Extracts from a submission from the Home Office Economics and Resource Analysis Group to the Parliamentary Under-Secretary of State for Crime Prevention entitled, 'DNA retention – Analysis of arrest-to-conviction data', July 2010.

You commissioned the Economics and Resource Analysis Group (ERAG) to undertake analysis of arrest-to-conviction data from the Police National Computer (PNC), to support your thinking on the appropriate retention regime for DNA profiles. This submission presents the initial and provisional results of our analysis, and identifies possible extensions to the work. The Annex provides more technical details.

Summary

ERA analysis of arrest-to-conviction data obtained from the PNC indicates that the time taken for the risk of conviction of individuals with no previous convictions who are arrested but not sanctioned to fall to the level observed in the general population is approximately three years. This assumes that DNA profiles are retained on arrest for all offences. If the scope of the provisions is restricted to 'Scottish list' offences with retention on arrest, the time taken for conviction risk to fall to the population level is 3¾ years. If scope is restricted to 'Scottish list' offences with retention only on charge, the time taken is 4¾ years (although this result is subject to significant uncertainty due to the small sample size).

We have also made provisional estimates of the outcomes of profile retention under each regime, assuming retention for three years. These suggest that restricting the scope of the retention provisions (from all offences on arrest to 'Scottish list' on charge) significantly reduces the number of 'innocent' profiles retained. The proportion of those profiles expected to receive a sanction before the end of the retention period does not change as scope is restricted, while the proportion of sanctions that are received which are serious (as defined by the 'Scottish list') rises slightly in absolute terms. Thus, the Scottish model appears to be effective in offering protection to individuals who are arrested but not sanctioned for any offence. The extent to which it is effective in ensuring that only the profiles of more 'serious' potential offenders are retained is debatable.

Consideration

Arrest-to-conviction analysis

We (ERAG) have produced provisional estimates of the time taken for the risk of conviction for an individual with no prior convictions who is arrested but does not receive a warning, caution, fixed penalty notice or conviction (from now on termed 'with no sanction' or WNS) to fall to the same level as in the general population with the same demographic profile. We have also estimated upper and lower bounds for this duration, based on a 95 per cent confidence interval. We have produced these estimates for a number of possible retention scenarios, as summarised in Table 1.

Table 1 Time to equalise conviction risks under different retention scenarios

Retention scenario	Intersection	Lower bound	Upper bound
All offences, retention on arrest	3 years	2 ¾ years	3 ¼ years
Scottish list offences only, retention on arrest	3 ¾ years	3 years	4 ¼ years
Scottish list offences only, retention on charge only	4 ¾ years	3 years	9 years

Table 1 indicates that it takes approximately three years for the risk of conviction of an individual with no prior convictions who is arrested WNS to fall to the level of conviction risk observed in the general population, with upper and lower bounds of 3¼ and 2¾ years respectively. This assumes that DNA is (temporarily) retained on arrest for all offences.

When attention is limited only to those who are initially arrested WNS for offences on the 'Scottish list', the time taken for the subsequent conviction risk to fall to the general population level is 3¾ years, with upper and lower bounds of 4¼ years and three years respectively.

Finally, when (temporary) retention occurs only when an individual is arrested *and charged* for a 'Scottish list' offence (the scenario which represents the actual Scottish retention model), the time taken for conviction risk to fall to that observed in the general population rises to 4¾ years, with upper and lower bounds of nine years and three years respectively.

There are a number of possible reasons why conviction risk takes longer to fall to the population level as the retention regime becomes more limited in scope. The explanation that the more restricted regimes might simply be identifying individuals with higher risks of conviction is possible, but not supported by further analysis on the outcomes of profile retention. Rather, it seems that more restricted regimes 'recapture' individuals much more slowly, because only a proportion of any potential contact with the criminal justice system is relevant to the temporary retention policy. However, we emphasise that these are only preliminary results and explanations, and further exploration of the reasons for the observed variations is necessary.

The principal explanation for the significant increase in the size of the estimated error bounds when we go from the first, most general scenario to the third, most restrictive, is the size of the associated dataset. The initial estimates are based on a sample size of 64,000 individuals who were arrested between April and July 2006. However, the number of those individuals who were initially arrested for a 'Scottish list' offence and then charged but not convicted was only 3,300, and this much reduced sample size makes the results subject to much greater uncertainty, and hence generates much wider confidence intervals.

Although our central estimates are subject to uncertainty, they are towards the lower end of the possible range within which the respective true values reside. This is because our estimates of the population conviction risk are, on balance, likely to be too high, because they are based on data which include individuals with convictions, and do not reflect the fact that conviction risk tends to fall with age. A lower population conviction risk would imply intersection points which are further into the future than the estimates presented in Table 1. In fact, it is possible that, if previous convictions and

aging could be built into the general population risk estimate perfectly, the conviction risk of those arrested might never fall to the level of those who were never arrested, or only after an extended period. This is a result that has been suggested by at least one academic study.

Implications of the arrest-to-convictions results for setting the retention period

If a policy of retaining the DNA profiles of individuals who are arrested WNS, but not the profiles of the general population, is to be justified on cost-benefit grounds, we need to be able to demonstrate three things:

- There are benefits to be gained from retaining DNA profiles, in terms of the likely impact on detection rates and, ultimately, on crime, or some other form of social value – otherwise, no retention is justified;

- The benefits of retaining the profiles of individuals who are arrested with WNS are higher than of retaining those of the general population – otherwise, there is no case for singling out the arrested WNS group specifically;

- The benefits of retaining DNA profiles outweigh the costs, in terms of, for example, database maintenance but also factors such as individual privacy.

Therefore, the time taken for the conviction risk of individuals arrested WNS to fall to the level observed in the general population only gives an indication of the maximum retention period which might be justified. The evidence currently does not exist in a form which would allow us to estimate the marginal value of retaining the DNA profiles of different individuals, in terms of the impact on crime. There is also no evidence of the cost of retention in terms of its impact on individual privacy. Both factors would imply shorter retention periods than the durations presented in Table 1, however. This means that the Table 1 results are potentially consistent with a three-year retention period for 'Scottish list' offences, whether retention is at arrest or charge, but are likely to suggest a retention period for all offences on arrest which is shorter than three years.

Profile outcomes analysis

We have made provisional estimates of the outcomes of profile retention under each retention scenario, assuming a three-year retention period. These results are presented in Table 2.

Table 2 Profile retention outcomes after three years under different scenarios				
Retention scenario		Innocent profiles	Retain on proven	Proportion Scottish list
All offences, retention on arrest		450,000	20.5%	5.6%
Scottish list offences only, retention on arrest		190,000	19.8%	7.0%
Scottish list offences only, retention on charge only		27,000	19.2%	6.9%

Table 2 indicates that, under a regime of temporary retention, for three years, on arrest for any offence, we would expect 450,000 profiles of innocent

individuals to be being retained at any one time once the policy had matured. Just over 20 per cent of these would go on to be retained indefinitely because the individual received a conviction before the end of the retention period. Of these, just over a quarter (5.6 per cent of the original 450,000) would be retained indefinitely on conviction for a 'Scottish list' offence.

It can be seen from Table 2 that, as the scope of the retention regime is restricted, the number of innocent profiles retained falls by around 60 per cent because of the reduced scope of offences, but then by a further 85 per cent (95 per cent in total) due to the restriction to retention on charge. The proportions of these profiles which are predicted to be retained indefinitely because of a sanction before the end of the three-year retention period is very similar for each regime, however, at around 20 per cent. Only a minority (between one quarter and one third) of these sanctions are predicted to be for a 'Scottish list' offence, although the higher proportions for the restricted regimes are statistically significant compared with the more general case.

The results of this outcomes analysis suggest that restricting the retention regime to the 'Scottish list' of offences appears to be effective in offering protection to individuals who are arrested but not sanctioned for any offence. However, it does not appear to result in the retention of profiles of individuals who have a higher risk of sanction, although any sanctions which do occur are slightly more likely to be serious.

Arrest-to-arrest analysis

Finally, we have also undertaken an analysis based solely on arrest-to-arrest data, as a check on the results of our previous analysis of this issue. Our best estimate of the time it takes for the risk of arrest to fall to the level of arrest observed in the general population is now between four and four-and-a-half years, compared with our previous estimate of six years. The main reasons for this reduction are as follows:

- Using a four-month period to select arrestees into the dataset (compared with our original one month) materially increases the sample size and the precision of the results;

- The limited follow-up period that is possible means that the future profile of arrest risk needs to be forecast, and this introduces possible error. The longer follow-up period (45 months instead of 39 months) which was possible for the current analysis reduces the need to forecast the profile of future risk, thereby reducing this source of error;

- A more robust check on existing criminal history was possible when compiling the dataset for this analysis, meaning that over 10 per cent of the individuals in the original (smaller) sample were found to have a prior conviction, and could be excluded;

- The previous analysis required us to estimate arrest risk for the general population, and this was only possible if a large number of assumption were made. This new analysis allows us to estimate what level of error these assumptions might have introduced.

These considerations lead us to conclude that our previous estimate of six years for the time taken for re-arrest risk to fall to level of arrest risk observed

in the general population, although the best available estimate at the time, was too high.

Full results: All offences, retention at arrest

Hazard rate analysis

These graphs show what the level of proven offending would be at the end of different lengths of DNA retention period. Proven offences include cautions, warnings, reprimands and fixed penalty notices as well as convictions. The national comparator shows the national level of proven offending for the same age and gender demographic and helps to put the proven offending level into context. The intersection point was used to justify the six year retention period in the Crime and Security Act.

The uncertainty caused by pending cases is reflected by having two hazard curves, one assuming all pending cases will be proven and another assuming none of them will be proven. A 95% confidence interval was calculated for each of these using a statistical technique known as boot-strapping, these confidence intervals overlap and only the two extremes are shown.
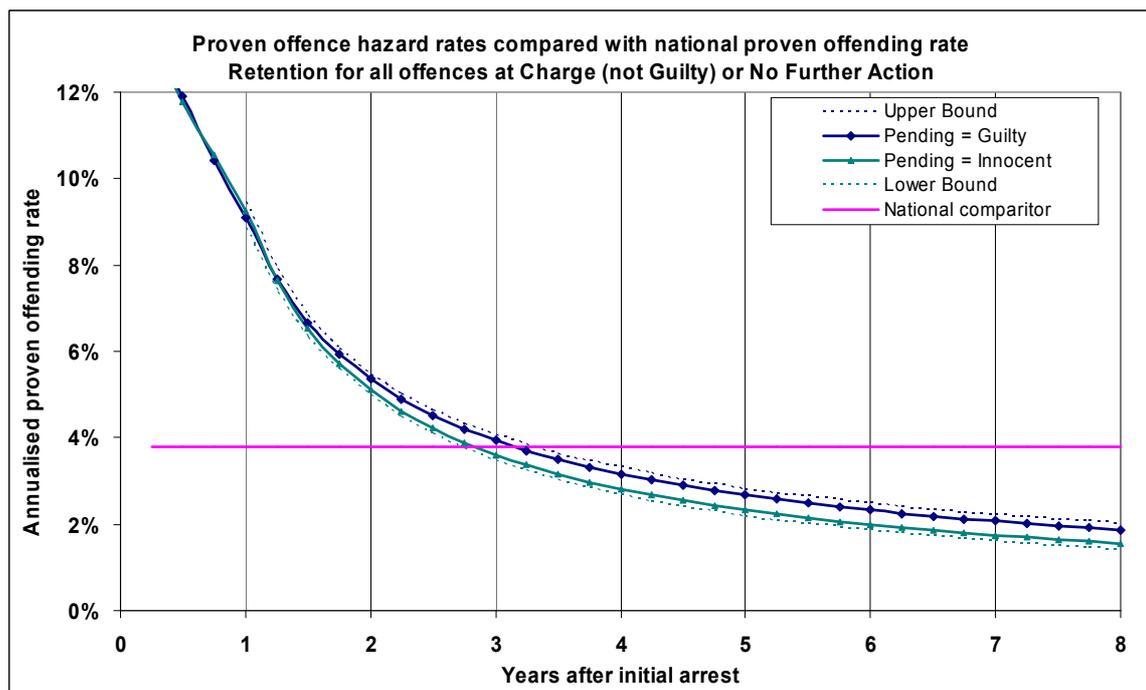
Following the initial arrest, the offending rate is equivalent to 13.2% of the sample committing one or more proven offence per year. The rate falls off sharply, halving in 18 months and dropping to 1.7% by year eight.

The national comparator, for the same age and gender demographic, is 3.8%. It takes around three years for the offending rate to drop to this level, with a lower bound of two years nine months and upper bound of three years three months.

Outcome analysis

The table below provides estimates of the number of innocent DNA profiles that would accumulate over time on the DNA database, and the status of profiles at the end of their retention period. The number of innocent profiles is not proportional to the retention length. This is partly because longer retention periods lead to more situations where an innocent is not added to the database because they are already on there and more instances of an innocent profile being retained indefinitely following a first proven offence.

Higher retain on proven rates, for a given retention period, suggests a policy is more affective in targeting individuals who are likely to receive a sanction for a proven offence or a sanction for a proven Scottish List offence. The similar rates, shown below and in the subsequent sections, suggest that the more restrictive retention policies are not materially better at targeting. This has been identified as a potential area for further analysis.

Retention period	Innocent Profiles on Database	Outcomes at the end of the retention period		
		Retain on proven	Retain on Reset	Dispose
1 year	178,000	11% (2.6% Scottish List)	8%	81%
2 years	322,000	16% (4.3% Scottish List)	11%	72%
3 years	448,000	21% (5.6% Scottish List)	13%	67%
4 years	564,000	23% (6.6% Scottish List)	14%	63%
5 years	671,000	26% (7.5% Scottish List)	14%	60%
6 years	773,000	28% (8.3% Scottish List)	15%	58%

There are around 178,000 arrests each year that could result in a new innocent profile being added to the DNA database under this policy. Around 11% of these will receive a proven offence of any type within one year, 21% within three years and 28% within six years. For Scottish List offences, these rates are 2.6%%, 5.6% and 8.3% respectively.

A longer retention policy will result in more retention periods being reset, following a second eligible arrest, and fewer profiles being deleted at the end of the original retention period. The proportion reset is 8% at one year, 13% at three years and 15% at six years. Deletions are 81% at one year, 67% at three years and 58% at six years.

Full results: Scottish List offences, retention at arrest

Hazard rate analysis

Proven offence hazard rates compared with national proven offending rate
Retention for Scottish List offences at Charge (not Guilty) or No Further Action
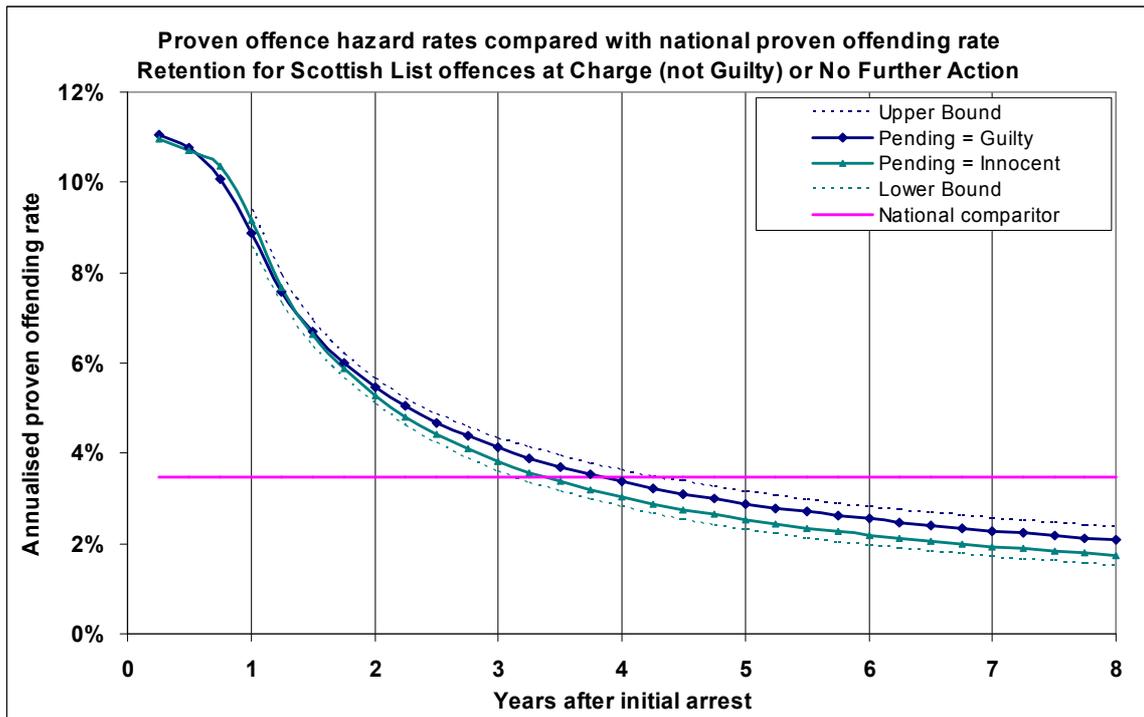
Following the initial arrest, the offending rate is equivalent to 11% of the sample committing one or more proven offence per year. The rate falls off sharply, halving in two years and dropping to 1.9% by year eight.

The national comparator, for the same age and gender demographic, is 3.5%. It takes around three years nine months for the offending rate to drop to this level, with a lower bound of three years and upper bound of four years three months.

Outcome analysis

Retention period	Innocent Profiles on Database	Outcomes at the end of the retention period		
		Retain on proven	Retain on Reset	Delete
1 year	74,000	10% (3.5% Scottish List)	5%	86%
2 years	137,000	16% (5.5% Scottish List)	6%	78%
3 years	194,000	20% (7.0% Scottish List)	7%	73%
4 years	246,000	23% (8.4% Scottish List)	8%	69%
5 years	295,000	25% (9.4% Scottish List)	9%	66%
6 years	342,000	27% (10.4% Scottish List)	9%	64%

There are around 74,000 arrests each year that could result in a new innocent profile being added to the DNA database under this policy. Around 10% of these will receive a proven offence of any type within one year, 20% within three years and 27% within six years. For Scottish List offences, these rates are 3.5%, 7% and 10.4% respectively.

A longer retention policy will result in more retention periods being reset and fewer profiles being deleted at the end of the retention period. The proportion

reset is 5% at one year, 7% at three years and 9% at six years. Deletions are 86% at one year, 73% at three years and 64% at six years.

Full results: Scottish List offences, retention at charge

Hazard rate analysis
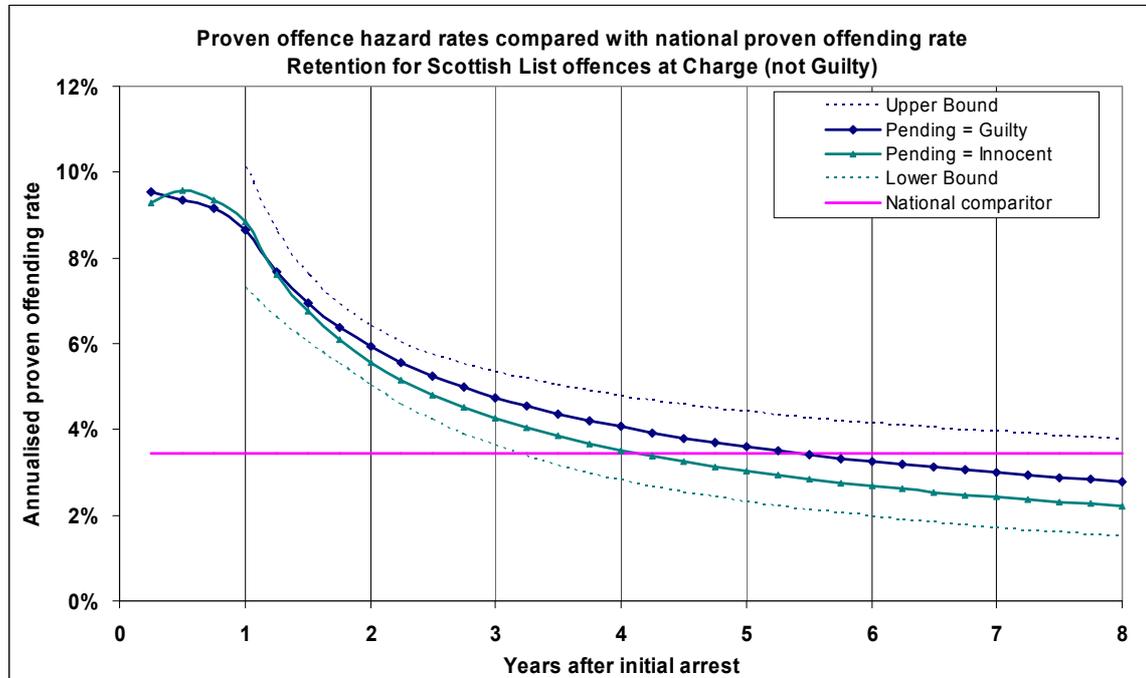
Following the initial arrest, the offending rate is equivalent to 9.4% of the sample committing one or more proven offence per year. The rate falls off, halving in two years nine months and dropping to around 2.5% by year eight.

The national comparator, for the same age and gender demographic, is 3.4%. It takes around four years nine months for the offending rate to drop to this level, with lower and upper bounds of three years and nine years respectively. The uncertainty around the intersection point is due to the small sample size. Only 5% of arrests with no sanction led to a charge for a Scottish List offence.

Outcome analysis

Retention period	Innocent Profiles on Database	Outcomes at the end of the retention period		
		Retain on proven	Retain on Reset	Dispose
1 year	10,000	9% (3.2% Scottish List)	2%	89%
2 years	19,000	15% (5.4% Scottish List)	2%	83%
3 years	27,000	19% (6.9% Scottish List)	2%	78%
4 years	34,000	23% (8.6% Scottish List)	3%	75%
5 years	41,000	25% (9.7% Scottish List)	3%	72%
6 years	48,000	28% (10.6% Scottish List)	3%	69%

There are around 10,000 arrests each year that could result in a new innocent profile being added to the DNA database under this policy. Around 9% of these will receive a proven offence within one year, 19% within three years

and 28% within six years. For Scottish List offences, these rates are 3.2%, 6.9% and 10.6% respectively.

A longer retention policy will result in more retention periods being reset and fewer profiles being deleted at the end of the retention period. The proportion reset is 2% at one year, 3% at three years and 3% at six years. Deletions are 89% at one year, 78% at three years and 69% at six years.