (as a bulk exchange) to other Member States

Using data from other MS to estimate the scale of true matches to be expected relies on an assumption that cross-border patterns of crime are the same between the Netherlands (for which we have the greatest granularity of data), France (for 6- and 7-locus matches) and the UK are similar. We cannot substantiate this assumption, and so the estimates of true matches provided should be treated with caution.

The two different methods of estimating an approximate scale of true matches from Netherlands data (one based on a proportion of the total number of international comparisons and the other on a proportion of national matches) gave results that were of the same order of magnitude (c. 3000 vs c.4000 per annum after bulk exchange), which provides assurance that the methods used were valid. This provides, however, no information regarding the cross-border patterns of crime.

Match Validation Arrangements

The most important recommendation in relation to match validation arrangements is that the possibility of DNA contamination of a result, usually a crime stain, should always be considered, and as far as possible eliminated, prior to reporting a match and ideally before the profile is even included in the data exchange.

In the UK, plans are in progress to create and maintain a high quality suite of elimination databases, covering forensic service provider staff, police staff, medical examiners, staff from manufacturers of consumables and unsourced contaminants. As of September 2014, the Forensic Science Regulator has an agreed protocol for England and Wales in place, which will be implemented from April 2015 [3]. Although individual countries and FSPs hold elimination databases for their own scientific staff as well as manufacturers, there is not at this point a pan-European equivalent database. The DNA Working Group of the European Network of Forensic Science Institutes (ENFSI) is continuing to work towards shared manufacturers and unsourced contaminants databases. An unsourced contaminants database is held by the International Commission on Missing Persons (ICMP, Sarajevo) [4]. This includes DNA profiles that are detected in control samples that must be due to extraneous contaminating DNA.

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/400002997

Many are later sourced as being from manufactured consumables and solutions used in the process of recovering samples for DNA analysis.

It is recommended that all possible steps are taken to eliminate the potential that a match is due to contamination before it is reported. This will include checking all UK crime scene profiles against an effective elimination database prior to comparison with other MS, and as far as possible, checking any matching crime scene profiles from other MS against available elimination databases prior to reporting matches. Where any gaps exist in elimination databases, reports should be caveated to ensure that the possibility of contamination is considered.

It is therefore recommended that all matching profiles be searched against the UK elimination databases for manufacturers and unsourced profiles before any further action is taken on the match.

Acknowledgements

This report was prepared with the financial support of the Prevention of and Fight against Crime Programme of the European Commission (Directorate-General Home Affairs).

The authors wish to give particular thanks to Kees van der Beek (NFI), Adam Shariff (NDU), Nick Apps and Gary Linton (SCJS) for useful discussion of database and Prüm exchange issues, and to Mathilde Huet (MoI, France), Roberto Puch-Solis (LGC Forensics) and Ian Evett (PFS) for their insight and assistance in relation to the statistical analysis.

We thank the following individuals for provision of data on which this report is based: Reinhard Schmid (AT), Marios Cariolou (CY), Alice Reslova (CZ), Aivi Sootla (EE), Emilia Lindberg (FI), Alain M. Mesmoudi (FR), Alexander Bachmann (DE), Zoltan Kormos (HU), Jelena Kolesnikova (LT), Kees van der Beek (NL), Jakub Mondzelewski (PL), Florin Stanciu (RO), Katja Drobnic (SI), Caroline Goryll (NDU, UK)

Abbreviations (and definitions)

Adventitious match	DNA profiles from two individuals, who are not identical twins, that match by chance.
Allele	Alternative forms of a DNA sequence at a particular locus

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

DNA17	DNA multiplex that contains all the loci specified by ENFSI
ENFSI	The DNA Working Group of the European Network of Forensic Science Institutes
FSP	Forensic Science Provider
ICMP	International Commission on Missing Persons
Locus (pl.loci)	Specific location of a DNA sequence on a chromosome; for forensic analysis it refers to areas that vary between individuals
MS	Member State
Multiplex	DNA system that simultaneously analyses several loci in a single test
NDNAD	National DNA Database
NDU	National DNA Database Delivery Unit (UK)
SGMPlus	Second Generation Multiplex Plus (standard UK multiplex from 1999 – 2014)
Wild card	An undesignated placeholder included where the presence of an allele is uncertain but needs to be considered

References

- [1] Huet, Mathilde. A study of the false positives in the French DNA database and simulations, presentation at PIES conferences, NICC, Belgium, 25 June 2014
- [2] van der Beek , Kees. (Custodian Dutch DNA-database). The implementation of the Prüm Treaty/EU-Council Decisions in the Netherlands, presentation at Metropolitan Police Prum meeting, London on 16 May 2014:
- [3] Forensic Science Regulator Codes of Practice and Conduct: Protocol: DNA contamination detection -The management and use of staff elimination DNA databases SR-P-302 ISSUE 1
- [4] International Commission on Missing Persons (ICMP) Online Elimination Database Matching Application (<https://edb.icmp.org/index.php?w=intro&l=en>)

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Appendix 1: Data Returns from Member States

Data Request: Austria

Provided by: Reinhard.Schmid@bmi.gv.at

Data provided as of 24/07/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	25.320 (open stain profiles for Prüm searches)
	b. Suspect profiles	179.772
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	1.361 (stains) 0 (reference)
	7	989 (stains) 2 (reference)
	8	1.197 (stains) 3 (reference)
	9	1.578 (stains) 208 (reference)
	10	11.871 (stains) 125.473 (reference)
	11	423 (stains) 0 (reference)
	12	435 (stains) 0 (reference)
	13	516 (stains) 1 (reference)
	14	617 (stains) 6 (reference)
	15	798 (stains) 376 (reference)
	16	5.535 (stains) 53.703 (reference)
	>16	0
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	SGM (1997-1998) SGM+ (1999-2010) NGMSE (since 2011)	
Standard Practice regarding upgrading potential matches (processing additional loci)	Upgrade of each reference profile in case of a hit (national as well as in Prüm) to actual used quality	

Prüm Feasibility Project HOME/2011/ISEC/AG/4000002997

	(presently NGMSE). Upgrade of stains if necessary and if biological material is available.
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	No duplicated profiles since 2004 possible because of one times acquisition policy (controlled with fingerprint checks by 24/7 realtime data transmission and AFIS search procedures in Austrian .BK). After profile upgrade the better quality profile will be searched automated again also in Prüm network but with same profile number (only additional underline version number changes. This number refers to number of quality upgrade).
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	Only identical twins and multiple siblings will be enumerated and controlled. No statistics about other status of relatives are claimed.
Historical growth rate and projected growth rate of database	Each year about 13.000 new reference profiles and about 2000 new loaded open stains with Prüm quality and without national hits to national reference profiles (Prüm stain profiles)
Any “binning” and wildcards (including, if applicable, rare alleles) used	Of course. Provided in Prüm in accordance with existing quality definition and data structure of Prüm Decision

Data Request: Cyprus

Provided by: cariolou@cing.ac.cy

Data provided as of 31 December 2013		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	10.765
	b. Suspect profiles	335 (only convicted persons)
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	
	7	
	8	
	9	
	10	
	11	~ 1.000
	12	
	13	
	14	
	15	~ 9.765
16		
>16		
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database		ProfilerPlus (~1.000) PowerPlex-16 (~9.765)
Standard Practice regarding upgrading potential matches (processing additional loci)		Additional loci may be typed on reference profiles or if additional crime scene profiles are available. This is done on serious cases.
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database		No duplicates allowed in Cypriot Prüm database.
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)		No available estimates but we expect that this should be negligible.
Historical growth rate and projected growth rate of database		Difficult to estimate. For crime scene profiles perhaps 100-300 year. Much less for convicted persons.
Any "binning" and wildcards (including, if applicable, rare alleles) used		No binning nor wildcards included in database.

Data Request: CZECH REPUBLIC

Provided by: alice.reslova@pcr.cz

Data provided as of 21/08/2013		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	14 576
	b. Suspect profiles	2 404 (Suspects) 121 822 (Offenders)
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	31
	7	31
	8	116
	9	250
	10	5 207
	11	485
	12	331
	13	609
	14	1 305
	15	66 536
	16	58 545
>16	6 060	
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	PowerPlex 16, Identifiler (2002-2010) PowerPlex ESI 17, ESX 17, NGM (2010-2014) 50 981 = PowerPlex 16 15 556 = Identifiler 58 513 = ESI 17, ESX 17 2 634 = NGM 11822 = reanalysed profiles (mix of several kits)	
Standard Practice regarding upgrading potential matches (processing additional loci)	Additional loci are analysed by all profiles from potential matches (if profiles are still available). If profiles are not available, raw data of profile are checked and the calculation of match probability is	

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

	provided to police authority with the note about necessity of other verification of all relevant case information.
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	Estimation is about 100 profiles, but duplicated profiles are continuously deleted from the database.
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	Profiles of close relatives are not the part of Prüm comparison database, but in the rest of the whole database there are currently 420 profiles of close relatives.
Historical growth rate and projected growth rate of database	2002 – 2006 = 17304 profiles included to the database 2007 – 2009 = 47259 profiles included (mass collection of DNA profiles from prisoners) Since 2010 to this day the increment of profiles in the database is cca 20 000 profiles per year. The current number of all profiles in Czech DNA database is 158 892.
Any “binning” and wildcards (including, if applicable, rare alleles) used	We do not use any “binning” or wildcards. All alleles including microvariants are inserted into the database and if some allele is questionable we do not insert it at all.

Data Request: ESTONIA

Provided by: aivi.sootla@eeki.ee

Data provided as of 10/06/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	1712
	b. Suspect profiles	26 088
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	-
	7	-
	8	-
	9	4
	10	21263
	11	-
	12	2
	13	2
	14	-
	15	6357
	16	7
	>16	165
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	Most of the profiles with 10 loci – SGM Plus. Most of the profiles with 15 loci – PowerPlex ESI 16.	
Standard Practice regarding upgrading potential matches (processing additional loci)	If external profile has more loci and if possible - always process additional loci.	
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	Person profiles – few, if any. Stain profiles – some, exact number not known.	
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	No estimates	
Historical growth rate and projected growth rate of database	1618 new profiles in 2014 (until 10.06.2014)	
Any “binning” and wildcards (including, if applicable, rare alleles) used	Rare alleles – numerical value in database, if possible.	

Data Request: FINLAND

Provided by: Emilia Lindberg, bio.rtl.krp@poliisi.fi

Data provided as of 21/05/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	17 029
	b. Suspect profiles	145 828
Number of profiles with 6...n loci, where n = maximum loci (excluding Amelogenin)	6	517
	7	808
	8	990
	9	1 423
	10	129 725
	11	148
	12	209
	13	415
	14	1 100
	15	27 522
	16	-
>16	-	
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	AmpF [®] STR SGM Plus: between Dec 1999 and Jun 2012 Investigator ESS Plex Plus: since Jun 2012	
Standard Practice regarding upgrading potential matches (processing additional loci)	Person profiles can be upgraded with Investigator ESS Plex loci	
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	None	
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	N/A	
Historical growth rate and projected growth rate of database	~25 000/year	
Any "binning" and wildcards (including, if applicable, rare alleles) used	Over marker range alleles marked with < or >	

Data Request: FRANCE

Provided by: alain.mesmoudi@gendarmerie.interieur.gouv.fr

Data provided as of 06/02/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	137,140
	b. Suspect profiles	2,586,727
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	4,997
	7	7,916
	8	8,158
	9	7,744
	10	175,312
	11	7,125
	12	8,281
	13	15,308
	14	75,794
	15	2,082,655
	16	20,572
	>16	17: 309,884 18: 121
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	AmpFℓSTR® Sefiler™ AmpFℓSTR® COfiler™ AmpFℓSTR® Identifier® Investigator™ Idplex® Investigator™ Idplex Plus® AmpFℓSTR® NGM™ PowerPlex® PowerPlex® 16 PowerPlex® 18D PowerPlex® 21 PowerPlex® ES PowerPlex® ESI 16 PowerPlex® ESI 17 PowerPlex® ESX 16 PowerPlex® ESX 17 AmpFℓSTR® Profiler® AmpFℓSTR® Profiler Plus® AmpFℓSTR® SGM Plus®	
Standard Practice regarding upgrading potential matches (processing additional loci)	No upgrading	

Prüm Feasibility Project HOME/2011/ISEC/AG/4000002997

Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	We know that we have some but we don't know how many
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	We don't have this information
Historical growth rate and projected growth rate of database	375000 new profiles in 2013
Any "binning" and wildcards (including, if applicable, rare alleles) used	For some profiles we have only one allele known, it is represented as "-". For instance, if for some locus we have one value known, let's say 17, the locus is set as "17, -"

Data Request: GERMANY

Provided by: alexander.bachmann@bka.bund.de

Data provided as of 05/06/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	245,408
	b. Suspect profiles	791,598
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	8,906
	7	250,481
	8	302,189
	9	3,340
	10	10,774
	11	124,962
	12	1,360
	13	13,146
	14	3,666
	15	8,755
	16	308,840
>16	587	
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	There was never a regulation determining the kits in use. The data were generated using practically every kit on the forensic market including self-made. Most of the data are results of two independent amplifications preferably with two different kits. The number of false homozygotes is expected to be low. Analysis kits used are not linked to the profiles.	
Standard Practice regarding upgrading potential matches (processing additional loci)	No Standard Practice for upgrading potential matches in Germany.	

Prüm Feasibility Project HOME/2011/ISEC/AG/4000002997

Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	None. Every profile belongs to a separate criminal case.
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	There are no estimates how many close relatives are criminal offenders and at the same time in the database.
Historical growth rate and projected growth rate of database	Historical growth rate: 8,000-10,000 profiles a month Actual growth rate: 5,000-8,000 profiles a month Projected growth rate: 4,000-5,000 profiles a month
Any “binning” and wildcards (including, if applicable, rare alleles) used	No “binning”. Wildcards are only used for rare alleles (values below the “normal” range = “1” and values above the range = “99”).

Data Request: HUNGARY

Provided by: Zoltan Kormos - dna.database@orfk.police.hu

Data provided as of 12 June 2014		
Number of profiles in Prüm comparison database	a. Scene of Crime profiles	2.387
	b. Suspect profiles	21.072
	c. Convicted Offender profiles	14.275
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	61
	7	92
	8	129
	9	592
	10	565
	11	62
	12	90
	13	124
	14	170
	15	35731
	16	112
	>16	6
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database		See Table 1.
Standard Practice regarding upgrading potential matches (processing additional loci)		Every match candidate originating from automated DNA data exchange among Prüm partners will be checked by qualified experts of the Hungarian National DNA Database. The validation process is carried out according to the ENFSI DNA database management recommendations, which in many cases contains additional DNA analysis before notification is made.
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database		0

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	No data
Historical growth rate and projected growth rate of database	See Table 2.
Any “binning” and wildcards (including, if applicable, rare alleles) used	We use multiplex kit specific > and < bins for out of ladder range alleles. No wildcards are currently in use.

Table 1.

Multiplex Kit	Number of Personal DNA Profiles	Number of Scene of Crime DNA Profiles *	Period in use
Profiler Plus	0	505	1998 - 2010
COfiler	0	34	1999 - 2006
Identifiler	12.533	17	2001 - 2014
MiniFiler	0	32	2007 - 2012
SGM Plus	0	757	2009 - 2012
NGM	9.007	800	2009 - 2014
NGM SElect	0	130	2012 - 2014
PowerPlex 16	0	7	2000 - 2008
PowerPlex ESI 16	13.807	165	2011 - 2014
PowerPlex ESI 17	0	34	2009 - 2014

* As some Scene of Crime DNA profiles managed in Prum database has been produced by the use of multiple kits, the sum of given profiles is more than the actual number of profiles they were generated from (2.387).

Table 2.

Date	30.06. 2012	30.09. 2012	31.12. 2012	31.03. 2013	30.06. 2013	30.09. 2013	31.12. 2013	31.03. 2014	12.06. 2014
Number of Scene of Crime DNA Profiles	115	358	478	482	480	473	1.734	2.142	2.387
Number of Personal DNA Profiles	18.291	14.441	13.864	21.588	23.681	27.096	30.242	32.849	35.347

First Prum search date: 19.09.2012 with AT.

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Data Request: LITHUANIA

Provided by: Jelena.Kolesnikova@policija.lt

Data provided as of 14/04/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	4080
	b. Suspect profiles	70541
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	34
	7	144
	8	364
	9	4510
	10	45399
	11	50
	12	116
	13	503
	14	352
	15	22870
	16	36
>16	243	
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	AmplF1STR SGM Plus (not in use since Dec 2011) 45399 profiles, AmplF1STR Identifier 1056 profiles, AmplF1STR Minifiler 3 profiles, AmplF1STR NGM 28127 profiles, AmplF1STR NGM Select Express 36 profiles.	
Standard Practice regarding upgrading potential matches (processing additional loci)	If possible, re-amplification with NGM kit (DNA extracts are stored up to 10 years)	
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	1433 duplicated profiles of suspect's	
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	N/A	
Historical growth rate and projected growth rate of	Approx. 700 stain's	

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

database	profiles and 8000 suspect's profiles per year are included into database.
Any "binning" and wildcards (including, if applicable, rare alleles) used	N/A

Data Request: NETHERLANDS

Provided by: k.v.d.beek@nfi.minvenj.nl

Data provided as of 10/07/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	38,678
	b. Suspect profiles	191,338
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	1,418
	7	934
	8	1,671
	9	4,214
	10	125,972
	11	519
	12	459
	13	1,205
	14	6,778
	15	85,919
	16	223
	>16	1,553
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database		NGM since May 2013
Standard Practice regarding upgrading potential matches (processing additional loci)		Each 6 or 7 locus match which is of interest to NL is typed with additional loci
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database		450
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)		No data
Historical growth rate and projected growth rate of database		Current total is c. 200000. Expected to include 25000 persons per annum from 2014 onwards, and from 2022, will start to remove profiles, leading to a steady state total number of profiles of c.625000 persons in total by 2034.
Any "binning" and wildcards (including, if applicable, rare alleles) used		None

Data Request: POLAND

Provided by: jakub.mondzelewski@policja.gov.pl

Data provided as of 30/04/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	4791
	b. Suspect profiles	33890
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	23
	7	69
	8	122
	9	240
	10	29584
	11	60
	12	37
	13	69
	14	168
	15	7941
	16	282
	>16	86
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	SGMplus – from 2007 to the end of 2012. NGM or NGMSelect since 2013	
Standard Practice regarding upgrading potential matches (processing additional loci)	If possible, profiles are upgrading to 15 or 16 loci. Kit: NGM or NGMSelect	
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	0	
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	32	
Historical growth rate and projected growth rate of database	c.a. 6000 profiles per year	
Any “binning” and wildcards (including, if applicable, rare alleles) used	No	

Data Request: ROMANIA

Provided by: Florin Stanciu, criminalistica@politiaromana.ro

Data provided as of 17/07/14			
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	801	
	b. Suspect profiles	suspects: 702 convicted offenders: 20,916	
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	7 crime	7
	7	16	16
	8	72	72
	9	160	160
	10		352
	11		380
	12		543
	13		1008
	14		2278
	15		16150
	16		1479
	>16		12
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database	ESSplex - 5974 ESSplex SE - 2413 Nonaplex - 32 Identifiler - 14028		
Standard Practice regarding upgrading potential matches (processing additional loci)	If we have a copy of the original profile, standard procedure implies reprocessing the profile		
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database	22		
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)	32		
Historical growth rate and projected growth rate of database	c. 5000 per year		
Any "binning" and wildcards (including, if applicable, rare alleles) used	-		

Data Request: SLOVENIA

Provided by: katja.drobnic@policija.si

Data provided as of 01/12/2013		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	6,356
	b. Suspect profiles	27,534
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	
	7	
	8	
	9	
	10	33,890
	11	
	12	
	13	
	14	
	15	
	>16	
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database		SGMplus until 2011, then NGM
Standard Practice regarding upgrading potential matches (processing additional loci)		
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database		
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)		
Historical growth rate and projected growth rate of database		
Any "binning" and wildcards (including, if applicable, rare alleles) used		

Data Request: UK

Provided by: Caroline.Goryll@homeoffice.pnn.police.uk

Data provided as of 01/07/2014		
Number of profiles in Prüm comparison database:	a. Scene of Crime profiles	170,175
	b. Suspect profiles	5,599,335
Number of profiles with 6....n loci, where n = maximum loci (excluding Amelogenin)	6	4,122
	7	5,910
	8	9,869
	9	17,282
	10	132,426
	11	
	12	
	13	
	14	
	15	
	16	
	>16	
Multiplex kit(s) used, with dates in use and associated number of profiles in Prüm comparison database		SGMPlus
Standard Practice regarding upgrading potential matches (processing additional loci)		
Estimated number of duplicated profiles (if any) profiles in Prüm comparison database		
Any available estimates of numbers of close relatives on the databases (siblings and parent/child)		Unable to provide
Historical growth rate and projected growth rate of database		<p>Previous years from annual reports 13/14 figures:- Crime Scene loads - 35005 Subject loads - 361933 Crime scene deletions - 6837 Subject deletions – 1,384,905</p> <p>Previous years from annual reports 13/14 figures:- Crime Scene loads - 35005</p>

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

	<p>Subject loads - 361933 Crime scene deletions - 6837 Subject deletions – 1,384,905 Previous years from annual reports 13/14 figures:- Crime Scene loads - 35005 Subject loads - 361933 Crime scene deletions - 6837 Subject deletions – 1,384,905</p>
<p>Any “binning” and wildcards (including, if applicable, rare alleles) used</p>	<p>From 1999 to 14th November 2008 the pre-3.3b rules were in place for Th01 binned alleles such that a Th01 10 would be assigned a value ‘R’ for loading to the NDNAD – full details of the legacy arrangements are in 2008 version of the Technical Standards document. (Both the legacy document and current document included below. Since this point FSPs have been back-converting ‘R’ to numerical designations for NDNAD retained records. A further minor change has been introduced since 1st February 2014 where vWA alleles 22, 23, 24 and 25, (and any variants of) though callable by the SGMPlus had until this point needed to be assigned a wildcard ‘R’ (to account for potential SGM vWA/FGA crossover) – from 1st February 2014 these alleles are to be assigned with the numerical value.</p>

Appendix 2: Assumptions & Simplifications

1. The numbers quoted are all based on the match probabilities for the White Caucasian population since it is assumed that this is the largest population group across Europe
2. The match probabilities are taken from US White Caucasian data for DNA17 systems and it is assumed that these are appropriate for the European White Caucasian population
3. The numbers are based on the average probability of a match
4. The calculations assume independent inheritance of DNA loci
5. We have adjusted the number of UK crime scene profiles to “remove” all those already matched to a subject profile. In doing so, we have assumed that the proportion of partial profiles in the remaining set mirrors that in the crime scene database as a whole
6. For the purpose of estimating matches in future when the UK uses DNA17 systems, it is assumed that the proportion of crime results that are partial profiles will remain constant. That is, it is assumed that the proportion of the 16 DNA17 loci (not including Amelogenin) obtained is the same as that of the 10 SGMPlus loci. So an SGMPlus crime result with 8 loci would be the equivalent of a DNA17 crime result with 13 loci. About 6% of current SGMPlus crime profiles have 8 loci so 6% of future loads of DNA17 crime profiles will have 13 loci
7. Each number of loci quoted refers to fully designated loci and not loci containing wildcards
8. We assume that relatives and duplicates present within databases are at such a low level as to have negligible impact on the analysis.
9. The eligible unmatched crime results from the UK have been compared against the crime and subject profiles of other MS with no breakdown of crime types
10. It is assumed that all partial profiles are from crime stains and not subjects for those MS where this information could not be provided. The actual figures are given for Austria and Germany
11. In estimating the scale of true matches, an assumption that the UK pattern of cross-border crime emulates that of the NL is required

Appendix 3: Supporting information and data

Prüm Inclusion Rules

The criteria that DNA results have to reach to be included in international comparisons are that the profile:

- Must include at least 6 of the 7 old ESS loci for subjects
- Must include at least 6 ESS loci for crime scene stains
- Must include any other of the 24 old1 Interpol loci
- One allele of a locus can be a wildcard
- No mixed profiles (a maximum of two values per locus) are allowed
- No profiles that have already matched a person are allowed
- No profiles that a country does not want to make available are allowed (e.g., DNA profiles of laboratory personnel kept for contamination detection purpose)

Prüm Matching Rules

The software produces a match when there are at least six fully matching loci between two DNA profiles. In addition, one deviation (wildcard or mismatch) is allowed, and this is called a near match. Any type of profile sent for a comparison will be compared to any type of DNA profile available for comparison, so the following types of matches can occur: stain-stain, stain- person or person-person. The matches can be of four different qualities:

Quality 1: All alleles of all loci that can match are identical

Quality 2: One of the two matching profiles contains a wildcard

Quality 3: One of the alleles of one locus contains a mismatch of one base pair (e.g., 9.2 ↔9.3)

Quality 4: One of the alleles of one locus contains a mismatch of more than one base pair (e.g., 22 ↔26)

Reports on visits

- **NDNAD Delivery Unit (NDU)**

Sue Pope met with Adam Shariff (DNA Technical Lead) and Caroline Goryll (data analyst) from the Home Office on 29 May 2014. They discussed the provision of data by NDU. The request was put to the National DNA Database Strategy Board at their meeting in June and accepted. The meeting also covered the current and planned rules on designation of wild cards, rare alleles and somatic mutations as well as policy on duplicate subject and crime profiles.

- **Netherlands Forensic Institute (NFI)**

Gill Tully and Sue Pope visited Kees van der Beek (Netherlands National DNA Database Manager) at the NFI on 12 June 2014. This included the opportunity to

With the financial support of the Prevention of and Fight against Crime Programme
European Commission – Directorate-General Home Affairs

Prum Feasibility Project HOME/2011/ISEC/AG/4000002997

watch and discuss the daily review, process and actions for the Prum database search hits involving the Netherlands.

We were also provided with data about the types and numbers of hits including a review of 100 Quality 3 & 4 cases that were followed up with further analysis. This real data was used to assess the reliability of the estimates produced using the adventitious match rate model.

Other data provided was the breakdown of matches in different categories, including the proportion of matches to foreign stains that were from NL residents born abroad.

- **National Crime Agency (NCA)**

Gill Tully and Sue Pope visited the NCA at Warrington on 18 June 2014, meeting representatives of the Interpol and SIRENE bureaus and the UK-Prum DNA and fingerprint Project. The process of validating European Arrest Warrants was discussed. The approach to scene to scene DNA matches uses a post search sift rather than pre-search limitations. Issues arise with partial matches to SGM reference samples that can no longer be upgraded since the S and Marper ruling has led to destruction of the stored samples.