

**Ethics Committee – WMP & Home Office Responses May 2020**

**NDAS Update**

Below table shows WMP Responses to the Committee's queries:

Questions & Notes From Committee	NDAS response
<p>On Modern Slavery, encouraged by good performance of algorithmic methods to identify MS outcomes. Support independent evaluation of methods by academic group. One important concern is the following: 'there are no plans to share the output of this model with immigration authorities' is NOT an outright commitment not to do so. The language needs to be sharpened much more if a commitment has to be made.</p>	<p>The response is intended to make it clear that this is a police tool. As per the original submission, we are constrained by the general ability of law enforcement and partners to safeguard and support victims, and victims of modern slavery are afforded certain protections support and advocacy that are well beyond the scope of this project. We cannot give a guarantee that no information from this model will be used by Immigration Enforcement. We do not believe that a theoretical possibility that a victim may be subject to immigration enforcement is a reason not to employ the model's potential to identify and safeguard victims of trafficking who may otherwise be unidentified.</p>
<p>The legal advice/guidance included on MS needs to be re-assessed. I am not sure if the author of this guidance is someone internal – it seems to be so. If it is, external advice should be requested.</p>	<p>The legal advice is from the West Midlands and Staffordshire Police's Joint Legal Services. I am not aware of any precedent for requesting legal advice from an external organisation and it is not clear what is to be gained by seeking separate pieces of legal advice at this stage.</p>
<p>I think it is very good we're building up a rapport between NDAS and the Committee - but in a point I repeat several times below, now is the time for NDAS to set out/begin its work on public engagement and transparency, just as the Data Lab have begun to approach this. I applaud the conscious use of proportionality analysis to guide the legal advice given on the Modern Slavery prevention and analysis project.</p>	<p>We appreciate these comments and agree that we have to have a public engagement strategy especially for the MSV use case although I suspect to be effective this will now have to wait for the Covid-19 crisis to abate. In relation to proceeding with the MS use case, it is my intention to proceed one step at a time, so the Ethics Committee will be able to review, comment and shape the development of the use case.</p>
<p>I am not clear on what the project team are looking for where providing opportunity for the Ethics Committee (EC) to input into the most serious violence project – should the team wish to engage with the operational team leads for the 7 West Mids local authorities then I Chair a Community</p>	<p>In relation to the MSV use case, we will seek to feedback to the Committee in due course the result of work we are doing to re-frame the problem statement and data discovery, I suspect the remainder of this point is</p>

<p>Safety forum for this group and can assist in facilitating this consultation.</p> <p>Ethical Framework: Appreciate the continued updates and the invite for the ethics committee to be involved.</p> <p>Legal advice provided in relation to the modern slavery project is clear and well structured, making it easy to understand and advise</p>	<p>similar to the public engagement point above. These are issues we can discuss next time we bring the use case to the Committee.</p>
<p>On a point of detail, what, please, does CDEI stand for?</p>	<p>CDEI is the Centre for Data Ethics and Innovation.</p>
<p>I observe that a consultation is to be launched to a wide range of stakeholders and experts. I very much suggest that a recommendation go forward that included in those to be consulted should be my own professional body, the Law Society of England and Wales. Reposing in certain of its specialist Committees are acknowledged experts in digital and other technology and its place in criminal law and its practice. The two Committees that most readily spring to mind are the Criminal Law Committee and the Technology and Law Committee.</p>	<p>I fear that if we are to approach other committees, particularly at this late stage, we risk having so many conflicting opinions we won't make any headway. For this reason I certainly wouldn't want to engage with two separate committees from the same organisation. The Technology and Law Committee would seem the closest fit but would risk re-visiting a lot of the work that is already being done around data protection, cyber security and the regulatory and ethical debate over the use of machine learning. I think we would need to be persuaded of what the specific benefits are likely to be before we use the taxpayers' money</p>
<p>Thank you again for another clearly laid out submission and for the legal advice.</p> <p>Ethical framework</p> <p>In terms of the NDAS proposals for building a specific ethical framework, I'm unclear on the need for a specific framework for NDAS when there are presumably ways that a more national and consistent approach can be developed (via NPCC, PCCs etc), ensuring that any framework can be overseen by appropriate committees/bodies, as recommended in the recent RUSI report (Babuta and Oswald 2020). My strong view is that any framework and process needs to be a) informed by the legal framework b) subject to clear terms of reference c) independent and transparent d) with teeth, in terms of a commitment to respond to the recommendations in a transparent way. It would make sense in my view (rather than reinventing the wheel) for any new national/regional process to use the WMP</p>	<p>As the submission sets out, our stated desire to set up a project-based ethical framework grounded in governance is based on existing sources of national guidance, made to fit the policing context. It is also unclear at this stage when a national ethical framework for analytics in policing will be developed. The ethical framework is not intended to replace any current framework but fill a gap that we perceive in the delivery of advanced data analytics on policing, in that there seems to a framework for the development of the technology, but this does not seem to go as far as guiding the tactical delivery of policing interventions based on the data analytics. In other words, once the technology has been developed and the insight is delivered to a force what guidance and constraints are put around the policing tactics that are put in place as a result.</p>

framework as a model, as this already has a) the law as a clear underlying principle b) clear and published terms of reference c) clear principles against which proposals are measured d) operating processes and procedures, and experience of adapting these for new contexts and e) a process by which the PCC and CC commit to responding transparently to recommendations.

Most serious violence

I would suggest that further thought is given to why, with the current data available, the model did not work. As interviewees in Babuta and Oswald 2020 pointed out, there is considerable value in knowing what doesn't work as what does. It is also not clear why further work to explore 'alternative problem statements' has immediately started, without such analysis being done and also without (it appears) a clear business need being identified. Again, Babuta and Oswald 2020 pointed out the issue that 'capability development is largely driven by data science, with comparatively little focus on the underlying conceptual framework, criminological theory or legal requirements.' There seems to be a need here to assess whether there could be some underlying issue with these types of models in these 'violence prediction' contexts e.g. that the police data available does not represent all the relevant information that needs to be taken into account (and not necessarily through data science/algorithmic methods), and to give some further consideration to the need for relevant theories to inform the models.

Modern Slavery

I continue to be supportive of this use case. As the legal advice pointed out, there is a clear public good in tackling this crime and identifying potential new leads within existing data. My main question is around the effectiveness analysis. If I have understood correctly, this appears to say that of, say, 100 cases being reviewed, 60 of those would be correctly identified as involving modern slavery. And therefore a considerable number would not be identified. If I have understood this correctly, then this would suggest that great care would need to be taken in how this tool should be operationalised, in particular to ensure that it is not introduced

With regard to the WMP framework, we are not aware of any framework that guides interventions and the only WMP ethical framework we are aware of is the terms of reference for the Ethics Committee, which will still be used as a guide for how the project and further use cases are developed. The overall goal, however, is for NDAS to stand up internal governance processes as well as be subject to oversight mechanisms (as it is currently).

We welcome the comments on the limitations of AI as a 'violence prediction' tool and also note that the RUSI occasional paper *Artificial Intelligence and UK National Security Policy Considerations* contains similar conclusions. We are looking to develop the MSV use case to be very much more of a tool to support existing decision-making processes, in line with Commissioner Cressida Dick's comments on 'augmented intelligence' at the RUSI annual security lecture in Feb 2020. We expect to be able to update the Committee on our progress in due course.

The figure of 60.6% quoted is the Recall figure (Recall = True Positive / (True Positive + False Negative)). The actual *False Negative* figure, i.e. number of cases of MS that are not identified by the model, is 17.02%. All practitioners we have engaged with agree that the model presents an opportunity to identify cases and networks that the partner force will be unaware of, with the obvious benefit of being able to safeguard victims, and it should be noted that the model shows an increasing number of cases tagged as modern slavery events which means that the importance of natural language processing in identifying events will decrease in terms of the overall proportion of events identified. However, the point that the model will cannot replace human decision making is well made, and the deployment of this and any other model of its type will come with a caveat that it is designed to augment rather than replace the identification of cases and allocation of resources by practitioners and decision makers.

<p>as a replacement for other methods due to the risk that therefore other cases might be overlooked. The final statement in the document 'The proof of concept showed...both suspects and victims' might therefore be rather over-egging the abilities of the tool.</p> <p>As a more general comment, I would suggest that the advice given by David Spiegelhalter in this article about FR 'accuracy rates' would be good to apply across the board, in particular the avoidance of terms such as 'accuracy' 'false positive' 'probability' etc, and instead 'explain everything in terms of what we would expect, or what was observed, in a specified group of people of defined size.'</p>	<p>In relation to the final paragraph and the Spiegelhalter and McConway article, I have double-checked our latest submission and cannot find any reference to 'accuracy' or 'probability' at all. The very point of using Recall and Precision rates is to use the statistical standard of a full understanding of the model. We will attempt to make our assessments more accessible however, perhaps in this format:</p> <ul style="list-style-type: none"> <li>▪ Out of a population of N (people in our data) we would expect to identify X people as involved in MS</li> <li>▪ Of those people, Y will have been falsely identified.</li> <li>▪ Within this population, there will be Z people involved in MS who the model will not identify</li> </ul>
<p>The finding that the MSV use case had coding errors which render it unviable should serve as a stark reminder around the risks of this kind of AI/tech – in the worst case scenario, inaccurate models could result in coercive or other sanctions against people for which there was no reasonable basis to have predicted their criminality – this risked harming young people's/anyone's lives despite the clear warnings – however, it is good to see the team having evaluated its own work and identifying flaws from which to start again.</p> <p>Regarding the modern day slavery use case:</p> <ul style="list-style-type: none"> <li>o The objectives of this use case are really positive in as much as they aim to address existing limitations in policing around a very complex issue</li> <li>o It would be helpful to provide the ethics committee with more information on the use case examples referred to</li> <li>o There would be real value in any evaluation of this model being tied to the effectiveness of referring victims to appropriate support services when discovered as a result of this AI</li> <li>o It would be really helpful to arrange a modern day slavery stakeholder advisory group as a priority, to gain the opinions and perspectives of the third sector and other experts at this early stage, which may well help in the design of the model</li> </ul>	<p>I agree that it is vital that models are thoroughly evaluated to ensure that there are no flaws before they are operationalised and it is important that we all recognise the fact that this might take some time. It should be noted however that due to the "predictive" nature of the MSV model we would have gone to every effort to avoid any sort of coercive intervention if decision-making was based solely on machine learning.</p> <p>In relation to this request: <i>'It would be helpful to provide the ethics committee with more information on the use case examples referred to'</i>, we will provide some examples of cases where the NLP has been applied to events to the Committee. However, these will not be suitable for publication.</p> <p>Although perhaps out of scope of the project, I agree that there is a real opportunity when identifying victims to understand what support is available to those victims.</p> <p>Since the original submission, we have engaged Justice and Care, a charity with a strategic focus on modern slavery, and we are also going to be engaging with the Office of the Independent Anti-Slavery Commissioner.</p>

	<p>The Commissioner has team members responsible for Victims and Partnerships, Prevention, and Research and Innovation who will be involved in this engagement. On a more local level, each force where the model is implemented will be part of an anti-slavery network (e.g. the West Midlands Anti-Slavery Network chaired by Robin Brierley) and the engagement with these networks will naturally lead to developments in the model.</p> <p>It is also worth noting that the operationalisation of the use case is not the end of the development – the concept of NDAS is that use cases will be constantly reviewed and refined.</p>
<p>I think the bullet-pointed resources on p. 4 are all fine as a basis of the NDAS, but I’m not sure how easy it will be to integrate these with the adapted Nolan code or other force statements of values. The Nolan principles were devised to respond to the “cash for questions” affair and, though widely used, do not seem to me to be well suited to the public sector in general.</p> <p>Data analytics in policing is easiest to justify for the prevention and prosecution of serious (high-harm) crime or high-volume medium harm crime. Data analytics does not need a special justification where it contributes to efficiency (as in automated searching and matching) except where it conflicts with data protection principles.</p> <p>There are special obligations on police not to behave arbitrarily and to be able to explain to the public what they do (Data analytics is not easy to explain; decisions involving detention or use of force should never be delegated to machines partly for this reason.</p>	<p>I am unclear on the rationale behind the comment on the 7 Standards of Public Life (‘Nolan Principles’). These standards explicitly apply to public office-holders including the police and have been adopted by the College of Policing (with the addition of ‘Fairness’ and ‘Respect’) for the police Code of Ethics. As such these principles need to be the bedrock on which we form any ethical framework.</p> <p>I think the second paragraph could potentially be the start of an interesting debate about where and how much justification and scrutiny different use cases require. In terms of data protection principles, the Committee can be assured that there is no conflict with these Principles for either use case, and there is constant scrutiny in these issues from relevant experts in each partner force. We are also engaged with the Information Commissioner’s Office and are due for a day-long workshop when the lockdown is over.</p> <p>As per previous submissions, we agree entirely regarding the need to be able to explain any model as well as the fact that we as the police will retain responsibility and accountability for any decision that is made, whether or not it is with the support of data analytics. With regard to the critical need to explain the output of NDAS models to not just end users within policing but also affected individuals, we are taking ICO guidance on the topic and embedding it into how we operationalise going forward.</p>

## Modern Slavery (MS)

Advice from the Committee:

- Outcome b) – proceed with minor amendments (see below).
- Care should be taken, when operationalising the tool, to ensure that it is designed to augment rather than replace the identification of cases and allocation of resources, and that the potential error rate is highlighted to decision-makers.
- The committee requests that the NDAS team returns to the committee once plans for operationalising have been developed for further advice, and at this point, provide more information about the planned use cases.
- The committee recommends a modern day slavery stakeholder advisory group as a priority, to gain the opinions and perspectives of the third sector and other experts at this early stage, to feed into the development of the model.
- In order to ensure that communications around the use cases are accessible, the committee recommends that the following format is used:
  - Out of a population of N (people in our data) we would expect to identify X people as involved in MS;
  - Of those people, Y will have been falsely identified;
  - Within this population, there will be Z people involved in MS who the model will not identify.

## Most Serious Violence

Advice from the Committee:

- Outcome e) – further information required (see below).
- As it is currently unclear how this use case will be developed (if at all), the committee requests that the NDAS team returns to the committee once a plan of action has been developed.
- The committee advises a need to assess whether there could be some underlying issue with these types of models in these 'violence prediction' contexts e.g. that the police data available does not represent all the relevant information that needs to be taken into account (and not necessarily through data science/algorithmic methods). The committee advises that this exercise is undertaken alongside consideration of any new MSV use-case(s) and the results discussed with the committee.

## WMP – Analytics Lab

### Serious Organised Crime Network

Below table shows WMP Responses to the Committee's queries:

Questions & Notes From Committee	WMP DAL response
<p>The method could have been better illustrated, with an example, in the briefing note.</p>	<p>Please see discussion in general remarks section.</p> <p>This project is at the proposal stage, so there are no examples to illustrate the method as yet.</p>
<p>The method also needs to be assessed on historical or retrospective data.</p>	<p>Any inaccurate data could be notified to the Lab and incorporated into any future runs of the analyses. Findings from initial runs of the analyses will be subject to checking by the Intelligence department.</p>
<p>There is not much information presented and the legal justification of public interest etc. is more than a little self-serving. We should seek more independent legal advice.</p>	<p>Reducing the harm caused by criminals, in particular through violent crime, is a priority for WMP and the PCC, as well as nationally. The National Crime Agency (NCA) states that, '<i>SOC affects more UK citizens, more often, than any other national security threat</i>' (National Strategic Assessment of SOC 2019). Analysis shows that those involved in SOC cause the greatest harm to individuals and communities. It is therefore in line with the policing purpose to pursue the identification of members of OCGs and to dismantle these groups.</p> <p>All proposed projects have been reviewed by Staffordshire and West Midlands Police Joint Legal Services. Their advice has been made available to the Committee. There appears to be no precedent for seeking legal advice outside of this service.</p> <p><a href="https://www.westmidlands-pcc.gov.uk/your-commissioner/police-crime-plan/police-and-crime-plan-2016-20/">https://www.westmidlands-pcc.gov.uk/your-commissioner/police-crime-plan/police-and-crime-plan-2016-20/</a></p> <p><a href="https://west-midlands.police.uk/about-us/vision-and-values">https://west-midlands.police.uk/about-us/vision-and-values</a></p> <p><a href="https://nationalcrimeagency.gov.uk/who-we-are/publications/296-national-strategic-assessment-of-serious-organised-crime-2019/file">https://nationalcrimeagency.gov.uk/who-we-are/publications/296-national-strategic-assessment-of-serious-organised-crime-2019/file</a></p>

<p>Details on errors in the process if automated require significant exploration here, specifically if information is used to direct resources etc.</p>	<p>Any inaccurate data could be notified to the Lab and incorporated into any future runs of the analyses.</p>
<p>What definition of “most harm” is utilised?</p>	<p>To calculate the levels of harm caused by each of the nominals linked to the networks we would use both the Crime Harm Index (CHI) developed at Cambridge University and the Crime Severity Score produced by the Office for National Statistics (ONS). The CHI essentially uses sentencing guidelines to score offences according to the level of harm they cause. For example the harm score of ‘Murder’ is 5475 whereas for ‘Abstract or use without authority electricity’, it is 1. The ONS score is a similar idea but using average sentences handed down by courts.</p> <p>We can calculate the harm caused by each network by aggregating the harm score of the offences committed by its members. This approach to understanding crime harm is mandated in the Police and Crime Plan.</p> <p><a href="https://academic.oup.com/policing/article/10/3/171/1753592">https://academic.oup.com/policing/article/10/3/171/1753592</a></p> <p><a href="https://www.westmidlands-pcc.gov.uk/your-commissioner/police-crime-plan/police-and-crime-plan-2016-20/">https://www.westmidlands-pcc.gov.uk/your-commissioner/police-crime-plan/police-and-crime-plan-2016-20/</a></p>
<p>“If a (potential) relationship exists between individuals (e.g. have they ever been arrested together, etc.)” – Can we have other examples?</p>	<p>As well as whether nominals have been arrested together, other evidence used is intelligence reports linking nominals and prison data as to whether nominals have been in the same part of the same prison for at least 30 days.</p>
<p>Does the exploration include analysis of information on victims?</p>	<p>The focus of the SOC analysis is on the offenders involved in organised crime. However, it is acknowledged that some nominals linked to SOC networks will have been both victims and offenders at some point.</p>
<p>If automated, how will errors in intelligence data be addressed?</p>	<p>Any inaccurate data could be notified to the Lab and incorporated into any future runs of the analyses.</p>
<p>This project seems very well thought through, and is clearly articulated. The only thing that needs more clarity is how the intelligence</p>	<p>Intelligence reports are graded according to the national ‘3x5x2’ process which is undertaken by trained Intelligence Officers:</p> <ul style="list-style-type: none"> <li>• The person submitting the intelligence assesses the reliability of the source of the information as ‘reliable’, ‘untested’ or ‘not reliable’. Reliable information</li> </ul>

will be sifted for reliability/credibility - some examples of how this will work would be reassuring and useful.

could be CCTV images; untested could be an anonymous report via Crimestoppers.

- The intelligence is also assessed based on how it came to be known; or can be corroborated by other sources; whether it is 'known directly to the source', 'known indirectly to the source but corroborated', 'known indirectly to the source', 'not known' or 'suspected to be false'.
- The third element deals with who should have access to the intelligence and how it should be handled.

The 3x5x2 system replaced an earlier grading system known as 5x5x5 in 2016. The table below shows the grades from the two systems which are perceived to be 'credible' and therefore used in the SOC Network analysis:

Included	Old		New	
	Source	Information	Source	Information
Yes	A - Always reliable	1 - Known to be true without reservation	1 - Reliable	A - Known directly
Yes	B - Mostly reliable	2 - Known personally to the source but not to the officer	1 - Reliable	C - Known indirectly
Yes	C - Sometimes reliable	3 - Not known personally to source but corroborated	1 - Reliable	B - Known indirectly but corroborated
No	D - Unreliable	4 - Cannot be judged	3 - Not reliable	D - Not known
No	E - Untested	5 - Suspected to be false	2 - Untested	E - Suspected to be false

IMS new and old intelligence grading system

The paper talks about the Serious Violence Strategy endorsing a "whole system approach" – will outcomes from the project be shared to partner agencies. If so, what level of information and for what purposes?

WMP is part of the West Midlands Violence Reduction Unit (VRU) which takes a collaborative regional approach to addressing violence, vulnerability and exploitation. This multi-sector body convenes a range of activities and initiatives designed to embed a 'public health' approach to violence across the system, underpinned by the conviction that 'violence is preventable, not inevitable'.

	<p>As part of this ethos in general, relevant information derived from WMP activity will be shared with partner agencies, using appropriate data sharing protocols. High level conclusions will be shared to assist with the development of regional strategic decisions. In addition, individual data may be shared, for example, where young people are identified as being at risk of being exploited by organised criminals, referrals would be made to agencies best placed to offer interventions such as mentoring.</p> <p><b>It should be noted however that it is not intended that findings from this analysis would be shared with anyone outside of WMP as it is aimed at feeding into relevant decision making regarding ongoing operations only.</b></p> <p>Information will also be shared with other law enforcement agencies such as the Regional Organised Crime Unit (ROCU) and the National Crime Agency (NCA) where networks are identified as working beyond the West Midlands region.</p> <p><a href="https://www.westmidlands-pcc.gov.uk/strategic-policing-crime-board/agendas-minutes-reports/">https://www.westmidlands-pcc.gov.uk/strategic-policing-crime-board/agendas-minutes-reports/</a> (see paper in November 2019)</p>
I note the lack of legal advice	All proposed projects have been reviewed by Staffordshire and West Midlands Police Joint Legal Services. Their advice was made available to the Committee within the papers of the last meeting.
Not entirely sure what the operational outcomes are to be	<p>The SOC Network analysis will feed directly into the SOCEX Intelligence Hub (Serious Organised Crime and Exploitation) and contribute to the intelligence picture of each network. The Hub will use this to assess the most effective strategies for dismantling the networks which cause the greatest harm.</p> <p>The operational outcomes are envisaged to include the recovery of firearms, drugs and assets from criminals; as well as the identification and safeguarding of people vulnerable to exploitation.</p>
The reference here to partner agencies and to joint agency working prompts me to broaden that “debate” to the proposal that there might be much mutual benefit in approaching and thereafter liaising with Professor Tim Grant and others of	The SOC Network Analysis focuses on criminal networks that operate physically within the West Midlands area. Whilst some of the nominals may also engage in online

<p>Aston University's Institute for Forensic Linguistics. I attended an all-day Symposium run by the Institute on Friday sixth of March just gone.</p> <p>The especial focus with the launch of a study co-authored by Grant and Doctor Nikki McLeod on the subject of a project focused upon the impersonation by specialist and thoroughly trained undercover police officers of aspirant members of the extensive cabals of sexual predators. In crude terms this complex undertaking has to do with entrapment by those impersonators being convincing enough to be able to infiltrate and then enable the arrest and prosecution of those predators. A highly significant network of those predators is the Pedo Support Community ("PSC") with an estimated membership of four thousand.</p> <p>Others involved include Doctor Andrea Nini of the University of Manchester and a senior academic from Swansea University. I am working from my trigger notes taken during the Symposium only but I have ready access to more materials and references should they be of interest.</p> <p>My reason for raising this strand of inquiry is that part of the underlying remit of these academic undertakings is the building up of offender profiles and I might well be that the potential for overlaps with this important undertaking by the West Midlands Police might well be identified.</p> <p>In parenthesis, I ought to add that my informal contacts with the Institute arise from my wife Doctor Yvonne Fowler's academic and teaching roles with this particular alma mater of hers. In further parenthesis, I am aware that Tom Sorell of this Committee also has close links with Aston University.</p>	<p>sexual exploitation of children, that is not the primary focus of this project (however, see the related CSE paper).</p> <p>The Online Child Sexual Exploitation Team (OCSET) investigates online paedophile activity and works closely with the Child Exploitation and Online Protection Centre (CEOP) which is part of the National Crime Agency (NCA).</p>
<p>Proposed methodology: Please might I learn what the phrase "edge values" means?</p>	<p>The edge values are related to the number of times two nominals are connected. For example if Person A and Person B are linked by 3 intelligence logs and 2 crime reports then then they would have a link with an edge value of 5.</p>

<p>The paper mentions 'gangs' so I'd suggest that we need clarification as to what categories of data/intelligence this analysis will be using i.e. limited to criminal activity? 'Gangs' could also refer to non-criminal activity.</p> <p>What definitions are being used?</p>	<p>WMP uses the following definitions of 'gangs':</p> <p>An <b>Organised Crime Group (OCG)</b> is defined as having 'individuals, normally working with others, with the intent and capability to commit serious crime on a continuing basis, which includes elements of: planning / control / coordination / structure / group decision-making'.</p> <p>An <b>Urban Street Gang (USG)</b> is defined as being relatively durable, predominantly street-based group of young people (17 and under, although the ages of relevant nominals varies) who see themselves (and are seen by others) as a discernible group, and engage in a range of criminal activity and violence. They also have at least one of the following features:</p> <ul style="list-style-type: none"> <li>• Identify with or lay claim over territory</li> <li>• Have some form of identifying structural feature</li> <li>• Are in conflict with other, similar, gangs.</li> </ul> <p>The focus of the SOC Network analysis is towards the OCG end of the spectrum. However, there is likely to be some overlap between these definitions, especially when analysing the development of a network over time.</p> <p>It is also of note that USGs often become OCGs</p>
<p>What are the links between gangs and criminal violence?</p>	<p>The NCA states that, '<i>SOC affects more UK citizens, more often, than any other national security threat</i>' (National Strategic Assessment of SOC 2019). The NCA assesses that:</p> <ul style="list-style-type: none"> <li>• There remains a strong connection between drugs supply and firearms use and recovery, with firearms used to protect and enable criminal interests. Firearms are regularly found at drugs incidents and alongside drug seizures.</li> <li>• Communities across England, Wales, Scotland and Northern Ireland continue to feel the impact of SOC-related violence. Victims may suffer directly at the hands of criminals, or indirectly, finding themselves caught in the cross-hairs of inter-OCG violence. County lines drugs supply involves the supply of Class A</li> </ul>

	<p>drugs from urban hubs to country towns, and continues to involve the use of violence and exploitation of children and vulnerable adults.</p> <ul style="list-style-type: none"> <li>Beyond this, growing violence has been observed across other crime types, such as organised acquisitive crime. Instead of avoiding contact, offenders in areas such as car key crime are increasingly seen confronting victims directly, thereby posing an increased physical risk.</li> </ul> <p><a href="https://nationalcrimeagency.gov.uk/who-we-are/publications/296-national-strategic-assessment-of-serious-organised-crime-2019/file">https://nationalcrimeagency.gov.uk/who-we-are/publications/296-national-strategic-assessment-of-serious-organised-crime-2019/file</a></p>
<p>I think it needs to be made very clear what happens to the outputs e.g. if it could result in the categorisation of someone as potentially involved in criminal activity even though they have not previously committed a crime, then this needs further thought.</p>	<p>Where people are linked to a SOC Network through an intelligence log; but do not appear as an offender in our crimes data because they have not previously committed a crime, they will not feature as prominently in the output (they also would have no harm scores attached to them).</p> <p>The aim of the analysis is to identify the individuals who have the greatest influence and cause the most harm.</p>
<p>What is the intention and safeguards in terms of business services picked up in this analysis, in particular solicitors, doctors etc?</p>	<p>Where businesses or professionals are identified as potentially involved in or facilitating SOC they will be investigated according to normal procedures.</p>
<p>Question: why is the data not categorised?</p>	<p>Categorising the data, for example by age or ethnicity, is not applicable for this type of analysis.</p>
<p>Would be helpful to understand a bit more about when someone is classed as being part of a serious organised crime group or gang – where is the dividing line? Essentially, an ethical issue could arise if the research inadvertently focussed</p>	<p>Please see the definitions of OCGs and USGs above. Over time, some USGs do transition into OCGs and equally, some members of USGs go on to become members of networks involved in more serious criminality.</p>

<p>on vulnerable people rather than genuinely seasoned offenders</p>	<p>It is acknowledged that some nominals linked to SOC networks will have been both victims and offenders at some point; and that some are vulnerable individuals who are being exploited and forced to commit crime.</p> <p>WMP's approach is to consider violent crime holistically in partnership with the VRU collaboration. The SOCEX Intelligence Hub is set up to focus on exploitation as well as SOC; and will feed knowledge into partnership safeguarding arrangements when appropriate.</p>
<p>Does this link up with CSE project, and, if so, how exactly?</p>	<p>The SOC and CSE projects are linked. We want to see where there is overlap between SOC networks and those involved in organised CSE.</p>
<p>What data indicates centrality in the network?</p>	<p>There are different measures of centrality which can be calculated from the structure of the network. For example, the degree of centrality is the number of other nominals a given nominal is connected to. The analysis will utilise a number of measures.</p>
<p>How is being central to a network different from having authority in a network? In communication networks someone could be a sort of messenger and receive lots of communication while being quite lowly in the network. Or is this led by intelligence reports bearing on place in a OC hierarchy?</p>	<p>An example of this could be a police unit:</p> <ul style="list-style-type: none"> <li>• The Inspector communicates to the sergeant, who then passes on messages to all the Constables.</li> <li>• In this case we know the Inspector has authority, but in this scenario they have very few connections (only the Sergeant), whereas the Sergeant has lots because they pass information on to the rest of the unit.</li> <li>• We're not attempting to identify authority as such; we are trying to identify nominals that are important in the context of the network. In this case that could be the Sergeant, if they were taken out of the unit how would it function without information being passed from the Inspector to the rest of the team. (In reality someone else may step in which is why sometimes other centrality measures will also be calculated).</li> </ul>

Advice from the Committee:

- Outcome b) – proceed with minor amendments (see below).
- The committee requests that the Lab clarifies its rationale for using both harm indexes mentioned, and considers whether the ONS index could be said to be fairer as it draws on actual sentencing information.
- The committee requests that the operational guidelines once drafted are returned to the committee for advice, and advises that these should deal with the rights of victims who have gone onto become offenders, and should contain a clear commitment (rather than an intention) regarding the sharing of the analysis as mentioned above.
- The committee requests confirmation of the data sharing arrangements that are in place with the prison service regarding the use of prison data.
- With regard the definition of gangs being used, the committee advises that more clarity is needed in respect to the focus on OCGs, rather than USGs who are mostly under 18, and the different forms of intervention that are intended to be deployed.
- In respect of vulnerable people, and individuals who have not committed a criminal offence, the committee requests further information as the project develops as to how a safeguarding approach will be taken in respect of such individuals who may appear within a network analysis.
- As the project proceeds, the committee requests further information as to the measures of centrality being used and which data are most indicative of centrality.

**Child Sexual Exploitation Network**

Below table shows WMP Responses to the Committee’s queries:

Questions & Notes From Committee	WMP DAL response
<p>What about working with the victims to enhance detection? What about supporting the victims?</p>	<p>Supporting vulnerable victims is a priority for WMP. Our Ambition Plan states that, ‘We will safeguard vulnerable victims &amp; witnesses and improve their confidence and satisfaction, embedding the Victims Code and working with the Victims’ Commission’.</p> <p>The Public Protection Department investigates CSE and where appropriate works with victims to investigate crimes. However, the right to privacy of these child victims means it may not always be appropriate to pursue lines of enquiry by working directly with them. The priority is to safeguard children who are identified as victims of CSE and we work closely with partner agencies in the Multi Agency Safeguarding Hub (MASH) to ensure that victims are supported. The <i>WMP Strategy for Tackling CSE (2017)</i> identified that there has been less focus on understanding offenders and locations. Therefore, the primary focus of this project is to make use of our data to identify offenders who are linked to organised CSE criminality. By</p>

	<p>understanding how to dismantle these networks we can prevent future victims being exploited.</p> <p><a href="https://west-midlands.police.uk/flysystem/public-sync/inline-files/Ambition_Plan_FINAL_0.pdf">https://west-midlands.police.uk/flysystem/public-sync/inline-files/Ambition_Plan_FINAL_0.pdf</a></p>
Briefing document for CSE very close to SOC briefing document.	The SOC and CSE projects are linked. We want to see where there is overlap between SOC networks and those involved in CSE.
Independent legal advice needed.	All proposed projects have been reviewed by Staffordshire and West Midlands Police Joint Legal Services. Their advice was previously made available to the Committee if required.
<p>Proof of usefulness of procedure should be demonstrated by trials first. As noted above:</p> <p>‘The project would aim to highlight the extent and membership of networks of individuals involved in CSE with the view to informing further intelligence related work. Other than direction of further intelligence work, no decisions would be based upon the findings from the project.’</p> <p>This again begs the obvious question. Decisions will have to be made eventually. Correct? So hedging bets in this way is unsatisfactory.</p>	<p>Please see the discussion in the introduction. The DAL project removes the necessity of manually reading all the records held on WMP systems and uses more advanced statistical methodology to assess the relative harm caused by each network; and to understand the nodes of centrality. This work can be achieved in a matter of minutes in comparison to weeks of work for an intelligence team.</p> <p>Any further requests based upon this work are possible; but these would be submitted to the Committee as a separate project.</p>
The term ‘Boyfriend Model’ has received significant criticism, specifically in excluding the abuse of young male victims. Recommendation for reviewing the terminology (for example ‘Relationship Model’).	Advice noted.
How accurate is the intelligence information?	Please see discussion of intelligence grading in the response to SOC Networks.
How will victim information be treated? (specifically, potential issues with the criminalisation of victims linked with other victims)	<p>There is potential for previously unknown victims to be identified by this analysis. It is also acknowledged that some nominals linked to CSE networks as ‘offenders’ will also have been victims and that an element of this exploitation may include forcing vulnerable individuals to act as facilitators.</p> <p>WMP’s approach is to consider violent crime holistically in partnership with the VRU collaboration. The SOCEX Intelligence Hub is set up to focus on exploitation as well as</p>

	SOC; and will feed knowledge into partnership safeguarding arrangements when appropriate.
Would the information obtained be used to direct investigations?	Please see discussion above. The aim of this project is to enhance the knowledge of the SOCEX Intelligence Hub, rather than to direct specific investigations.
The only thing that needs more clarity is how the intelligence will be sifted for reliability/credibility - some examples of how this will work would be reassuring and useful.	Please see discussion of intelligence grading in the response to SOC Networks. Only credible intelligence will be used in the analysis.
To what extent will the outcomes be shared with other partners?	High level conclusions will be shared to assist with the development of regional strategic decisions, or to feed into regional and national understanding of this type of criminality. In addition, individual data may be shared, for example, where young people are identified as being at risk of being exploited by organised criminals, safeguarding referrals would be made to the relevant agencies, however this would be in line with normal, current, policing activity and this project would not lead to individual level data being shared.
What is the intended operational use? Some concern that if the findings are based on intelligence logs therefore operational actions based on untested information	The aim of this project is to enhance the knowledge of the SOCEX Intelligence Hub, rather than to direct specific investigations. Please see discussion of intelligence grading in the response to SOC Networks. Only credible intelligence will be used in the analysis.
I note the lack of legal advice	All proposed projects have been reviewed by Staffordshire and West Midlands Police Joint Legal Services. Their advice was previously made available to the Committee.
The reference at its paragraph one to working with partner agencies prompts me to broaden the “debate” to include a brief description of the contributions to an all-day Symposium of Friday the sixth of this month.  This event had been offered by the Institute of Forensic Linguistics of Aston University.	The Public Protection Department and OCSET are the subject matter experts in relation to the issues discussed at this Symposium.  There is also the Evidence Based Practice Team which liaises with academic institutions and cascades learning to relevant departments.

In part it had been to launch a book co-authored by the Institute's head Professor Tim Grant and by an alumna of the Institute Doctor Nicci MacLeod, now of Northumbria University.

The subject of the book was that of the project run under the aegis of the Northumbrian Police Service and of the enabling of infiltration by specialist undercover police officers into the extensive cabals of child sexual predators by their masquerading convincingly as "fellow travellers".

Whilst that subject may seem specialised and a far cry from the West Midlands Police Service's project under consideration, a significant amount of effort has been put into building up profiles of such groupings and a prime example featuring in one of the sessions had been that of the notorious Pedo Support Community ("PSD").

I mention this event since it does seem plain that a considerable amount of work has been undertaken both nationally and internationally on the developing of offender profiles and, in the interests of perhaps comparing notes and avoiding as far as possible any "re-inventing of the wheel", I wonder if it might be of value and of mutual benefit for some contact to be made with the Institute with a view conceivably to liaison and more.

There were presentations on closely associated and well established research projects from of course members of the Institute but also from Manchester and Swansea Universities.

There was a further paper delivered by Matt Sutton, the Senior Manager of the National Intelligence Hub concerning the elaborate and painstaking investigation which had only

<p>recently led to the conviction and sentencing at the Crown Court sitting in Birmingham on the nineteenth of February 2018 of the serial sexual predator and cabal member Matthew Falder, a lecturer at the University of Birmingham in Edgbaston.</p> <p>I respectfully suggest that we might recommend that appropriate contact might be made by the West Midlands Police with the Institute in order to ascertain to what extent all parties engaged might be able to share endeavours. For all I know this may already be happening but I thought it better to mention the subject rather than merely to assume that all entities engaged on related tasks are fully aware of each other's' activities.</p> <p>There is surely an ethical dimension here in the sense that the obligation of all must be to maximise opportunities to engage with these complex and nefarious practices.</p>	
<p>Proposed methodology: Noted but what, please, is the meaning of "edge values"?</p>	<p>The edge values are related to the number of times two nominals are connected. For example if Person A and Person B are linked by 3 intelligence logs and 2 crime reports then then they would have a link with an edge value of 5.</p>
<p>Does 'having had involvement in CSE' mean that the persons in question have been proven to have engaged in such activity, or are they merely suspected of doing so?</p>	<p>The analysis will use both crime data, where an offence has been reported to the police; and credible intelligence logs – please see previous discussion of how this is determined.</p>
<p>Has thought been given to how WMP will ensure the identification of potentially new victims might arise and how this information will be used to follow through with an investigation or referral? For example, would WMP consider whether another young person was a victim if they were arrested with a known victim? There is an argument to say that a failure to actively look for a potential pattern of wider victimisation would be an ethical issue in its own right, i.e.</p>	<p>The <i>WMP Strategy for Tackling CSE (2017)</i> identified that there has been less focus on understanding offenders and locations. Therefore, the primary focus of this project is to use our data to identify offenders who are linked to organised CSE criminality. By understanding how to dismantle these networks we can prevent the future exploitation of children. However, there is potential for previously unknown victims to be identified by this analysis. It is also acknowledged that some nominals linked to CSE networks as 'offenders' will also have been victims and that an element of this exploitation may include forcing vulnerable individuals to act as facilitators.</p>

willingness to ignore potentially relevant information around actual levels of abuse	
How are the relationships between inappropriate relationships, boyfriend model, network/organised exploitation understood? For example, couldn't groomers also be engaged in the exchange of images with other groomers, meaning that some people fitting the BF model might also be in networks.	This project should assist in developing our understanding of how these exploitation models operate.
Diagram makes it appear as if suspected or predicted Section 15 (Sexual Offences Act) trigger inclusion in data. Is this right?	The analysis will use both crime data, where an offence has been reported to the police; and credible intelligence logs – please see previous discussion of how this is determined.
Are the networks worked out independently of criminal-to-criminal communications data	The analysis used crime data and credible intelligence logs. Communications data will not be used.
Is there enough data to justify inferences to network connections with high probability?	There is a large amount of data available for network analysis.

Advice from the Committee:

- Outcome b) – proceed with minor amendments (see below).
- It is recommended that the 'Boyfriend Model' terminology and any assumptions behind it are reviewed, and the conclusions returned to the committee.
- The committee recommends increased clarity as to the operational uses of the model, in particular around the sharing of individual data where required for safeguarding purposes. The committee can envisage some beneficial uses but these need to be specifically articulated.
- The committee requests that the operational guidelines once drafted are returned to the committee for advice, and advises that these should deal with the rights of victims.

## Community Tensions

Below table shows WMP Responses to the Committee's queries:

Questions & Notes From Committee	WMP DAL response
<p>I am not sure what the committee is being asked here. The paper is highly speculative in the sense that no algorithms have been proposed as such, no testing been undertaken, no ethical implications been explored. This seems to be a request for an approval for a project for which few precise details are provided.</p>	<p>This project is at the proposal stage and therefore the precise algorithms to be used (if any) have not been determined. These will be determined after the Exploratory Data Analysis (EDA) has been undertaken when the benefits and limitations of the data are better understood.</p>
<p>I am also concerned that the links being explored are likely to be very fuzzy – what do we mean by 'community tensions' and what do we mean by 'violence'? What violence? Why should community tensions be related to knife crime for example? Both input and output need to be defined more precisely.</p>	<p>Community tensions refer to any emerging issues between different groups within our communities, which have the potential to escalate into disorder and violence. We would like to see if the DAL can use information from intelligence logs that relate to community tension issues to particularly focus on our priority of reducing youth violence and the use of knives.</p>
<p>It would be useful to have further information on the processing of individual data?</p>	<p>Whilst it is likely that data relating to individuals will be processed as part of the project (notably in relation to crimes), the outputs from the project (including any predictions) would not relate to individuals. Therefore it is unlikely that the privacy of individuals would be affected to any greater degree than would occur in normal police activity.</p>
<p>It would be useful to understand the source/accuracy of intelligence logs?</p>	<p>See previous details regarding the grading of intelligence.</p>
<p>What interventions/results would follow predictions?</p>	<p>The purpose of this proposal is to enhance the work of the Intelligence Department's assessment of the likelihood of violent incidents occurring in the future. Project Guardian, whose remit is to reduce serious violence and knife crime amongst young people, would receive estimates of the probability of a violent incident occurring in specific locations as a result of this work to inform their intelligence picture.</p>

	This would contribute to resourcing decisions, such as where to focus patrols or engage in preventative work with partner agencies.
Potential biases associated with the model would require consideration.	<p>The community tension intelligence originates from a range of sources as described above. Only credible intelligence would be used in the model.</p> <p>The EDA phase would include processes to identify the presence of any bias and to ensure that no bias is built into the analyses / any resulting model.</p>
Specificity/accuracy of the model would require significant consideration.	If a predictive model can be built, its accuracy would be tested on separate data which had not been used in its building. This can be done using historic data – where we know what happened and would tell us if the model would have predicted events or not. If the model is not deemed to be accurate enough, then the project would not be pursued.
It is an important safeguard in relation to this proposal that it would only predict at a small spatial scale and not in relation to individuals.	This project is to assist with resourcing decisions and to provide estimates of the probability of a violent incident in specific locations. There is no intention to predict the future behaviour of specific individuals.
One thing, again, that needs more clarity is how the intelligence will be sifted for reliability/credibility - some examples of how this will work would be reassuring and useful.	See previous details regarding the grading of intelligence.
Additionally, serious consideration, and an explanation should be given to the Committee as to how this project could be and should be communicated to local communities, given the potential for stigmatising people residing/working/socialising in the areas deemed to be at high risk.	<p>There are a number of existing mechanisms for reporting back to our communities in order to enable them to scrutinise our decision making. These include the OPCC Strategic Policing and Crime Board (SPCB); IAGs on each geographical area and scrutiny panels for Stop and Search and Use of Force. These existing arrangements could be considered as the channel for communicating the output of this project.</p> <p>An alternative could be specific community focus groups facilitated by an independent organisation as was the case with the IOM model.</p>

<p>What is the criteria for labelling a log as related to “community tensions”? Are we assured that the labels are accurate and therefore the information being used to inform the project sound?</p>	<p>Intelligence officers receive training in the collation, evaluation and dissemination of intelligence logs. They use their professional judgment to assess whether a log relates to community tension. The continual engagement with the national process for reporting community tensions also informs their thinking.</p> <p>Only credible intelligence would be used in the project (please see earlier discussion).</p>
<p>I note the lack of legal advice</p>	<p>All proposed projects have been reviewed by Staffordshire and West Midlands Police Joint Legal Services. Their advice was previously made available to the Committee.</p>
<p>Not entirely sure what the operational outcomes are to be</p>	<p>The purpose of this proposal is to enhance the work of the Intelligence Department’s assessments of the likelihood of violent incidents occurring in the future.</p> <p>Project Guardian, whose remit is to reduce serious violence and knife crime amongst young people, would receive estimates of the probability of a violent incident in specific locations as a result of this work to inform their intelligence picture.</p> <p>This would contribute to resourcing decisions, such as where to focus patrols or engage in preventative work with partner agencies.</p>
<p>I am afraid that I am unable to grasp what a “separate test dataset” means in this context;</p>	<p>We might build the model using data that only relates to Wolverhampton. In order to test the accuracy of the model we need to see if it works just as well on any data set; so we might test it on a separate data set relating to Birmingham. Both these would be historic data sets so that we know whether or not the model accurately identified offending patterns.</p> <p>Only when the accuracy of this process has been assessed, would we be able to apply the model to current intelligence logs to predict future crime patterns.</p>
<p>nor do I understand the words “beta testing”</p>	<p>Beta testing is where we believe the model has been built and completed; but before it is made available to all end users it is tested by a few individuals who can provide feedback on how well it works for them as practitioners. Beta testing is considered the last stage of testing.</p>
<p>What, please, does the acronym “AUC” stand for?</p>	<p>AUC stands for ‘area under the curve’ of a ROC (Receiver Operating Characteristics) curve. Essentially, ROC is a probability curve and the AUC</p>

	tells us how good the model is at distinguishing between different groups within the data. This is a statistical test for the accuracy of the model that has been built.
I would like to learn more about the violence reduction project “Guardian”; I have never previously encountered it.	In 2019, the Home Office gave the force £7.62million in police surge funding with the mandate that it is to be used to reduce serious violence in public spaces, with a focus on reducing knife crimes among young people.  The force’s response has been to create a two year project called Project Guardian. See the introduction to the briefing paper for the Knife Crime project.  <a href="https://www.westmidlands-pcc.gov.uk/wp-content/uploads/2019/07/190618-SPCB-Item-9-Violence-Crime.pdf">https://www.westmidlands-pcc.gov.uk/wp-content/uploads/2019/07/190618-SPCB-Item-9-Violence-Crime.pdf</a>
I fear I have forgotten what “DCMS” stands for.	Department for Digital, Culture, Media and Sport <a href="https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework">https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework</a>
“trade-offs, etcetera”: I am unclear what the description “trade-offs” means in this context.	This is a question from the Algotcare framework. In this case, trade-off would principally relate to the ability to identify true positives and true negatives (sensitivity and specificity); whilst either can be increased, this is at the expense of reducing the other.
“post-implementation, etcetera”:  The nature of the checks described needs to be detailed. Might this be a stage for the independent and/or academic scrutiny I have identified in my preliminary observations above?  Accuracy and accuracy of the algorithm: See my comment as above.	If the model was automated, there would be a process for checking its accuracy; comparing the predictions against what we know actually happened. These would be statistical accuracy checks that would flag up a decline in accuracy which could then be investigated and rectified.  Once the model is built, the methodology paper would be available via the OPCC website for other data scientists to evaluate.
Again, unclear as to the categories of data/intelligence that would be analysed - community tensions could include non-criminal activity, but I	Community tensions refer to any emerging issues between different groups within our communities, which have the potential to escalate into disorder and violence. We would like to see if the DAL can use information from

<p>appreciate that some of this could be relevant to what might then happen in terms of escalating violence. Suggest this is explained in more detail.</p>	<p>intelligence logs that relate to community tension issues to particularly focus on our priority of reducing youth violence and the use of knives.</p>
<p>What exactly is meant by social tensions? Some examples of what you mean by this, and the kinds of logs you would expect to be included as part of the study, would be helpful.</p>	<p>Please see introduction</p>
<p>Will there be any analysis about who makes these kinds of intelligence logs, and therefore a consideration of any bias?</p>	<p>As discussed in the section describing how we handle intelligence; the source of the information is removed in order to protect their identity. However, Intelligence Officers grade the credibility of the source as part of this process. For example, where there is a known neighbour dispute and one neighbour continually submits malicious and unfounded reports about the other, this would be regarded as 'unreliable' and graded as such. Future recipients of this intelligence would know to treat it with caution, even though they do not know where the information originated from.</p>
<p>Will thought be given as to whether the community tension is a proxy for some other activity? What if it is police activity causing tension? That could create a self-fulfilling prophecy</p>	<p>Our assessments of community tension are made in conjunction with other streams of information such as the IAGs. The interaction between the police and different elements of the community is always considered as part of our assessments.</p>
<p>Would be helpful to see if links are identified in the exploratory exercise, so we could consider ethical issues on a more granular level should you have enough to consider a predictive model</p>	<p>If the proposed project is approved it would be submitted to the Ethics Committee again, in greater detail and with results for further discussion.</p>
<p>A predictive model could have the same issues of labelling and stigmatising certain communities as other predictive proposals around youth violence have raised? Could it result, for instance, in police officers targeting youths they know within particular areas that come up as flagged?</p>	<p>WMP's strategy for reducing youth violence and knife crime is to work with partners and engage in a 'public health' approach. Any results from this project would feed into this over-arching strategy.</p> <p>Existing arrangements, such as the Stop and Search Scrutiny Panels would be able to consider whether certain areas were being unfairly targeted.</p>
<p>Will any assessment be conducted comparing any patterns identified or predictions made, with what police officers already suspected, i.e. the value added or not by this model?</p>	<p>If a model can be built it will shared with Subject Matter Experts (SMEs) for sense-checking.</p>

<p>Will there be a coordinated effort to connect this work with the surge team and the VRU team?</p>	<p>Yes, this will feed into the work being done by Project Guardian and the Violence Reductions Unit (VRU).</p>
<p>Can one explain how individual data is looked at?</p>	<p>Whilst it is likely that data relating to individuals will be processed as part of the project (notably in relation to crimes), the outputs from the project (including any predictions) would not relate to individuals. Therefore it is unlikely that the privacy of individuals would be affected to any greater degree than would occur in normal police activity.</p>
<p>What happens if there is no relevant community tension intelligent log in relation to an incident? Is there a risk this model disproportionately focusses on communities for which intelligence logs are more likely, perhaps due to other complex and sometimes historical social tensions regarding those communities?</p>	<p>Considerations of this type will be explored in the EDA phase of the project and used to determine its feasibility.</p>
<p>Does violence =knife crime? Is other crime included?</p>	<p>For this project we would be focusing on offence types reviewed by Project Guardian: Violence with Injury offences, robbery, threats to kill, attempted murder, homicide and possession of a firearm; in addition to knife related offences. This focuses on violent offences which are not related to domestic abuse and on those under the age of 25.</p>
<p>Are intelligence reports sufficiently standardised to be managed in a data science project? Is there automated analysis of intelligence reports, with semantic markers for references to “tensions”?  Or is this done manually? Might manual analysis be inconsistent, patchy?</p>	<p>See above re: the grading of intelligence. Logs are flagged by Intelligence Officers / Analysts on a manual basis.</p>
<p>Is this a supervised machine learning project?</p>	<p>If a model could be built, it would be a supervised model.</p>

Advice from the Committee:

- Outcome e) – further information required (see below).
- The committees advises the Lab to undertake an exploratory exercise, so that the ethical issues can be considered on a more granular level at future meetings should results indicate the potential for a predictive model. The committee advises the Lab to take into account the various questions raised by committee members in this exploratory stage.

### Domestic abuse

Below table shows WMP responses to the Committees queries:

Questions & Notes From Committee	WMP DAL response
<p>Again here there is no detail on the algorithm other than it will be geared towards improving domestic abuse prosecutions using a range of data. I think we need to know more about the details of what is proposed – modelling, testing and ethical implications. I also think – as I have remarked several times already – that there should really be less hedging of one’s bets. Either these algorithms are useful or they are not. If successful they will be useful in actual decision making and there is no point in being coy about this. Statements that keep coming up like ‘It is not intended that this becomes a means by which investigations are determined to proceed or not (or other forms of prioritisation).’ are not really at all useful.</p>	<p>The terms of reference requires the Data Analytics Lab (DAL) to submit briefing papers to outline proposed projects ‘in principle’ in order to highlight any immediate major issues prior to a project commencing. All the papers submitted to this meeting were such proposals for new projects.</p> <p>Subsequently, once the analyses have been completed, the projects are presented to the Committee again so that findings and methodology can be examined. An example of this was the RASSO findings presented at the last meeting in January.</p> <p>There is potentially merit in reducing this to just taking the findings following analyses (and methodology) to the Committee, as this would reduce the burden on members and allow them to make a more in-depth assessment when the findings are clear.</p> <p>However, this would require a change to the terms of reference.</p>
<p>Can you explain if there are any other areas (beyond the age groups) of categorisation?</p>	<p>Any categorisation (feature engineering) would be undertaken during model build if it was useful for the creation of the model. This can only be</p>

	assessed as part of the EDA / model build. Some examples of potential features that based on categories are within the RASSO report.
Will the analysis/research contain information on suspected DA incidents?	Data relating to individual incidents of DA and their associated outcomes will be used to conduct the analysis. However, there will be no reference to individual cases in the output which will draw high level conclusions to inform WMP policy.
What is the definition of a 'successful conclusion'?	<p>Outcomes 1-10 are generally referred to as 'positive outcomes' (not all are relevant for DA so do not appear in the chart in the introduction – but see full list at the end of this document). These include a charge, caution or community resolution.</p> <p>In the case of DA, a 'successful conclusion' will vary according to the needs of each victim's circumstances; and a charge may not always be the desired outcome. However, the high rate of offences which result in Outcome 16, where the victim does not support prosecution does deserve attention.</p>
Can we have further information on the examination of 'outstanding offenders'?	Please see explanation in the introduction. The project aims to understand the impact that outstanding offenders have – for example how quickly we arrest them after the reported incident and the effect this might have on the subsequent actions of both the offender and the victim.
<p>There is a strong ethical driver to undertake this important work.</p> <p>Since this proposal builds on earlier RASSO analysis project works, there needs to be a similar focus on addressing how the use of victim data will result only in positive outcomes for victims, not the use of the explanatory findings to 'streamline' investigative policing work away from difficult-to-address cases/scenarios type by type, as the models built under this proposal reveal them to be.</p>	The aim of this project is to inform policy by identifying areas for improvement in the way we conduct DA investigations. It is not intended as a tool for determining whether or not particular investigations should be pursued.
The paper refers to a previous paper. For completeness and transparency, as well as showing what the committee have reached its decision based on,	The RASSO paper referred to and the minutes of the January 2020 Committee meeting are available here:

<p>I think it would be helpful to have the previous paper as an appendix or reproduce sections within this report.</p>	<p><a href="https://www.westmidlands-pcc.gov.uk/ethics-committee/ethics-committee-reports-and-minutes/">https://www.westmidlands-pcc.gov.uk/ethics-committee/ethics-committee-reports-and-minutes/</a></p>
<p>Whilst I appreciate that it cannot be answered with certainty, it would be useful to have examples of what changes to process might look like</p>	<p>Until the analysis is complete it is impossible to predict what policy changes would follow. However, the analysis primarily focuses on how we manage the issue of outstanding offenders; and what factors could improve the outcomes of our investigation for victims. Therefore, it is likely that any changes to our processes will focus on these two areas.</p>
<p>It is difficult to determine necessity based on this paper alone; it doesn't illustrate the extent of the issue or current operational challenges etc.</p>	<p>Please see the introduction.</p>
<p>"method chosen": More detail will, surely, be required as to the other methods employed at the appropriate juncture.</p>	<p>As with the RASSO project the methods employed will be discussed with the Ethics Committee once the analysis has been conducted.</p>
<p>Re-creating a model that 'examines the factors that contribute to or detract from the probability of making a charge in DA cases.' Is looking at in terms of 'probability' really what is being done here, or is it trying to find factors that impact on an ability to bring a successful prosecution?</p>	<p>The analyses would be similar to the previous RASSO project.</p>
<p>Also there could be cases where a prosecution was not possible, but the victim was safeguarded in another way.</p>	<p>As discussed above, it is acknowledged that a so called 'positive outcome' may not always be in the best interest of every victim and there are other outcomes which may meet their needs better. The Public Protection Department who investigate DA work closely with partner agencies via the MASH.</p>
<p>As with the sexual offences research, it's important to ensure no inadvertent impact of identifying features that make certain investigations harder to carry out – just because certain cases are harder to carry out, does not necessarily mean they should not receive as much focus – what safeguards will be put in place for this?</p>	<p>The aim of this project is to inform policy by identifying areas for improvement in the way we conduct DA investigations. It is not intended as a tool for determining whether or not particular investigations should be pursued.</p>
<p>It would help to have more detail about how this research will be carried out, as it is tricky to identify sensitivities without knowing more about how it will be done</p>	<p>Once the EDA has been undertaken the methods of research can be determined and shared with the Committee.</p>

Hampshire Police have prior work in this area, I believe. Marion will know the officer leading it. I heard of his work at a RUSI meeting.	The force is engaging with Hampshire Police and the College of Policing over some aspects of managing DA.
Does 'outstanding offender' mean 'conspicuous offender' or 'unapprehended offender'?	Please see introduction.
Are elements that add or detract from probability of making a charge in a DA case all connected with process? What about a victim's change of mind with regard to giving evidence? Is that a process element?	External events will affect a victim's decisions about whether or not to pursue a prosecution. The focus of this project is to understand the elements which are in our control – for example how long it takes us to arrest an offender – and the impact they might have.

#### Advice from the Committee:

- Outcome b) – proceed with minor amendments (see below).
- The committee advises the Lab to ensure that the aim of this project is to inform policy by identifying areas for improvement in the way WMP conducts DA investigations. WMP should ensure that it will not now or in the future be used as a tool for determining whether or not particular investigations should be pursued.
- It is understood that the Lab will return to the committee for further discussion once methods of research have been determined.

#### Knife Crime

Below table shows WMP responses to the Committees queries:

Questions & Notes From Committee	WMP DAL response
The aim of the project, to develop methods to reduce knife crime, is of course worthwhile. However, we are again not given sufficient detail on the models being proposed – except to be told it is to be a spatio-temporal model supplemented by univariate modelling. This is simply too vague as it is not specified what would go into the model, what the predictive output would be, what methods will be used etc etc. It is also for the ethicists to	<p>The terms of reference requires the Data Analytics Lab (DAL) to submit briefing papers to outline proposed projects 'in principle' in order to highlight any immediate major issues prior to a project commencing. All the papers submitted to this meeting were such proposals for new projects.</p> <p>Subsequently, once the analyses have been completed, the projects are presented to the Committee again so that findings and methodology can</p>

<p>determine what ethical issues are involved but these are not likely to be different from above.</p>	<p>be examined. An example of this was the RASSO findings presented at the last meeting in January.</p> <p>There is potentially merit in reducing this to just taking the findings following analyses (and methodology) to the Committee, as this would reduce the burden on members and allow them to make a more in-depth assessment when the findings are clear.</p> <p>However, this would require a change to the terms of reference.</p>
<p>“Individual data will be processed... there would essentially be no interference with the privacy of individuals” – it would be useful to have further information on the processing of individual data?</p>	<p>Crimes data relating to individuals (as offenders or victims) will be the basis of this analysis. However, this will be aggregated and no individual will be identifiable from the analysis.</p>
<p>What interventions would be used following the prediction of locations etc.? (for example, Use of stop-search? Targeting of specific individuals?)</p>	<p>The results of this analysis would be used to inform resource allocation via Project Guardian every 4 weeks, using a range of policing tactics as appropriate. The decision to use stop and search tactics may be partly influenced by this analysis alongside other intelligence. This analysis would not influence the targeting of specific individuals.</p>
<p>Potential biases associated with the model would require consideration.</p>	<p>The EDA phase would include processes to identify the presence of any bias and to ensure that no bias is built into the analyses / any resulting model.</p>
<p>Specificity/accuracy of the model would require significant consideration.</p>	<p>If a predictive model can be built, its accuracy would be tested on separate data which had not been used in building and training the predictions. This can be done using historic data – where we know what happened and would tell us if the model would have predicted events or not. If the model is not deemed to be accurate enough, then the project would not be pursued.</p>
<p>It is an important safeguard in relation to this proposal that it would only predict at a small spatial scale and not in relation to individuals.</p>	<p>Crimes data relating to individuals (as offenders or victims) will be the basis of this analysis. However, this will be aggregated and no individual will be identifiable from the analysis.</p>

<p>However, serious consideration, and an explanation should be given to the Committee as to how this project could be and should be communicated to local communities, given the potential for stigmatising people residing/working/socialising in the areas deemed to be at high risk.</p>	<p>There are a number of existing mechanisms for reporting back to our communities in order to enable them to scrutinise our decision making. These include the OPCC Strategic Policing and Crime Board (SPCB); IAGs on each geographical area and scrutiny panels for Stop and Search and Use of Force. These existing arrangements could be considered as the channel for communicating the output of this project.</p> <p>An alternative could be specific community focus groups facilitated by an independent organisation as was the case with the IOM model.</p>
<p>The detail in this paper appears quite vague and therefore it is hard to comment on. It would be helpful to understand more about the data that is being used (and how accurate it may be e.g. is intelligence being used, arrests without charge etc.)</p>	<p>The data input for this model is crime data, specifically offences where a knife was used causing injury. The key components of the data will be the location and time of the offence.</p>
<p>We also know that dealing with serious crime requires a partnership approach; will the information be shared with 3rd parties and does this create any risks?</p>	<p>WMP is part of the West Midlands Violence Reduction Unit (VRU) which takes a collaborative regional approach to addressing violence, vulnerability and exploitation. This multi-sector body convenes a range of activities and initiatives designed to embed a ‘public health’ approach to violence across the system, underpinned by the conviction that ‘violence is preventable, not inevitable’.</p> <p>Project Guardian is closely linked to the VRU which is the mechanism through which a partnership approach to violence is delivered.</p> <p>More information about the range of support via the VRU can be found in the paper presented to the SPCB in November 2019.</p> <p><a href="https://www.westmidlands-pcc.gov.uk/strategic-policing-crime-board/agendas-minutes-reports/">https://www.westmidlands-pcc.gov.uk/strategic-policing-crime-board/agendas-minutes-reports/</a></p>
<p>Proposed methodology: To comment usefully, I would need to understand better (or at all) the concepts of a “spatial-temporal model “and of” a separate univariate time.”</p>	<p>Spatio-temporal models use patterns evident over space as well as patterns evident through time – “everything is related to everything else, but near things are more related than distant things” (Tobler’s first law of geography). This is often true in time as well as over space.</p>

<p>What does “productionised” mean?</p>	<p>To ‘productionise’ means that once we are satisfied that the model works well, we would automate the process of providing predictions every 4 weeks.</p>
<p>The reference to the checking of accuracy leads me back to my remarks about independent/ academic appraisal.</p> <p>Noted throughout but under “Responsible”:</p> <p>At the stage of teasing out “biases” should there not be an element at this crucial stage of independent appraisal of the kind I have touched upon before?</p>	<p>If the model was automated, there would be a process for checking its accuracy; comparing the predictions against what we know actually happened. These would be statistical accuracy checks that would flag up a reduction in accuracy which could then be investigated.</p> <p>Once the model is built, the methodology paper would be available via the OPCC website for other data scientists to evaluate.</p>
<p>I’m unclear as to how this is a 'predictive' model - is this intended to find links such as if X &amp; Y happen in one location on Monday, then A &amp; B is likely to happen in another location on a Thursday? Or if certain organised crime activity happens in one location, it is also likely to happen in another location?</p>	<p>The project would only look at knife crime and principally use data regarding knife crime.</p>
<p>Is it possible that young people who used a knife under mitigating circumstances can be unfairly affected by this?</p>	<p>This model will not identify individuals, it will predict where and when knife related offences are more likely to occur. A desired outcome is that we should be policing ‘in the right place, at the right time’. Therefore, a young person using or carrying a knife is more likely to be identified as a result of policing activity. This would result in normal investigative procedures for which there are a range of outcomes.</p> <p>However, a key part of Project Guardian is to support prevention work which includes a new out of court disposals pilot which will focus on diverting rather than criminalising children and making the most of ‘teachable moments’ in custody to divert offenders from committing further crime.</p>
<p>Can you explain a little more about the variables being considered in this model and what you anticipate as having predictive value? I’m not sure I understand what exactly is being looked at here?</p>	<p>Using data regarding the location and date of knife crime to predict future knife crime.</p>

<p>Is this essentially a form of geographical hot sport predicting? If so, what assumptions are being made on establishing a model?</p>	<p>It is indeed a form of hot spotting. This has been undertaken by WMP for a number of years using past events to identify hot spots. This project would make predictions as to future hot spots.</p>
<p>Can we have some indication of the kinds of policing response that might follow from such predictions? The response given is key to understanding ethics; disproportionate police presence for instance might create community tensions, whereas findings used for a renewed business case for neighbourhood policing could be well received and sensitive for community/ethics concerns</p>	<p>The results of this analysis would be used to inform resource allocation via Project Guardian every 4 weeks, using a range of policing tactics as appropriate. The focus will be on prevention and enforcement activities in hotspot areas; in particular providing more capacity for neighbourhood officers to engage in preventative work with young people.</p> <p>Supporting the prevention work will be a new out of court disposals pilot which will focus on diverting rather than criminalising children and making the most of 'teachable moments' in custody to divert offenders from committing further crime.</p>
<p>Is this basically an attempt to predict future knife crime hotspot locations from location data or is it based on more informal intelligence info?</p>	<p>Yes, the aim is to predict where and when knife crime is more likely to occur, based on knife-related crime data. There is no intelligence data included in this model.</p>
<p>What are the independent variables in the model?</p>	<p>Any other features used in the model would likely relate to general crime numbers.</p>

Advice from the Committee:

- Outcome e) – further information required (see below).
- The Lab is asked to provide more details of its proposed model including what would go into the model, what the predictive output would be, what methods will be used, and the potential interventions that would follow including the proposed out-of-court disposal.
- The Lab is asked to clarify in particular the range of example objectives that they hope to develop or discover from the data in terms of the reasons behind certain hot-spots occurring.
- The Lab is also asked to clarify the evidence of efficacy around this type of hot-spot policing.

## Home Office

### Retrospective Assisted Facial Recognition (FR) - Trial on Historic Criminal Case Data

#### Bias

Below table shows Home Office responses to the Committees queries:

Questions & Notes From Committee	Home Office Response
<p>In the past some facial recognition technology has had difficulties recognising people of colour, I just wondered whether this had been considered?</p>	<p>The issue of bias in algorithms is important and complex, and we agree that it should be considered thoroughly before any operational deployment of AFR. However, this is a tightly-controlled non-operational research study and seeks to address a different and possibly more fundamental gap in the publicly available evidence base. This study can be described as investigating “Does the technology work at all?” which logically must come before questions like “Does it work differently for different people?”.</p>
<p>Racial, gender or other bias:            -This is an integral issue raised in multiple ethics forums. Arguably the pilot is not fit for purpose if this isn’t an essential component of what is being tested: could result in more disproportionate interference with certain demographics; this would be unjust but might also harm public confidence in policing and give rise to legal issues            - In Cardiff’s research, will it not look at the potential difference in recognition between age?</p>	<p>However, while the trial design is focussed on achieving the key study aim (to compare the speed and accuracy of AFR reviews compared with an original manual review), we don’t want to miss the opportunity to gather any additional learnings on bias which could be achieved with minimal impact on the trial design – hence the intention to record possible indicators of bias.</p>
<p>How will indicators with bias be recorded/managed?</p>	<p>Researchers will record the following information about the person being sought: race, sex and approximate age. This data will be anonymised by WMP before sharing with HO researchers. In its anonymised form, it will be analysed by researchers in conjunction with the results of the research sessions, and may form part of the final report. How “race” is defined in this context will be determined by whatever information WMP is able to share.</p> <p>It is important to remember that – in accordance with data minimisation principles - this study will only look at the minimum number of cases</p>
<p>It is important that potential indicators of bias are to be recorded; but there should be a commitment to a follow up study to establish whether there is statistically significant bias, and if so in what manner, in the use of this approach. It is clear there is wider work on this area in the Home Office programme, but a clear commitment to this follow-up work needs to be made by the team responsible here.</p>	

	<p>required to achieve its key aim, which is to compare the speed and accuracy of AFR reviews compared with an original manual review. We believe we can gather this learning from a sample of 1-2 cases per force.</p> <p>This will give us a very small sample of target people to assess for indication of bias – probably a sample size of 2-4 target people. The characteristics of these people (race, sex, age) will be random depending on the cases selected. It is also not possible to control all the additional factors that will have an impact on accuracy: footage quality, lighting, camera angle, posture, face-coverings.</p> <p>The Cardiff academics will analyse factors which may have an impact on accuracy – race, age, sex, footage quality etc – in case it is possible to draw any interesting or useful findings, but due to the sample size these will not be conclusive and will be heavily caveated.</p> <p>The question of bias in algorithms is being considered by various groups across HO and policing, and it is not possible for the Innovation team of the Law Enforcement Portfolio to commit to being the team to conduct further work on bias. However, the team have worked with statisticians to develop an initial idea of the scale of a trial that might be required to investigate bias, and expect to use this to inform discussions with relevant parties. It is expected this would require filming test footage with actors, in order to control for all other factors and to include a sufficient quantity and range of characteristics.</p>
<p>I note for instance that the Glasgow face matching short test, which is stated as being used as part of the trial to test the human officers' abilities, contains no BAME faces (at least the test downloadable from this site does not <a href="http://www.facevar.com/glasgow-face-matching-test">http://www.facevar.com/glasgow-face-matching-test</a>), raising considerable questions about how appropriate it would be to use this test in the trial process</p>	<p>The proposed research study is not designed to investigate bias, and will not claim to draw solid conclusions on this point. However, where it is possible to potentially increase learning with minimum impact on the trial design, the ambition is to take these opportunities. It is in this spirit that the Glasgow Face Matching test is included as a method of assessing recognition ability. As noted, there are caveats about the design of the Glasgow Face Matching Test, and these caveats will be taken into account in the analysis and findings.</p>

<p>I venture to raise the additional factor of the age, sex and ethnicity of the reviewer.</p> <p>Since the proposal is to tease out any potential for discriminatory distortion of any matches reported, it is necessary, surely, as far as possible to explore all avenues which might lead to that distortion.</p>	<p>The member is correct that the characteristics of the reviewer are understood to have an impact on their recognition ability (for example, reviewers are more likely to be better at recognising people from their own ethnic group). Again this is another complex factor that would need to be investigated in a much larger trial specifically concerned with bias and with a greater number of participants and target people.</p>
---	---

### New Evidence

Below table shows Home Office responses to the Committees queries:

Questions & Notes From Committee	Home Office Response
<p>a) I am VERY troubled by this: ‘There is a chance that information that did not come to light during an original investigation may become available during the study as a result of processing footage through the tool. If this is believed to have occurred, the results of the trial would be sent to a senior officer in the force for them to determine if further action was required.’</p> <p>(b) Is there not scope for double jeopardy if the tool is used in this way? Comparison of AFR versus manual methods, if I understand correctly, at the present time is only being with a view to testing the algorithm under various criteria. How can it right to use such a comparison to re-open, for example, a closed case? Seems to infringe on all kinds of rights.</p> <p>(c) I am completely un-reassured by the following statement which appears in several places: ‘As the tool is being used for research purposes there is no risk of harm resulting from incorrectly spotting a person, as no action would be taken (excluding the scenario outlined under risk 4).’ I think the whole point is the tool, once implemented, will be used for non-research purposes and in such circumstances getting things wrong would be a serious matter. This is seeking comfort in places where there should be none.</p>	<p>We are acutely aware of the need to maintain and build public trust in the use of this technology and therefore it is essential that in the unlikely event that new evidence comes to light that there is a transparent and robust process in place to deal with such situations.</p> <p>The finalised details of the approach will be documented in an MOU which we will share with the Committee as soon as it is available. However we expect the process to be the following:</p> <p>In the event that any new evidence comes to light, whether that relates to the original offence or a new offence not previously identified, this will be brought to the attention of a review officer who is not a research participant in the trial. Having reviewed the CCTV footage in question, if the new material is considered to be of potentially material significance then the Senior Investigating Officer for the offence in question, or an officer of the equivalent rank, will be informed to consider whether the new evidence warrants further investigation. This is consistent with the decision making process utilised as part of the original investigation.</p>

<p>Use as evidence is mentioned in the trial protocol – how will this be assessed in line with the requirements around use of expert/scientific evidence/disclosure of material useful to the defence? As it is proposed that the output could be presented as evidence, how will the methodology behind the tool be assessed, and any uncertainties communicated? What will be the process for communicating to the defence relevant information regarding the FR search of the footage e.g. other individuals detected? Kotsoglou and Oswald (2020) argue that 'the means by which the identification took place must be disclosed to the defence, if Article 6 right to a fair trial is to be upheld, together with information regarding disregarded 'matches' and error rates and uncertainties of the system itself. Only then would the defence have the ability to effectively cross-examine the system via the expert witness, in the same way as they would cross-examine an eyewitness.'</p> <p><a href="https://www.sciencedirect.com/science/article/pii/S2589871X20300024">https://www.sciencedirect.com/science/article/pii/S2589871X20300024</a></p> <p>Face templates are stated as being deleted as soon as a result is returned – but should they be if there is an evidential context? I'd suggest that CPIA needs to be considered.</p> <ul style="list-style-type: none"> <li>• It was positive to see the question of when a case should be reopened addressed in the trial protocol. Regarding the research question 'What proportion of the possible matches presented to participants by AFR that were not documented in the original manual review were deemed correct by the majority of reviewers?', if any are generated, then presumably they would be referred for review using the process set out.</li> </ul>	<p>We have agreed with CPS that, where appropriate, they will then be consulted. In the event of the CCTV footage leading directly or indirectly to a future court case, its existence and how it came to light would be disclosed under the CPIA, whether submitted in evidence or as part of the unused material bundle. West Midlands Police and the Home Office would provide the necessary supporting material in accordance with CPIA and Article 6 as it would for all other forms of evidence.</p>
<p>Risks highlight that a case could be re-opened if a new line of enquiry is uncovered. Can we have further detail on the process for dealing with this situation?</p>	
<p>It is ethically important that closed cases can be re-opened if there are evidentially relevant matches.</p>	
<p>Reopening cases: if you found someone was potentially innocent and convicted, would you bring this to the relevant authority's attention?</p>	

## 55% probability threshold

Below table shows Home Office responses to the Committees queries:

Questions & Notes From Committee	Home Office Response
<p>I am struck and concerned to be reading that the “safe” percentage threshold for the exercise described should be as low as 55%. As I currently view it that proposition rather detracts from the purpose of the exercise and calls into doubt at least in mind the inherent reliability of the technology as at present operating. I am naturally open to being put right or receiving reassurance on this aspect.</p>	<p>The 55% confidence threshold has been set as a result of experimentation with test footage, however this is another factor into which this research study will start to give us more insight.</p> <p>It should be remembered that this is not “live” AFR, where a decision must be made swiftly under time pressure, and may result in interference with an innocent person. This research study replicates retrospective use of CCTV – investigators examining CCTV after a serious crime, in slow-time, trying to piece together evidence and identify investigative leads.</p>
<p>Isn't the 55% probability threshold too low?</p>	<p>Our current hypothesis is that when an investigator is reviewing footage, it is preferable for them be presented with a large number of possible matches for review, and therefore have an increased likelihood of finding any genuine matches. If the confidence threshold is set very high, the tool may discard genuine matches – which may not have a high confidence threshold for a variety of reasons, including footage quality, lighting or the face being partially hidden. It should be remembered that, in a real-life operational setting, police investigators would be able to bring in additional information to their assessment of these possible matches, such as recognising clothing, hair or carried objects.</p>

## Data Protection Act 2018

Below table is Home Office response on the Committees queries on the Data Protection Act 2018

Questions & Notes From Committee	Home Office Response
<p>Although data protection is mentioned, and an extract from the draft DPIA set out, it would be useful to explicitly state the justification under DPA 2018 for processing of the footage, and in particular in relation to the</p>	<p>The team commit to sharing the finalised DPIA for review, including justification under DPA 2018. This should be available to be shared by the end of June 2020.</p>

<p>biometric data of the individuals who are not of police interest (sensitive processing).</p>	
<p>An opinion on the legal position of using personal data (here without consent of data subjects) under Part 3 of the Data Protection Act 2018 is required for the Committee to give a more positive outcome from the one recommended below.</p>	

Below table are Home Office responses to the Committees queries on other topics:

Questions & Notes From Committee	Home Office Response
<p>The researchers outline the fact participants can withdraw during the data collection period - will a specific date for withdrawal be outlined clearly in the consent form?</p>	<p>The consent form will make clear that the intention is to publish a report using anonymised data within an estimated time frame, and that consent can be withdrawn prior to the end of that time frame.</p> <p>Participants will each take part in a single research session, lasting 2-3 hours. Prior to the session they will be given material explaining the activity, and asked whether they are content to participate. This will be confirmed by signing a consent form.</p> <p>Any personal identifiable data will be deleted within 90 days of the start of the research study, if not before. This information will only be known to the research team conducting the sessions. Participants can contact the team at any point after the session to request personal identifiable data is deleted immediately.</p> <p>Anonymised data – for example, anonymised quotes - will not automatically be deleted after 90 days, and will form part of the published research findings. Participants can contact the team at any time prior to publication to request their anonymised data be withdrawn and it will be removed from the report. After publication, it will not be possible to revise the report, but any internal records of anonymised data will be deleted on request.</p>

<p>When will risks on personal data identified in the Data Protection and Impact Assessment that is currently being drafted be available?</p>	<p>We estimate this will be available by end of June 2020. The Home Office team is currently engaged in drafting a DPIA with Kent Police (the other force participating in this research) and the ICO. When this is finalised, it will form the basis of the WMP and Home Office DPIA, as agreed with the WMP data protection team.</p>
<p>Assuming that the participants won't have details about the target like name, location, original offence?</p>	<p>It is correct that police participants in the research will not be given name, location and offence details. This information will also not be shared with the HO researchers.</p>
<p>Are the static images custody images? What happens when custody images were obtained long before the CCTV imagery? Won't this raise the probability of false positives and false negatives? Why are MP images used?</p>	<p>The static images will be whatever was available to the original manual reviewers – this could be a custody image. However it is likely the initial facial image may actually be a “screen grab” from the footage being analysed – that is to say, in the original review, the target was initially identified as an unknown person of interest within the footage, a static frame was taken from the footage for reference, and then the rest of the footage was reviewed attempting to track the movements of that target person.</p> <p>The impact of characteristics associated with the image are some of the many variable factors that we hope to gain more insight into in the course of this research study.</p> <p>The reference to Missing Persons is an error and will be updated – thank you for pointing this out. This study was initially expected to be conducted using footage from Missing Persons cases, but it was discovered that this footage is not generally retained long enough or documented in enough detail for the purposes of this research. However, it is possible that the research may include cases that started as a Missing Persons investigation and later became a serious criminal case.</p>
<p>The fact that ‘new manual reviews will not be conducted as part of this study’ is acknowledged in the study limitations. However the research question ‘Do reviewers using facial recognition technology come to the same conclusions as those who reviewed footage manually in a live</p>	<p>Home Office researchers have not yet had access to the relevant case documentation. From wider contextual research it is our understanding that viewing logs are unlikely to hold this kind of data, however the above advice is noted if this is available.</p>

<p>investigation?' may then miss the crucial question of how other (incorrect/uncertain) alternatives/matches were dealt with in the manual process. Regarding the original manual reviews 'Their success and time spent will be obtained from the original case viewing logs.' – however how far do the records disclose the full process that actually went on e.g. uncertainties about a face and discussions with colleagues etc.? Baseline should include any uncertainties from the original investigation.</p>	
<p>The trial refers to assessing whether the technology is 'useful'. Although there is probably minimal Article 8 interference in the trial scenario, there could be considerable in the operational context, as incorrect detection could result in police action being taken against individuals. Therefore, is 'usefulness' the correct standard? From an Article 8 perspective, 'necessity' and 'proportionality' should be considered. This then raises question of efficacy. What does review of 'performance' mean? How will it be assessed?</p>	<p>The word "usefulness" is used to indicate the potential benefit to policing, which would form part of the assessment of proportionality. Performance will be measured as specified in the "Outcomes" section of the protocol:</p> <ol style="list-style-type: none"> <li>1) <i>Compare reviewers' conclusions when reviewing CCTV to find people using facial recognition technology with reviewing the footage manually</i> <ul style="list-style-type: none"> <li>• <i>What proportion of sightings made by investigators in the manual review were also made by participants when using AFR?</i></li> <li>• <i>How many sightings did participants using AFR make that were not documented in the original manual review?</i></li> <li>• <i>What proportion of the possible matches presented to participants by AFR that were not documented in the original manual review were deemed correct by the majority of reviewers?</i></li> </ul> </li> <li>2) <i>Compare the speed at which investigators can review CCTV footage to find people using facial recognition technology, with manually reviewing footage</i> <ul style="list-style-type: none"> <li>• <b>Review time</b> - <i>The time from starting to review CCTV to having made a decision based on the results of the review, including any wait time during processing of footage.</i></li> </ul> </li> </ol> <p><i>This will be calculated by a researcher observing the participant completing the task and timing the episode from logging on to the system to the participant having made a decision on all the results presented to them. The data will be verified by cross referencing timings with system data extracted from the tool. This outcome seeks to understand if the technology helps investigators reach a conclusion about the footage quicker.</i></p>

	<ul style="list-style-type: none"> <li>• <b>Time on task</b> - <i>The time a participant spends interacting with the footage. This includes uploading footage and viewing footage and/or results provided by the tool. It excludes any wait time during processing.</i></li> </ul> <p><i>This will be calculated by a researcher observing the participant completing the task and measuring the time they interact with the system. The data will be verified by cross referencing timings with system data extracted from the tool. This outcome seeks to understand if the technology frees up investigator time to conduct other tasks.</i></p> <p><i>The data will then be evaluated against a reasonable best-case scenario. As footage is taken from historic investigations and investigators do not routinely record how long it takes to review CCTV, direct comparisons cannot be made. Previous research suggests that reviewing behaviour can vary considerably dependent on factors such as the quality of footage, level of activity in the footage, software functionality and severity of the crime<sup>1</sup>. However, it is reasonable to assume that it will take at least as long as the length of the footage to review footage in investigations that justify the review of large volumes of footage (e.g. serious crime or high-risk missing persons cases). Therefore, potential time saving benefits will be assessed by comparing time outcomes to the total duration of the footage.</i></p>
<p>The above leads to questions about what the police operator will be told about the result from system, in terms of its uncertainties etc? How has the threshold for a match been set? Will the trial analyse the influence of a machine tool on whether investigator is likely to accept the match as valid, even if it is uncertain? This may be something for a future trial but it would be useful to see these questions acknowledged.</p>	<p>An explanation of what “confidence rating” means is included on the results screen. The 55% confidence threshold has been set as a result of experimentation with test footage, however this is another factor into which this research study will start to give us more insight. We acknowledge human-computer interaction and the influence of this on reviewers’ decision-making is an important factor in the performance of AFR, and hope to investigate this further in future research.</p>
<p>I can see that the paper discusses ethics of the trial. However, my view is that the ethics of eventual use is essential to any trial itself – it should be one of the main questions of the trial or subsequent stages before roll out. Otherwise, without thinking these issues through now, the trial wouldn’t necessarily get one to a position of knowing whether it should be rolled-out</p>	<p>This is a tightly controlled research study, designed to have no operational impact as far as possible, with external academic consultancy and the intention to openly publish its findings to improve the publicly available evidence base. If there were to be further research studies or an actual operational trial, this would be treated as an entirely new consideration and accordingly be brought back to the Ethics Committee for scrutiny.</p>

<p>Error rates?: If often making errors, there is a risk that the harm caused by interfering with innocent people might happen too often for this technology to be considered proportionate – this needs discussion or at least highlighting as a key question of the trial</p>	<p>This study is an exercise in starting to understand accuracy in a realistic setting – with real case footage and real police users. There is no risk of interfering with innocent people, aside from the very small risk around uncovering new evidence, for which a process will be agreed as specified above.</p>
<p>Proportionality: What offences will this be used for? Prioritisation: What governance will there be to ensure the prioritisation of where this technology will be focussed/will not involve discrimination or bias? For instance, public assurance/commitment that this will be used for, for example, high level violent offenders would help allay fears that it would not be used to disproportionately police high volume but low level crime which might repeat some of the failures of the past of criminalising marginalised and poorer communities</p>	<p>The cases used in this study will be limited to serious crime cases, for example, murder or serious sexual assault.</p>
<p>Technical query: are we confident a full surveillance of the market for this technology has been undertaken? Accuracy is an ethical issue, and so understanding how this particular tech product was selected is helpful for giving assurance this is the most ethical option – simply the cheapest option would arguably be an ethical concern in its own right, if higher accuracy were forsaken</p>	<p>Limited internal testing has confirmed the supplier to be used for this research study is among market leaders in terms of accuracy for uncontrolled imagery. This testing is proportionate to the scale of the research study being undertaken.</p>
<p>Training data: the HO says that they are not aware of the training data as this was undertaken by third parties. I would expect HO to ask these questions themselves, as they give rise to important ethical issues, such as around discrimination. Would HO consider looking at this issue again?</p>	<p>Having access to this information would not change the design of this limited research study. However, we do agree that if there were to be further relevant activity, such as an operational deployment, this should be reconsidered.</p>
<p>Should we not be supplied with a draft of the Memorandum of Understanding when to hand?</p>	<p>This will be provided when ready. It is expected to be ready shortly after the DPIA by end of June 2020.</p>
<p>I fail to understand the description “technical research into the maturity of the facial recognition market”. Is it possible to elucidate?</p>	<p>This consists of desk research into the accuracy, cost, speed and scalability of a sample of facial recognition suppliers. In addition to this high-level theoretical research, we have conducted some practical research through software development and testing – for example, investigating how the tool might feed CCTV image frames to an algorithm in a faster and more efficient way.</p>

<p>I am unclear as to how the results of the projected exercise are to be fed into the planned trial; I wonder if that might be made somewhat clearer?</p>	<p>The word “trial” is perhaps misleading – the entirety of the proposal for this board’s review is the research study described in the protocol: 9 WMP participants using AFR technology to review footage prepared for them by researchers.</p>
<p>Trial Protocol: TBDs When and on what basis are these at present “missing” numbers to be decided?</p>	<p>The “to be decided” numbers in relation to number and length of footage clips cannot be confirmed until the Home Office researchers have actually had access to the proposed case footage and relevant documentation.</p>
<p>Trial Protocol: Under Analysis, please might I understand the meaning in this context of “interquartile”?</p>	<p>The interquartile range is a way of describing how "spread out" the results are. It is an alternative to standard deviation and it's somewhat better at describing the shape of the distribution.</p>

Advice from the Committee:

- Outcome e) – more information required (see below).
- The committee awaits further information in the form of the DPIA, the legal opinion and the Memorandum of Understanding relating to the possibility of uncovering new evidence.
- The Home Office is also requested to consider the following initial advice, which will be supplemented once further information is provided:
- The committee advises further investigation of the selected commercial tool’s performance in respect of all populations, including disclosure of details of how the tool was trained, and known biases/errors. The Home Office is advised to explore expanding the study in terms of footage analysed so that issues of bias do not have to be ‘heavily caveated’.
- The committee advises exploring using new manual reviews as part of the study which would help address the crucial question of how other (incorrect/uncertain) alternatives/matches were dealt with in the manual process; as the Home Office acknowledges, viewing logs are unlikely to hold this kind of data. Otherwise, the research would struggle to compare the two methods. Furthermore, the research protocol indicates that a major focus of the trial will be ‘review time’ and ‘time on task’. However, issues around false positives (particularly as the confidence threshold has been set at 55%), and how these are dealt with, are missing from the trial. Therefore, the trial would be unable to draw any conclusions around whether the technology ‘works’ in a wider sense.