THE FUTURE OF FORENSIC BIOINFORMATION

Seminar Two, 25th March, 2009- Headlines

Session One – Data Sources

This session considered currently available national and local data sources which provide evidence on the current uses of forensic bioinformation. Discussion focused on an attempt to begin to address the following questions:

a) What methodologies are available that might enable us to make meaningful assessments with regard to impact or effectiveness within law enforcement or criminal justice?

b) What information is needed if we are to use such methodologies?

c) Are there any special features of forensic bioinformation that may require us to adapt existing methodologies or treat data differently?

d) What standardised data are collected on the effective uses of bioinformation?

e) Who holds these data, how have they been analysed, and what reports exist that summarise and interpret them?

f) What uses are made of these data/interpretations and by whom?

Discussion was preceded with a presentation outlining: sources of data and their presentation; prevalent criticisms of lack of data; and some examples of the mis-representation of forensic bioinformation by the use of poor data.

1.1. The police have more detailed data sources contained in the annual returns that each force is required to submit to the Home Office. These used to be submitted to Police Standards Unit (PSU) to create graphical representations (referred to as "the stripey charts"). There were questions about the disbanded PSU and what has happened to their data and data collection role. Will NPIA be collecting same (I-QUANTA) data?

1.2. However, even within the police annual returns there is systematic underestimation of the contribution of forensic bioinformation because not all instances where forensic bioinformation is collected are recorded. For instance, if a non-police officer recovers biological material (i.e. a pathologist/ forensic medical examiner) then this will not be recorded in the force annual return (e.g. in the Suffolk murder investigation it was a pathologist that recovered DNA from the third body and so that

is not counted). This means that the actual utilisation of forensic bioinformation in police investigations is underestimated in these annual force returns.

1.3. The ‘purpose’ of the annual returns as a force, and also at an individual level, is largely to monitor the performance of SOCOs in police forces.

1.4. All data collated will only ever go as far as ‘detection’ in the criminal process. The multiplicity of factors involved, and involvement of other agencies means that the police cannot collate data on ‘convictions’. At any rate, it would be near impossible to make judgements (it would require human interpretations of lots of factors/ decisions) about whether the forensic bioinformation had an impact on the conviction/ non-conviction.

1.5. The NPIA do not collect figures on forensic bioinformation utilization, and they do not have the capacity to process such data. The data that are collected are based on ‘counting cases’ in any quarter – but these are not followed further down the process chain throughout the criminal process.

1.6. There remain difficulties over what to do about ‘multiple evidence’ supported detections. How can they be recorded? What if one piece of forensic bioinformation led a suspect to confess to multiple TICs – is this one detection from forensic bioinformation or many?

1.7. A difficulty is that a detection is simply reported as that: a ‘detection’ – there is no depth to this reporting. It is not therefore possible to discern whether the evidence was determinative or whether there was other compelling evidence and the conviction did not hinge/ or even involve the forensic bioinformation at later stages of the criminal process (was it relied upon at trial? Did the jury rely upon it?).

1.8. A difficulty in demanding ‘deeper’ information is that the collation of data has always been about the performance management of SOCOs/ forces/ forensic provision etc. rather than the ‘effective uses’ or otherwise of the evidence collected. It now presents a “long knotty problem” for the police to try and prove (demonstrate?) the ‘effectiveness’ of forensic science because this is not what the data is collected for. It has previously been data collected for performance management and budgetary purposes and it is highly complex to now manipulate this data into being able to reveal different types of information that may now be being sought. The ‘obsession’ has always been with performance management and being able to compare the ‘performance’ of forces rather than making informed operational decisions or setting/ changing priorities.

1.9. Previously, the onus on collating data was to be able to facilitate ‘activity analysis’. Would it now be possible to add different methods of data collection to give more information on different ‘outputs’? This would depend on what the ‘outputs’ were considered to be and whether they were easy to measure? (general feeling that ‘outputs’ were impossibly complex, fluctuating and subject to change and that they would not be easy to measure).

1.10. The ‘ultimate issue’ centred upon the ‘value’ of what forensic bioinformation can add/ does add to the criminal process/ criminal justice system? It may not be complex to create new measures and collate data but would require substantial resources. It would require significant human intervention (which is costly) because the type of data required would necessitate a judgement of values. This also introduces subjectivity. It is very difficult gather to information when humans are needed to

make judgements, and requires a great deal by way of guidance to try and remove variation in decision-making to ensure comparability.

1.11. Any 'rich' data collection would also require the 'buy-in' of other agencies – in particular the CPS. However, the CPS previously have expressed no interest in collection/ returning data.

1.12. It would be argued that all data are difficult to obtain as a certain degree of 'pain' in setting up any new system. The most important element (which is that which is most often not done adequately) is thinking *ahead* of setting up any new system, what the questions are that need answering.

1.13. The issue of large enough sample sizes were mentioned with regard to serious cases – though it was pointed out that there are on average 700-800 murders a year which is a large number. However, it was also suggested that 'dip sampling' could overcome size of sample issues and also lessen the burden of data collection. There could be other 'sampling' strategies designed as with medical trials. Data collection could be done in stages, in particular by considering the uses of forensic bioinformation in the construction of 'intelligence packages'. Forensic science contributes massively to intelligence.

1.14. There was considerable debate over whether there could be an agreed definition of 'value' in criminal investigations/ the criminal process. Is the 'value' that we are looking for merely limited to a 'detection'? The definition of 'value' will depend on who you are considering as the beneficiary (the police, victim, CPS, court, society *et al*?) How to measure the value of an exclusion from an investigation? Is the cost of closing down a major murder inquiry days/ weeks early to be counted? The concept of 'value' will be vague and will remain contestable – adding to the complexity of collating any data that would be seeking to measure the 'value' added by forensic bioinformation.

1.15. However, despite such debate, it was suggested that organizations are routinely measured on their 'effectiveness' so they must have some ideas of what they consider to be 'effective' performance in their context. In previous studies ('Using Forensic Science Effectively') 'value' was defined as 'utility' (this may be a case of simply swapping one contestable concept for another without solving the problem at all).

1.16. An additional complexity is measuring the 'significance' of the forensic bioinformation. Was it crucial to the detection? Was it marginal? Was the 'value' added by other evidence/ other factors in the investigation which the forensic bioinformation may have had a very small/ no impact upon?

1.17. All of these debates may be had, yet there is a political backdrop of moving toward fewer measures, and perhaps simplification of performance indicators (in health and education for example). In policing terms, the 'new' goals/ targets are framed in terms of 'public confidence' and 'customer satisfaction'. These are the new goals that forces will have to work toward (not clear how either of these concepts are to be 'measured' themselves).

1.18. In terms of the contribution of forensic bioinformation to 'public confidence' or 'customer satisfaction', there is an important element of 'reassurance policing' (which is presently underplayed/ not 'captured' in the rudimentary focus upon 'detections' at present). However, there could also be an reversal of the 'value' in that members of the public will have heightened expectations (driven by media – fictitious or not) that

1.19. There are also periodic shifts in emphasis in policing, with changing priorities. There are also significant differences between types of crime and where ‘value’ may be most important – i.e. serious violent crime versus acquisitive crime. These shifts/ priorities will have an impact on definitions of ‘value’ and on the measures put in place to gauge ‘value’.

1.20. The Police National Computer (PNC) is designed to answer single queries very quickly in large volumes. That was it’s specification and it does this very well. However, it was not designed to facilitate any data collection or deliver large ‘chunks’ of data. The PNC is purely an operational police tool, and not useful for statistical analyses. The PNC data is very limited and only really holds demographic/ operational and conviction data. It is therefore of little use for collating any more complex data.

1.21. The South Yorkshire police trialled a project to track cases from crime scenes to court but this project proved very difficult indeed.

1.22. The challenge will be how to change the data IT systems in place – do we need new ones – and if it is possible to ‘future-proof’ such IT systems. For instance – the priority now is ‘public confidence’ and ‘customer satisfaction’ that may lead to certain definitions of ‘value’ that will prompt the collection of particular data, but what will happen if (when?) these priorities are changed – the new system may then again be useless and collecting the ‘wrong’ data.

1.23. How might the police go about determining whether the money that is sent on forensic bioinformation is giving value? In terms of budgetary discipline, this is a question that must be asked and so the police must have a method for answering it (or must rapidly design one to protect the ‘forensic spend’ part of their budgets).

1.24. With specific regard to the retention of forensic bioinformation and how it might be determined what is the most ‘valuable’ length for retention – the NPIA asked other (European?) countries about the basis on which they made the decisions regarding their retention rates? The (apparent) answer was that no other country had any evidence base for their decision on their retention regimes.

Session Two – Research Studies

Session Two considered the availability of operational and research studies of the uses of forensic bioinformation in England & Wales and elsewhere. Key questions included:

a) What studies currently exist?

b) What methodologies do they use?

c) What are their findings and what efforts have been made to collate and compare such findings?

d) What uses are made of these operational and research studies, and how are their results disseminated to professional communities and the wider society?

Discussion was preceded with presentation outlining the range of recent social research on the uses of forensic bioinformation in support of criminal investigations. The presentation was not concerned with the scientific literature that underpinned or described technological innovations in forensic science practice.

2.1. It was pointed out that it is very difficult to get a sense of an accumulating body of knowledge in this field. Most research to date (what little there has been) has been narrowly focused on 'police uses' of forensic bioinformation (though this may not be the most useful term and may exclude other uses). Research has been mostly undertaken/ funded by the Home Office and disseminated through Home Office channels.

2.2. The "glossed" answers to parliamentary questions are fundamentally right – the 'rhetoric' employed does reflect organisational reality. There are lots of cases where DNA is useful but no official 'counting' and there is an 'air-gap' – so you cannot tell whether the forensic bioinformation was relevant further down the criminal process – so you cannot answer these parliamentary questions asking for convictions from DNA.

2.3. The NPIA are looking to more accurately represent the contribution of DNA to detections (because of the 'under-estimation' that currently occurs). There remains a recognition however that no crime is 'solved' by DNA alone. Of course there can be occasions when DNA is the first link that police make with a suspect (i.e. in the Mark Dixie case) but further evidence then goes on to be collected once that suspect is named. Forensic bioinformation can also be used at early stages of an investigation or during the interview stage but then may play no further role, or the case may cease at this stage. There could be compelling evidence against someone which is just forensic science however – so the 'case' against a defendant is DNA + fingerprints + fibres for example.

2.4. Whilst NPIA manuals may seek to represent the 'best' utilisation of forensic bioinformation, this may depart from actual practice in many cases, and such manuals are not a substitute for empirical research on this topic..

2.5. Much of the research which has been done has relied upon 'performance data' produced by the police – but these data have already been shaped organisationally

2.6. There is a conventional distinction between the use of forensic bioinformation in major (serious) and volume crime. The cost of forensic support is always relative to

the severity of crime. Budgetary considerations will always make this the case. With major/ serious cases, there is allowance for more expenditure on forensic science and more leeway for 'speculative' or perhaps 'experimental' forensic analysis to be undertaken, even if costly. This would not be acceptable in volume crime (operationally). In volume crime the investigator will be working within very limited budget and within boundaries of 'knowledge' (so nothing 'experimental/ speculative' undertaken) etc. but in serious crime they are allowed far more leeway to spend/ use lots of different technologies.

2.7. It may still be considered that DNA is expensive but fingerprints are 'cheap'. However, once proper calculations are made – there is very little difference if considering *real* costs of both.

2.8. The NPIA have an ongoing research project looking at all the research literature on DNA (the Campbell Collaboration project) and also one on forensic science more generally (same researchers).

2.9. There was debate over whether there would be publication bias present in the research but it may be that this can be discounted because such a tiny amount of research is actually undertaken. However, publication bias is a real issue that must be accounted for.

2.10. Perhaps there is a new 'orthodoxy' and if so, consideration is needed of how knowledge flows into the system. Where is the knowledge coming from? Who is creating it? At this time, there are cultural issues surrounding the research because nearly all is 'internal'. It is undertaken within the police, by the police, for the police. This is then not peer reviewed etc. and is only published internally (in some instances there may be good reasons for this – perhaps we do not wish everyone to know how certain crimes are investigated for instance as this may be 'useful' to perpetrators of these crimes). These internal documents are then disseminated widely just within the industry. It is not published within 'academic' circles but kept within the profession. There is very little collaboration with external researchers/ academics. The emphasis on current research is on 'lessons learnt' from case work.

2.11. This situation in England and Wales was contrasted with policing research in continental Europe. There they have integrated/ collaborated more with external researchers and there are 'academic' publications. More knowledge is flowing 'in'. Again, this is not the case in the US, where the FBI produces lots of reports but these are primarily only accessed internally – or externally available ones are 'normalised' before publication.

2.12. There remain lots of questions about the quality of the research undertaken 'in-house' by the police and perhaps a lack of partnerships. What may be needed for the future is a more open and confident organisation that is happy to partner with researchers and publish externally the research.

2.13. There is a recurring theme of how hard it may be to undertake research in this field, there are lots of data, most of which are not easily collated (contrasted perhaps with court – based research where a researcher can simply go to a courtroom).

2.14. It was commented that if you are going to "open a can of worms", you have to do the work properly and dedicate sufficient resources to making sure it is done properly. The level of resources that would be required is an issue. Indeed, the difficulty here is that there are no resources. Previously Home office funding has focused on volume crime (for good reason – the UK had very poor record in terms of prevalence

and also detection rates) but the funding is again shifting, in particular to support counter-terrorism.

2.15. A question was raised about what academics can bring to research in this area? Would there be enough cases? Are there enough murders? However, it was pointed out that in medical trials, sampling is undertaken (indeed randomized control trials are undertaken – similar to that attempted by the NIJ, so it can be done). It is possible in medical arenas so while it may be politically/ pragmatically difficult in the CJS, it may not be impossible.

2.16. Prior research undertaken by the police may have had certain alternate 'agendas' (motivations?) and it was suggested that some published works had only a limited evidential basis There was cynicism that in many instances, the funders/ decision-makers were more interested in the current 'buzzwords' – which were largely politically driven. Research would be undertaken if someone could promise to 'deliver' what the funders/ decision makers wanted (i.e. with funding applications you choose those 'buzzwords' that you know will get your applications responded to positively.) This would be the same in research in this field, you need to tap into the current concerns. This may not be the same as gathering the information that you really want. The information that is gathered may largely be dictated by what others want you to find out or know.

Sessions Three & Four: *Knowledge Gaps & Futures*

These final sessions considered the existing gaps in knowledge of the effective and cost-effective uses of forensic bioinformation. Particular attention was paid to:

a) Knowledge of effective management strategies for controlling the collection and use of forensic bio-information in both volume crime investigations and serious crime investigations;

b) Methods for measuring the cost-effectiveness of current and proposed uses of forensic bioinformation;

c) The relationship between knowledge of the effective uses of forensic bioinformation and knowledge of the effective uses of other intelligence and investigative methodologies.

d) The current work and future plans of key stakeholders to request or collect more information on the effective uses of forensic bioinformation;

e) Government expectations of data and the future funding of forensic bioinformation.

f) Whether the currently available range of data allow a sufficiently wide range of stakeholders and interested parties to come to well supported conclusions about the effectiveness and cost-effectiveness of the current uses of forensic bioinformation.

g) What additional kinds of data can be collected to improve knowledge of the uses of forensic bioinformation;

h) Whether lessons can be learned from efforts to collect such data in other jurisdictions

Discussion was structured around a framework derived from two key studies: on the economic evaluation of policing (Stockdale et al. 1999) and a scoping document on the measurement of productivity in the CJS (ONS 2008)[1]. The interest in CJS consumption of GDP had been manifest from the mid-nineties onwards. While there appeared to be a retreat from the use of a whole battery of performance indicators (PI) in policing, evaluation was likely to intensify as recession intensified fiscal constraints. With the highest level of CJS expenditure in G8 countries when measured as a proportion of GDP,[2] UK police forces and forensic providers could expect increasing scrutiny from budget managers and their accountants.

These indicated the likely methodology for government evaluation of the effectiveness and cost-effectiveness of the current uses of forensic bioinformation within overall CJS decision making. The former document also provided a useful summary of earlier attempts to apply economic evaluation to policing and the difficulties in doing so identified by its authors. The organizers emphasized that these approaches were not necessarily the best way of evaluating the effectiveness or cost-effectiveness of policing.

There was risk that such evaluations, reflecting their origins within business and budgetary control were simply two dimensional, incapable of allowing users to imagine the multi-dimensional reality of a CJS that often lacks tidy parameters. For example, the skills and scientific infrastructure required to populate forensic bioinformation data bases might contribute to identifying the dead. In addition to the victims of crime scenes, this might include fatalities in natural disasters such as the 2005 tsunami. Such activity is likely to be well outside the CJS scoping study, but it is something where forensic scientists and forensically aware police officers are the only means that society may have to instantly organize a large-scale response to such calamities.

Hence there two sessions were intended to help indicate the issues and expectations to which forensic science managers and other CJS decision makers would have to have to respond. Hopefully the proposed systematic analysis of the issues would contribute making evaluation as realistic and robust as possible, and less vulnerable to manipulation.

3.1. Stockdale et al. identified the problem of measuring inputs as a major obstacle to cost evaluation in policing. Since the publication of their study in 1999 ACPO had effectively created a competitive market for many forensic science services. At first sight this would result in better input information. It had been suggested during informal discussion at the seminar, however, that one of the results of the procurement exercise had been to create a false sense of the cost of forensic provision, with inadequate account by procurement managers of the ability of smaller providers to guarantee timely delivery and quality of results. The procurement exercise may have led some to believe that 'value for money' was synonymous with 'lowest price'?

[1] Stokedale, J.E., Whitehead, C.M.E. and Gresham, P.J.,(1999) 'Applying Economic Evaluation to Policing Activity', *Police Research Series Paper 103* (London, Home Office) and Office of National Statistics (ONS),UK Centre for the Measurement of Government Activity (2008), *Criminal Justice System: Scoping document* (Newport, ONS).

[2] Research by the organisers suggests that in 2004-05 it was equivalent to 2.5% GDP, of which police expenditure accounted for 61% of overall CJS expenditure. The equivalent figures for the USA (next highest) in 2005 were 2% and 41% respectively, although the costs per 100,000 population adjusted by the 2005 OECD purchasing power parity index were significantly higher.

3.2. The key measures examined by Stockdale et al were ‘cost-effective analysis’ (CEA) and ‘cost-benefit analyses (CBA)? The former is used to compare input costs per unit of output with a decision rule to choose the option with the least cost per unit. The latter is more complex. Having valued all inputs, outputs and outcomes in monetary terms used to compare competing options in order to identify the highest net benefit.

3.3. ONS in seeking to address how to measure CJS productivity define this as ‘a ratio of CJS outputs to CJS inputs at constant prices’. This approach appears to have been based on CBA, but it acknowledges the importance of outcomes in addition to outputs. Moreover, the discussion in the ONS paper contains two caveats about their proposed approach:

- It may be unclear which observed or recorded changes can be attributed directly as outcomes to improved CJS productivity or result from other factors (e.g. is a reduction in the theft of or from vehicles the result of improved CJS detections resulting the detention of offenders or, alternatively, better security by manufacturers and owners)?

- In contrast to the ONS approach, the Administration of Justice (AoJ) methodology developed by the Home Office recognizes the interdependence of CJS agencies, for example the better detection of offenders achieved by the police may be negated an the absence of effective follow-up by the CPS, courts and NOMS.

3.4. Comparative analysis may help to avoid incorrect attribution. While the decline in burglary in England and Wales may have largely coincided with the DNA Expansion Programme, a similar trend can be observed in other countries that did not expand their forensic DNA databases at the same pace. Yet more detailed analysis indicates the much greater prevalence of burglary in the UK compared with other EU member states and that the proportion of DNA detections for individual offences appeared to decline more slowly (in one case increased) than the overall volume of such offences.

3.5. The ONS study also recognized that ideally evaluation should reflect value weighting or quality-adjustment instead of basic numbers/costs. Not all outcomes are of equal concern or importance (e.g. offences against vehicles (13.3% of 07/08 recorded crime) are more tolerable than burglary (5.7%) Perhaps cost-utility analysis in health economics (e.g. ‘quality adjusted life year’ that reduces inherent discrimination in CBA against people with lower earning capacity) offered models that could be adapted for the CJS?

3.6. Moreover the very notion of standard measures of CJS productivity seemed to be potentially rife with paradoxes? For example, a decline in productivity might reflect a switch of resources to more serious/harder to detect crimes. Conversely more detections might simply reflect more crime or changed priorities/data manipulation; the great majority of drug offences are recorded through the act of detection.

3.7. The most promising area for the evaluation of what can be achieved through forensic science might be major crimes. Case reviews provided an opportunity to obtain data from case support and scientific support units. With the analysis of a sufficient number of cases it might be possible to robustly estimate the cost-benefits of forensic science inputs. Although the full range of inputs/outputs might be overlooked, miscounted or unquantifiable (e.g. a scene of crime visit that did not yield trace bioinformation but resulted in effective crime prevention advice being acted on).

3.8. There are always going to be different levels of ‘value’ in such evaluations. There may always be a cash value for the decisions facilitated by particular types of forensic evidence. There will also be intangible/ (intrinsic) ‘values’ in terms of the ability to successfully detect crimes (community values) and these may be harder to capture and measure.

3.9. There can be strategic targets which have secondary benefits, for instance, targeting drug factories may have a wider positive impact by taking dealers and associated crime off the streets. It is also known that there may be less benign secondary impacts that are sometimes not accounted for – criminals turn to other crimes (eg from street robbery to shoplifting) or other areas (displacement). How are these secondary impacts to be measured in addition?

3.10. However, there may be a danger that causality is assumed. There can be regression to the mean and and impact on the quality of trying to get the right data. Much evaluation is not robust enough. There may always be an element of ‘randomness’ in crime data in particular which cannot be accounted for or monitored, but ‘half-decent research should seek to eliminate such problems.

3.11. The greatest difficult in assessing CEA/ CBA is the ‘hyper-variability’ of the problem. There are so many factors that need controlling/ accounting for. Essential, therefore, to state assumptions and hypotheses so that the methodology is clear.

3.12. Research may be undertaken but may require many caveats because of the complexity of the social and legal dimensions that may affect interpretation. Moreover, many pilot studies or focused work produced results that reflected a level of resource that was not available when working normally and such results were unlikely to be replicated in national trends.

3.13. It was pointed out that the switching of resources/ priorities from one crime type to another is extremely difficult. There is always an element of policing that has to be demand led. To ensure any public satisfaction the police have to respond to what the public see as important, such as burglary and car crimes. It is difficult to respond to new offences such as cyber crime. This is because of the need to redeploy resources from existing activities. Some of the latter may have an emotional resonance with the public. Most often the work is actually quite static. There could be an element of management that monitored levels of crime and then would dedicated resources to ‘problems’ and then they may see a reduction in that crime/ problem. However, this is more often than not temporary and it may be that what is required is long-term strategic changes introduced incrementally over a significant period of time.

3.14. The ‘problems’ with research that has been undertaken to date is that much of the research was on a small scale and may not have big enough to detect what may be achieved by different approaches. At this scale there was also a risk that more dramatic results were accidental.

3.15. Also much research does not answer the questions that we are now asking. The attrition studies were intended to identify where cases fell out of the system not crime reduction or displacement, and for all forms of investigation not just forensic science. We are now trying to fit bits of research into very different agenda and asking questions of earlier research that it cannot answer.

3.16. A further problem with much of the discussion about data is that it is necessary to distinguish between information collected for two different quite different purposes:

management information intended to drive better behaviour and other data needed to qualify the value of work.

3.17. Instead of collecting what is really useful we have tended to count what is easy to count. It is making sense of this readily data available that is often a problem. There is also a cost associated with the very act of counting.

3.18. Economic analysis can have the effect of driving dysfunctional behaviour and affect decision-making. There is a need to understand the limitations of decision-making on economic bases and the possible distortions this may lead to. For example, it is know that measuring productivity against particular measures/ targets can impact upon the performance in one area but have knock-on impacts. It is vital that the flaws in 'performance indicator' methodologies are understood. The Royal Statistical Society have previously researched the area of PIs. It is well understood for example that 'what gets measured gets done' – there is a need to be able to spot that behaviour and minimise the possible negative impacts (e.g. perverse incentives etc.).

3.19. One difficulty with evaluations is the lack of stability within the criminal justice policy-making arena – very often politicians are short-termist (responding to almost cyclical media pressures, such as the theft of mobile telephones, firearms offences, knife crimes, and counter terrorism or drugs policy, irrespective of what the most reliable statistical analyses and clinical research suggest)[3]. There is a need to look at outcomes/ outputs and how they are valued and not be deluded perhaps that politicians will respond 'correctly'. The experts can build better models and can make evaluations about long-term values/strategic objectives for the criminal justice system etc (one example that longer-term CJS transformation is possible was the abolition of capital punishment against the opposition of large segments of the press and the instincts of some senior politicians.,

3.20. While these political problems existed in the CJS there are lots of models from health research (though of course there is a lot more money in health research) and economics. This area of public policy was even more complex in terms of the difficulties of making decisions, the personal consequences for large numbers of people and local/national politics. Constructing the model and doing the research will take lots of money. It could be that we require something along the lines of a measure similar to that used in health of 'quality adjusted life year'. Indeed there are lots of other areas of econometrics where research combine objective and subjective measures and seek to evaluate short-term and long-term impacts.

3.21. Also potential lessons from health research in terms of how research is done and how lessons are promulgated. There was an issue arising from the commercialization of forensic science research over the difficulty in bringing new products to market. Do we need a (National Institute for Clinical Excellence) 'NICE' for the criminal justice system. Should this be the role of the NPIA? Does or should NPIA have the same degree of technical, scientific and professional independence? The role of/influence over ACPO in relation to NPIA may be significantly greater than that of the GMC in relation to NICE? While it has some way to go the Forensics21 program, however, could be a move towards the health model?

3.22. It was argued that NICE remains fairly conservative, it does not tend to fight battles with large drug companies nor does it always adequately asses the economic utility

[3] A further point is that decisions to change priorities and redeploy resources may be presented to give the impression that additional resources have been provided for CJS purposes.

of outcomes. However, what NICE does is makes the decision-making process transparent. It also relies on a large knowledge-base and below NICE there are other layers of knowledge base including the MRC trials base. The NPIA could not be comparable because of the lack of depth to their knowledge base and the recurring question of where their knowledge is coming from. This is bound to be constrained by the absence of an equivalent to laboratory conditions for, as an example, viral research and ethical constraints on the ability to treat part of the population as a control group.

3.23. The police have relied on 'Professional practice guidance' which forms a police doctrine. This again does not have a large knowledge base and is largely non-evidential in nature. Again a medical analogy would be to look at the intention to introduce regular audits of clinical practice and in various areas enhance peer review. What had been earlier in the seminar about the review process at specific time intervals might be seen as a move in the same direction.

3.24. There is now a movement towards 'competencies'. The Forensic Regulator is concerned with demonstrating 'competency' (using UKAS). UKAS does use blind trials but these could be improved. It is not the quality of the science perhaps that needs auditing/ regulating – but the decision making and economics that surround the use of the science. What impact is the creation of a forensic marketplace having? What will happen in the further and the impact on forensic science is perhaps what needs to be monitored more importantly. Within England and Wales, the purely commercial provision of forensic science means that economic pressures are going to drive decisions made by the forensic companies because they have economic imperatives. For example, most obviously, R&D – the FSS was expected by the Home Office to invest £10 million each year to ensure innovation but there were fears that this could get squeezed by procurers seeking the lowest price against competitors who were not operating under a similar obligation – so one potentially valuable 'output' could be lost because of the commercial economic model. This suggested that the CJS productivity measurement of reduced input values should qualified by any consequential loss of output in terms of research and development investment.

3.25. Questions remain about whether the ONS approach will operate at the right level of detail? An alternative approach – followed by Roman et al – when evaluating the impact of enhanced forensic activity was to take existing policing inputs and outputs as fixed. The evaluation might only attempt to measure the costs of changing the use of forensic science in terms of both the science itself and any additional policing cost directly resulting from the different approach.

3.26. The different Home Office initiated approach to productivity measurement was based on a broad typology of crimes (distinguished between indictable and summary offences) to achieve a measure of homogeneity in inputs and outputs within the CJS as a whole. Violence against the person, sexual offences, burglary and robbery were potentially more relevant than most to the use of bioinformation, but even then the were significant questions about the range of crime covered by the category of violence against the person and the different level of resource that might be invested in an inquiry. Any research reflecting such categories would need to be based on further sub-division and while a domestic murder with a confessed perpetrator would not require the same level of forensic work, this might still be needed to foreclose what might otherwise be potential lines of defence although as a genuine and critical output it would not register as a detection and possibly not as contributing to case construction.

3.27. Timeliness of effective intervention did not appear to have been considered yet as a possible output or weighting for outputs. Sometimes this might be relatively straightforward or uncontroversial, such as the estimated 3.5 hours of processing time saved in Operation Lantern detections. In most instances, however, simply working out 'costs' will be difficult – for example – with a 'spree offender', what is the 'benefit' if you stop taking a prolific burglar off the streets in two weeks rather than the seven indicated in the early SWIM results from Derbyshire discussed at the seminar might result in a much lower measure of productivity for that detection than a slower forensic and investigatory process. It was agreed that notional costs could be calculated for police costs saved (as sunk costs not cashable savings using the analysis based costing data collected by the Home Office in recent years) and estimated economic consequences for the victims of burglaries available from Home Office research. This would be highly contestable – similarly to arguments over assumptions made about youth justice interventions – because it is looking at individual behaviour retrospectively and in an isolated manner and to which is then applied standard values for whole populations that will not be valid in all instances. There is a requirement perhaps that we should adopt a sensitivity analysis making transparent the underlying assumptions used and a range of values calculated from these different assumptions rather than create a 'too' precise a number which could end up being wholly unrealistic and lacking in credibility?

3.28. There is a pressure on forces to demonstrate 'cashable benefits'. But how then to factor in public confidence? How can we say that with current methods we can catch a prolific offender, on average, in 70 days, but if we spend more money we can do it in 14 days. The public will then surely demand the 14 days to become standard and then there can be 'customer satisfaction' but at what cost? Moreover, today (perhaps not 10 years ago) most public concern does not reflect volume crime, even serious crime, or even investigative failings such as with Napper and the black cab rapist.

3.29. In attempting evaluations of forensic bioinformation contribution - what do we define as 'benefit' – who is the benefit for? Is there a link that can be made between improved forensic performance and public satisfaction? (Public confidence and customer satisfaction being the current 'imperatives' of the police).

3.30. Any evaluative model can only be a fairly abstract tool and will necessarily be retrospective. It will have to be based upon assumptions about behaviour, i.e. how different types of people / crimes/ criminals behave. Extending such assumptions would be highly contestable (e.g youth offenders versus career criminals?). There are also very different approaches in England and Wales and Scotland in what they think can be done about offending – all emanating from the same data. Any generalisation of the model will be questionable.

3.31. Prior to 'SWIM' there was a linear model in place developed for PSU. The assumption was that simply attending more scenes to collect additional forensic material would result more detections. SWIM was developed in response to concerns that where there backlogs in processing or acting on bioinformation and other forensic intelligence greater scene attendance would only add to the backlogs and would not necessarily improve detections or to the degree suggested. Forensic scenes work had to be matched by equally rigorous and efficient processes all the way through the CJS. This linear model is clearly too simplistic. Any statistical analysis will also have to figure in changes in significance of forensic bioinformation as progress is made through stages of the criminal process. Within the SWIM analysis decisions made during the stage of acting on the identification was critical in

terms of outcomes achieved from the forensic work but was outside the control of those providing forensic services.

3.32. Other problems in evaluation arose when benefits took time to be realised or were dispersed. The Dutch were able to demonstrate back in 2004 that without necessarily achieving immediate detections that could build up via their DNA database patterns (by offence and location) of collaboration between different criminals that might later prove to be significant later in the light of new (possibly quite different) information. While the Belgian database had been used to demonstrate that DNA could link serious offences committed in that country, the Netherlands, Germany and France, yet the inputs and outputs for this activity would be scattered over four jurisdictions.

3.33. In the US a lot is being learnt from the deconstruction of DNA exonerations (where innocence is largely now uncontested). It may be that a similar de-construction in England and Wales of miscarriages of justice may be useful. You could map where 'mistakes' occurred in the criminal process. The study in question[4] not only exposed the limitations of techniques such as serology and bite marks compared with DNA, but also how scientific testimony might not be credible or be distorted in cross-examination or when the case was summed up. However, this is very difficult approach (and we do not have similar 'uncontested' exonerations – at least only a handful and might require that restrictions flowing from commercial confidentiality in respect of techniques and individuals as to matters of personal competence.).This is something within the remit of the CCRC but they had not undertaken a great deal of research except a study of 80 initial referrals within which sample there were said to be 26 examples of flawed or problematic forensic science and the absence of a database that could be interrogated had been an obstacle to systematically following-up concerns about expert evidence exposed by the *Cannings* appeal. The Omagh trial was seen to be a clear example where significant errors had been made during the trial – not just in respect of the actions of police officers but in understanding the scientific issues.

3.34. The communication of forensic evidence in court is crucial and yet under-researched. The focus of what research there is all on police stations/ laboratories. Yet what is going on in a courtroom? The significant issues with how forensic bioinformation is used/ portrayed/ dealt with during trials is highly complex, and for raised questions about the ability of the adversarial system to cope properly with expert evidence. There are a multiplicity of other factors in play, that mean that research into forensic bioinformation can only ever talk about detections and not convictions.

3.35. One of the interesting outcomes of the day – linking this seminar with the first in the series - had been the extent to which discussion had covered issues of governance, independence of decision making (both policy and resource allocation) and transparency. The Government is going to have to make decisions In terms of effectiveness and costs, but also with regard to the Marper ruling and the Information Commissioner report on retaining data relating to previous convictions (the Five Constables Case). Pre-2007 the government decided that general policy would always be to keep data. That now has to change and retention has to be justified in terms of proportionality. This presents a huge problem for forensic bioinformation because retention is fundamental and is an area that demands actual research before the sort of decisions required can be made.

[4] Garrett, B.L. and Neufeld P.J. (2009) 'Invalid Forensic Science Testimony and Wrongful Convictions' *Virginia Law Review*, 95:1-97

3.36. This returns the debate back to ‘knowledge’ – who knows what is going on and who is going to take the decisions? The people making recommendations about forensic bioinformation often know nothing about the topic. Indeed, far fewer people know or understand forensic bioinformation than those that are making decisions about it. However, it was pointed out that the public do not necessarily need to understand the science, it can be put in layman’s terms. For the most part, good judges or politicians can get understand that which is explained well. What is required is a richness of communication between a range of stakeholders.