THAMES VALLEY TOGETHER

Data Ethics Committee
26 April 2022
MS TEAMS

Present:

Mark Sheehan	(MS)	Chair of Ethics Committee
Lewis Prescott-Mayling	(LPM)	Inspector, TV VRU Data & Targeting Lead
Tori Olphin	(TO)	Data Scientist, TV VRU
Eric Twigg	(ET)	Ethics Committee
Tim Lowe	(TL)	TV VRU Researcher
Daniel Whiting	(DW)	Ethics Committee
Chris Lloyd	(CL)	Ethics Committee
Mark Warner	(MW)	Ethics Committee

Apologies:

Sylvia Simmonds	(SS)	Ethics Committee

Please note, when performing their duties as chair they will be referred to as the chair in the minutes. When commenting on the specifics of a use case that is presented they, like others, will be referred to as a committee member to sustain anonymity.

1) **Welcome – Chair**
   The chair welcomed everyone to the meeting.

   a) **Observers and conflicts of interest**
      There were no observers present and no conflicts of interest declared.

2) **Thames Valley Together Data Programme Presentation and committee questions – Lewis Prescott-Mayling**
   A copy of the presentation had been provided prior to the meeting. The presentation was shared on screen.

   - **LPM** gave an overview of the Thames Valley Together Data Programme to inform the committee of the data architecture that will be used to present other products to the committee.
   - A **Committee Member** commented that it was interesting that the data source came from a non UK government source. And yet, this was all off the back of a UK Government Strategy. **LPM** stated that the Centre for Disease Control (CDC) was used being the most complete and included world-wide data. **LPM** agreed that there would be some cultural differences and would not be without debate.
   - A **Committee Member** asked if any work had been carried out on the connection between the increasing rates of violence and society in general. **LPM** replied that it would require identifying cohorts of people, already involved in violence, for tertiary prevention. It would also require identifying people who are risk of becoming involved in violence and that would identify the primary prevention. This would be discussed later in this meeting. **A Committee Member** added that it raised questions on how social cohesion is measured.

They quoted from the presentation "violence is a public health issue". There will be health consequences of everything we do, however it does not make it a health issue. Focusing here on a public health model of carrying out policing. Keeping these two things separate seems really important. There appears to be confusion regarding the duty. **LPM** agreed and added that many police officers say this is the same as problem solving. However, these are just conceptual frameworks for identifying a problem, identifying tactical solutions where a public health approach does add value as it talks about understanding risk and protective factors in people rather than place based approaches. Fundamentally, within a proper public health approach there are two things, understanding the problem for a commissioned service, and then use opportunities to interact. The duty states that each Community Safety Partnership, or where there is a Violence Reduction Unit, must publish a public health violence reduction plan. The plan must be delivered over the coming year, and part of that is data sharing. The plan must be informed by risk and protective factors for violence.

3) New Phase One Presentation

a) Reminder of phase one content/focus from ToR

TL stated that the first stage is early stage consideration and process scoping. To consider the broad ethical questions and problems that might present themselves in the use case of high level initial stage. The second is to consider what extra things the committee might want to involve in the process of making a recommendation. This could be community members or extra expertise. The Committee needs to consider that process, a mixture of ethical questions and process of making that recommendation or decision. **The Chair** reiterated the two processes. One to get a general sense of the ethical issues that are involved, and secondly to think about what the process is for gathering more evidence and engaging with broader stakeholders.

b) STAR- *Lewis Prescott-Mayling*

LPM shared the presentation on screen (also provided prior to meeting). ToR Step 1 presentation around violence contagion.

- **LPM** stated that the VRU had been award monies from STAR to fund a research project to see whether they could model the spread of violence using a social network methodology. Two full vetted researchers from the University College London were conducting this work. They are vetted to NPPV3 level, however the data they are working with is anonymised.
- The research was presented to the Thames Valley Together Programme Board and is still a proof of concept. The statistical analysis and modelling has been started. Part of the funding agreement was that the project was presented to the Data Ethics Committee regularly.
- The World Health Organisation declared, in 1996, that violence was a public health problem and in 2002, they said it was a preventable health problem. There are a number of studies which have been conducted, particularly in the US, focusing on the spread of violence and whether it is contagious like a disease. There are several statistical models, some of them around diffusion of benefits or diffusion of ideas. In summary, they have been able to model that within a group of people who has either been violent, or has had a violent act conducted against them. There is a particular pattern. That might be because

of retribution or because others have been exposed to that event, and are more likely to themselves be exposed to violence as a perpetrator or victim.

- One way of modelling this is similar to earthquake modelling which is when you have aftershocks within a network. The closer you are to a person, the more likely you may be to be the perpetrator or victim. You can model the effects of violence in the social network similar to how you may do with how wide a disease spreads.
- How this would work within a statistical model is that we would take the risk and protective factors and build them into the statistical models. There would be individual level factors including the history of the person themselves to perpetuate violence or how much violence they have been exposed to. You may have factors that are at relationship and community level. Do they have family members who they have offended with in the past, or have they grown up in an area that is exposed to high levels of crime.
- **LPM** shared a graphic onscreen showing the co offending networks.
- **LPM** continued stating that the data and risk factors that would feed into the model would change over time. The more information that is acquired about people would make it better to understand who co offends with who.
- If there is an assumption at this stage that we can model this kind of contagion of violence in social networks, what will we do with it? Ideally we will be able to identify people that are at risk, which may be that they are the same person, or there may be a large overlap between people that are victims and people that perpetrate. We know from our analysis that there is a lot of overlap between perpetrators and victims of particularly serious violence.
- There is a project in America, Cure Violence, which is not without some criticism, but basically looks to interrupt the spread of violence in the social networks by identifying people that are more risk and then targeting community outreach to those people to get them into either peer support programs, or other support that that person needs.
- We already use Community Navigators, usually people of lived experience to try and interrupt that spread of violence. The results of this work is positive. There are probably some ethical questions about how this approach is operationalised, which will be for this committee to discuss.
- With reference to Serious Violence Duty, the level of data being used would be level 3 as we would be talking about people potentially. Potentially, as this may be something we debate whether we identify people. Or should it be risk stratification, and that would be level 1. It may be useful enough to have the general level of risk in a network and just commission support services at community level?
- **LPM** proposed some starter questions for the committee.
    - Need to avoid stigmatising communities or individuals
    - Who has access to the data and what they do with it?
    - How do we analyse whether it is effective?

b) Questions and Comments from the Committee

- **A Committee Member** stated that there would be challenges around data sharing and how the individual level of support can be framed. **LPM** replied that some work had been carried out using just police data and PCSOs have been used to visit the individual and offer support. Stating that they know who they are associating with and asking if they needed anything. This is not carried out anymore, we were not comfortable about developing it to make it business as usual. However it became a proof of concept. Ideally it wouldn't be a police officer or member of police staff, but, as previously commented, it would be someone from the community or from a charity we have commissioned. They

will carry out one to one coaching and signposting to support regarding employment, education, and drug or alcohol misuse or services depending on their need.

- **A Committee Member** asked if there was a way forward to see what was being taken from those already rehearsed and real life empirical examples, from the work carried out with UCL. What is it from those cases studies that TVT were seeking to replicate? Regarding the TVP's mock-up of the Slutkin work, should the committee have sight of these ready outcomes? Would this help mapping? **LPM** replied 1) work carried out in the US was on statistical models, particularly around the view of networks. It is taking static views of the network and makes the assumption that everybody who knows each other on day X always knew each other. When you model the migration of violence across that network, from a date before day X, you will see an effect because it was a violent event that created the link between those two people. There are issues around how you time slice social networks to model this contagion. Random models work for groups of 20 – 30 people, not for groups of 1,000.000 people. 2) What is the intervention going to look like? It is equally an ethical challenge. **LPM** agreed with the **Committee Member** that TVT could present more to discuss what the intervention would look like.

  A committee member stated that it would be interesting to understand what the plan was to evaluate the effectiveness of the model plus the intervention. He agreed that it would be useful to understand the long terms effects of prior deployments. To understand what the interventions were in the US and what the long term consequences were, both positive and negative. The connections changing across time is an important issue. **LPM** commented that the first stage that has been funded is to develop the model. There is no funding to assess and develop the model. The model will be shared and published and the code will be freely available. A presentation could be made of the previous research. However, the question is how do we evaluate? There is not a plan at present. We could evaluate in certain areas, identified those at risk, but do not intervene because there was a potential backfire effect of a PCSO knocking on their door. It needs to be planned.

  A committee member commented that it was a trade-off between having an effective intervention and reducing the amount of resources required, but also a risk of intrusion on identified individuals.

  LPM added that there will be statistical analysis if the model works, where the intervention is targeted.

  TO commented that if two people have been identified as part of a network, they then become known to police and therefore are more likely to be policed and therefore more likely to have bad outcomes as a result. If someone gets arrested for having a bit of weed on them because they were identified by us as being part of a network and police visit that person, that person is more likely to have bad outcomes in future. It is a problem that comes up a lot in hotspot policing. Community navigators are a good way of dealing with this, then the community do not realise that there are being intervened with. It can be blinded from the enforcement, which can make it more ethical.

  TO commented that there was a risk to having an open model. Individuals can direct their behaviour around the model if they know how it works. In addition, if a model is being retrained, it is possible to flood that model with data so that it is blinded and never picks up on a problem. Thought needs to be given around whether or not being completely

open around how the model works and exactly what feeds into it and how it links. This also applies to DARAT.

A committee member asked how it is determined when a police office is sent to a house, and if that could cause harm.

TO replied that if an officer attends someone's home who already had interaction with police, the officer is unlikely to nudge that person to have worse outcomes by them interacting with the police again. The difference of a first offence, first interaction with police is the difference that it can make, especially in hotspots.

LPM agreed with **TO**s comments. He confirmed that he did not intend suggesting that all of the algorithm was disclosed. The issue of drag netting rather than safety netting, could make it worse for the person. In the first instance, the resource and an evaluation strategy is needed in place to test the different models. Which could be 1) send a community member 2) send a Community Navigator and 3) send a police officer. To separate out from the confounding variables, a large enough experiment would be needed.

A committee member stated that this was consistent with their point around a public health model versus health. They invited LPM to speak further around the contagion metaphor. In addition to the contagion metaphor, there was also an aftershocks metaphor. How do these metaphors shape how thought is given to the models and responding interventions that may make a difference.

LPM stated that people are infected by a common source idea. Growing up in a high crime environment, they are exposed to the same sort of environmental factors. He added that because people are attracted to likeminded people and so are more likely to experience environmental factors that lead to violence as a result. LPM commented on research carried out regarding the contagion of smoking, that people are attracted to other people because they already smoke. However, this research is not robust regarding criminology around separating out the causal pathways. However, the contagion could be modelled with a statistical model, but how accurate will it be and does it inform the theory.

A committee member stated being open was important, and added, regarding **LPM**s last presentation slide, there was a thought around qualitative evidence or work which would unpack relationships between people and their neighbours and friends. This seems important for how the model is applied. In addition **the committee member** added that there was some evidence that the idea that violence is contagious and asked if there was scope for allowing that evidence to inform both the modelling and the interventions. The developed model is continuous, more data is fed into it and individuals and groups are identified within society. There is some value of learning from a single point in time to inform strategies not just within the police but for external agencies.

LPM stated that the model was a way off generating an operational output and a list of people to be intervened with. It has not been agreed by the VRU Strategic Board, who own the data, and there are other barriers that could stop it happening. **LPM** agreed with **the committee member** around a static point in time and asked how these people would be engaged with, younger people, young adults and late teens, people disproportionately

affected by violence. Engaging to ascertain what the intervention may look like and where the gaps are in the data.

A Committee member commented that the point around metaphors that were being used and how that is shaping this thinking, the idea of contagion, or is there an alternative? Which sort of metaphor this aligns to would have a significant impact on the type of intervention that would be applied or whether the need to use a machine learning model going forward.

LPM replied when the model is generated, with a contagion type variable, it will be a potential infection from a single source type variable relative model. We should separate out the different variable that may be more aligned to different theoretical pathways, this would be presented back to the Committee and you may request more community based work as it may be related to the environment and not people.

A Committee member asked if there was value trying to model on an individual basis and their families.

LPM replied that work would have to be carried out with individual data as that was how the model worked. However, the Committee may find that this data is not accurate enough and would be unethical. Predictive algorithms around individuals are quite accurate at a more ecological level.

TO added that people can have a lot of information about their past. There are clusters of risk factors and produce an idea of whether something might happen, and these can be relatively accurate.

A Committee member commented that the ethical use is linked with what eventually is done with it. This is not at the intervention stage, but is relevant as there are a lot of parallels being spoken about with what is already in mental health services. Identifying needs is often spoken about rather than the risk of someone committing a future offence is the same as having a need around that? How much of this kind of work and community policing is already being carried out without a model?

LPM replied that some risks factor will never be a need. However he agreed that the needs need to be identified otherwise the risk cannot be prevented. The Community Policing Team are so overwhelmed with work they are unable to continue interacting with people, unless they are specifically tasked to interact with someone.

A Committee member returned to the point of the difference between individual levels versus community level. They added that there was a difference between the effectiveness of a model at different levels and the effectiveness of a model plus intervention. This could change things quite significantly. There is a need to be mindful of how it integrates into existing practices and not just the effectiveness of the model. It needs to be thought about in combination. There is potentially a contagion effect in reverse.

LPM replied that the police currently engage at community level. The is more primary prevention than targeting and singling people out

- **Committee Requests**

LPM: Further work around engagement with young people and young adults. Understanding how they are going to manifest themselves within the data and how that will inform the modelling. Committee will be updated as to what that looks like and what results are achieved. This will also shape what is carried out in the longer term.

2) **Updates**

a) **Update on DARAT – Tori Olphin**

TO stated that work was ongoing getting DARAT built into the Cloud. Much further along than previously, however not at the point we had planned due to technical constraints. The next models built will be more fluent and fluid.

b) **Update on Compass – Lewis Prescott-Mayling**

LPM updated the committee that Paul Gresty will be the new Business Lead for the VRU and that Compass may be delayed slightly in how it moves forward, as there was a lot of work to carry out. LPM added that there was two elements to the Compass project, one was exploring social impact bonds with or without data, and the second was around building a model for unit cost calculating on events in the system. For example the cost of a missing persons, a victim of crime, an arrest, or to hold a Strategy Meeting. The second part having the more ethical implications. Compass is still in early stages, has not been scoped out or commissioned. The committee will be updated with the detail of how the unit cost calculator will work. When it has been scoped out.

A committee member questioned that the social impact bonds would not be a consideration for the committee, whether or not they contained data. **LPM** agreed as the data would evaluate the impact and would not involve Thames Valley Together data.

TO commented that regarding ethics and social impact bonds, often the resources that deliver the action are often the same ones that would deliver something else. This may make a difference to the measures and the unit cost calculator. The overall question may be whether it is value for money? And would be dependant where the funding is coming from.

3) **AOB**

The chair Suggested a topic for further discussion – the value of holding a live Data Ethics Committee meeting in public.

ACTION: LPM, MS and TL to discuss and present options at next meeting.

4) **Closing**

The chair thanked everyone for attending.