WMP Briefing Paper

Serious Violence Precursors

Ethics Committee (13 September 2023)

> This project is at the proposal stage and is presented to the committee 'in principle' so that any immediate concerns can be raised.
>
> The finer details of the methodology, exact data to be used and mode of communicating the results will not be determined until after the exploratory data analysis (EDA) phase has been undertaken.
>
> Once the analyses have been completed the project will be presented to the Committee again so that the data used, methodology, findings, intention for deployment and communication plans can be examined in more detail.
>
> Legal opinion has been sought and the Data Protection Impact Assessment (DPIA) is being reviewed by the Force Data Protection Officer (DPO).

Tasking

This project was requested by Chief Superintendent Pete Henrick in April 2023.

The aim of the project is to use police data as a starting point to identify the precursor events that cause an individual to be involved in serious violence (non-domestic abuse related).

The senior responsible officer is ACC Claire Bell, Local Policing.

Purpose

The purpose of the analysis is to enable West Midlands Police (WMP) to comply with the Serious Violence Duty (SVD), statutory guidance which requires relevant services to work together to prevent and reduce serious violence within their local communities.

The SVD acknowledges the importance of understanding and addressing the factors that cause someone to commit violent crime in the first place, including where coercion is a factor regarding vulnerable children and adults.

For WMP, the first step in building this multi-agency approach to prevention is to ensure that we are able to articulate the insight that can be derived from our own data sources about the likely precursors which contribute to an individual committing violent offences.

Note that this analysis relates to people who go on to commit non-domestic abuse related violence, but that domestic abuse will be considered as a potential precursor (for example as a child present in a domestic abuse setting).

Context

Serious violence has a devastating impact on the lives of victims and families, instils fear within communities and is extremely costly to society. Incidents of serious violence have increased in England and Wales since 2014. For example, offences involving knives or sharp instruments increased by 84 percent between the year to June 2014 and the year to June 2020.[1]

In July 2019, the Government announced that it would bring forward legislation introducing a new SVD on a range of specified authorities. This will ensure relevant services work together to share information and allow them to target their interventions, where possible through existing partnership structures, collaborate and plan to prevent and reduce serious violence within their local communities.

The Government also announced that it would amend the Crime and Disorder Act 1998 to ensure that serious violence is an explicit priority for Community Safety Partnerships (CSPs) and by making sure they have a strategy in place to explicitly tackle serious violence.   Local areas are encouraged to adopt the World Health Organisation's definition of a public health approach, which can be summarised as follows:

- Focussed on a defined population
- With and for communities
- Focussed on generating long term as well as short term solutions;
- Based on data and useful information to identify the burden on the population, including any inequalities
- Rooted in evidence of effectiveness to tackle the problem

A Strategic Needs Assessment (SNA) has been undertaken which provides baseline information about the volume of offences and individuals involved. However, there is insufficient information to identify and address the precursor events that lead individuals to become involved in serious violence, hence the need for this Data Analytics Lab (DAL) project. This is a statutory requirement and we need to ensure that we optimise the use of partnership and police resources by basing interventions on a firm evidence base.

Intended activity resulting from the project

This DAL project will support WMP to fulfil its obligation to contribute to the development of partnership strategies, governance arrangements and tactical plans to reduce serious violence through early intervention, which are to be published by January 2024.

The analysis produced by the DAL will be used to support all seven Community Safety Partnerships (CSPs) to develop their strategy and tactical plans with a sound understanding of the precursor events that can be identified within police data.

In addition, this analysis will ensure that Local Policing Areas (LPAs) understand the precursors to violence that can be identified in police data to ensure they undertake policing activity based on evidence. In particular, the analysis will include consideration of the effects of coercion, where vulnerable children and adults are concerned.

The specific policing activities that are likely to result from this analysis will be determined once the precursors are identified.

---

[1] Serious Violence Duty (accessible) - GOV.UK (www.gov.uk)

Ethical considerations

It is acknowledged that the analysis of precursor events in an individual's life may well occur before they appear in police data sources as either an offender or as a victim.

- Experience tells us that a substantial proportion of offenders who go on to commit serious violence do appear in police data early in their life, either as a youth offender, or as a victim, or both. Therefore, this analysis will ensure that we have extracted all possible insight from our own data to inform our policing activity as well as contributing to the development of the wider partnership collaboration. This analysis should be considered as the starting point for a wider partnership understanding of the impact of precursor events on future behaviour.

How will the results of this analysis affect individuals who are identified as having experienced any of the precursor events?

- Whilst data relating to individuals will be processed it will be used to develop an explanatory model to inform our approach to reducing violence when developing strategic and tactical plans. It is not intended that this analysis would be used for operational decision making about individual data subjects. This would occur once the local partnerships have developed their strategy, governance and tactical plans and when there is an evidence base of which interventions are effective.

The DAL has undertaken similar analysis in the past to understand the causes of violent crime.

- In 2020 the DAL undertook analysis of the *Factors that contribute to the transition to most serious violent crime by young people.*[2] This updated analysis will consider the recommendations of the Committee which required major amendments and will consider all age groups rather than focussing specifically on youth violence. It will be used for the specific purpose of supporting WMP's response to the SVD which is a statutory requirement.

Police data does not reflect protective factors which may offset the identified precursor events.

- For this reason, this analysis is the starting point for wider partnership work to understand how to prevent violence using evidence-based intervention strategies. Other agencies across the partnership will hold data reflecting protective factors which may counteract the identified precursor events.

---

[2] Ethics Committee Reports and Minutes - West Midlands Police & Crime Commissioner (westmidlands-pcc.gov.uk) January 2020 Agenda Item 4

Data

Data to be used:

- Connect (crimes, custody, intelligence)
- Compact (missing person data)
- Open source deprivation and census data such as populations (i.e. number of violent offenders in an area per 10,000 people)

Level of analysis:

☑ Individual
  Individuals aggregated?
  ☐ Yes
  ☑ No
☐ Specific Area:
  ☐ Output Areas
  ☑ Super Output Areas - Lower
  ☐ Super Output Areas - Mid
  ☐ Wards
  ☐ Districts
☐ West Midlands
☐ Other

EDA will look at (aggregate) data at a neighbourhood level

Home location (first and latest offence) (Neighbourhood/sector) is likely to be in the final model (possibly replaced by other geography-based metric such as deprivation).

Reliability of data:

The data are sourced from WMP systems. An extensive exploratory data analysis (EDA) phase will be undertaken to examine the extent of any data quality issues. Any data quality issues, such as missing values, outliers and potential biases will be noted and where applicable will be addressed.

Sample or entirety: Entirety

Type of analysis:

☐ Exploratory
☑ Explanatory
☐ Predictive
☐ Optimisation

Proposed (broad) methodology:

1. Serious violence (non-domestic abuse) based on Violence with Injury offences and taking into account Harm.
2. Undertake extensive EDA; incorporating both spatial (LPAs and LSOAs) – includes choropleth maps, analysis of global and local spatial autocorrelation as well as temporal information.
3. Likely to be some form of relaxed LASSO or Bayesian regression with regularising priors;
   - Investigate utility of splining quantitative variables, etc.
   - Keep any transformed quantitative variables as transformed quantitative
   - Include variables relating to coercion (county lines, modern day slavery, etc.)

4. Bootstrap the modelling analysis (if relaxed Lasso) to create confidence intervals / show credible intervals if using Bayesian
5. The final model needs to balance fit with explainability.

Include other modelling types to qualitatively assess findings of main analysis (e.g. DAG, single tree, GBM with var importance, etc.)

Will the project eventually be automated:

☐ Yes
☑ No

Means of evaluation:

The project will test other modelling types to qualitatively assess findings of main analysis (e.g. DAG, single tree, GBM with var importance, etc.)

ALGO-CARE considerations

As this project is at the proposal stage and is presented to the committee 'in principle' in order that any immediate concerns can be raised, the finer details of the methodology will not be determined until after the EDA. Once the analyses have been completed the projects will be presented to the Committee again so that findings and methodology can be examined in more detail.

Advisory	
If applicable, are the outputs from the algorithm to be used in an advisory capacity?	Yes, the findings of the analysis will be used to inform senior decision makers and partners as they develop strategic and tactical plans for reducing serious violence.
Does a human officer retain decision-making discretion?	This work will not determine case by case operational decisions, but will assist with the development of strategic and tactical partnership plans.
Lawful	
What is the policing purpose justifying the use of the algorithm (means and ends)?	The work supports WMP in its statutory obligation to work in partnership to reduce serious violence under the Serious Violence Duty.
Is the potential interference with the privacy of individuals necessary and proportionate for legitimate policing purposes?	Whilst data relating to individuals will be processed, it will be used to develop an explanatory model to inform our approach to reducing violence when developing strategic and tactical plans. It is not intended that this analysis would be used for operational decision making about individual data subjects. This would occur once the local partnerships have developed their strategy, governance and tactical plans and when there is an evidence base of which interventions are effective.
In what way will the tool improve the current system and is this demonstrable?	This analysis will provide the most up to date and comprehensive survey of precursor events which are likely to impact on future offending behaviour, as far as can be derived from WMP data. This robust statistical analysis will provide the evidence base from which police and partnership activity can be developed. This is intended to build on the baseline information provided by the SNA.
Are the data processed by the algorithm lawfully obtained, processed and retained,	The data are from WMP systems and collected to enable their normal day-to-day operations.

according to a genuine necessity with a rational connection to a policing aim?	
Is the operation of the tool compliant with national guidance?	The analyses proposed would accord with the Government Digital Service Data Ethics Framework 2020[3]
Granularity	
Does the algorithm make suggestions at a sufficient level of detail given its purpose and the nature of the data processed?	This explanatory model will provide information at a level sufficient for developing strategic and tactical plans to support our partnership responses to the SVD in the West Midlands.
Are data categorised to avoid broad-brush grouping and results and therefore issues of potential bias?	As the underlying data are at the level of individuals, there should be no categorisation.
Do the potential benefits outweigh any data quality uncertainties or gaps?	The aim of the project is to support WMP to work with partners under the SVD. The analysis will include an extensive EDA element and this should highlight areas of heightened uncertainty in the data or where particular gaps exist. Should any such issues be identified, these would be addressed as a part of the project. Given the benefits of reducing serious violence it is not expected that any data quality issues would be of such a magnitude as to warrant not undertaking the project.
Is the provenance and quality of the data sufficiently sound?	The data will be those that are gathered during the day-to-day work of WMP and will enable analyses of the type envisioned for this project.
If applicable, how often are the data to be refreshed?	Not applicable
If the tool takes a precautionary approach in setting trade-offs, what are the justifications for the approach taken?	Not applicable

---

[3] https://www.gov.uk/government/publications/data-ethics-framework

Ownership	
Who owns the algorithm and the data analysed?	WMP owns the model and the underlying data.
Does WMP need rights to access, use and amend the source code and data?	No
Are there any contractual or other restrictions which might limit accountability or evaluation?	No
How is the operation of the algorithm kept secure?	The analyses developed would be kept wholly within the secure WMP computing environment.
Challenge	
What are the post-implementation oversight and audit mechanisms, e.g. to identify any bias?	Not applicable
If the algorithm is to inform criminal justice disposals, how are individuals notified of its use?	Not applicable
Accuracy	
Does the specification of the algorithm match the policing aim and decision policy?	The nature of the analyses initially chosen have been determined to be the best means of addressing the research question.
Can the accuracy of the algorithm be validated periodically?	Not applicable, this is not a predictive model.
Can the percentage of false positives / negatives be justified?	Not applicable, this is not a predictive model.

How was the method chosen as opposed to other available methods?	Due to the nature of the research question (explanation as opposed to prediction), the envisioned approach is appropriate.
What are the (potential) consequences of inaccurate forecasts?	Not applicable, this is not a predictive model.
Does this represent an acceptable risk?	Not applicable, this is not a predictive model.
How are the results checked for accuracy and how is historic accuracy fed back into the algorithm for the future?	Not applicable, this is not a predictive model.
How would inaccurate or out-of-date data affect the result?	If data were to be wholly inaccurate then the analyses would essentially provide inapplicable findings. The DAL will seek to minimise this potential through a thorough analysis of the data in an extended EDA phase including their pitfalls, issues and overall nature. Discussions with subject matter experts (SMEs) should also highlight effects that may arise due to erroneous data.
Responsible	
Would the operation of the algorithm be considered fair?	The analyses will be fair in that each data point will be considered on its own merits.
Is the use of the algorithm transparent (taking account of the context of its use), accountable and placed under review?	The nature of the intended method(s) is such that the end-use is to provide information rather than predictions that feed into a process on an on-going basis.
Would it be considered to be used in the public interest and to be ethical?	This project supports WMP to fulfil its statutory requirements under the SVD. Therefore, it would be considered to be in the public interest and to be ethical.
Explainable	
Is information available about the algorithm / decision-making rules and the impact of each feature?	The nature of the envisioned methods is such that information regarding the features and how they relate to each other is capable of explanation.

Appendix 1: Glossary of Terms

WMP / Law Enforcement Terminology	
CSP	Community Safety Partnership
LPA	Local Policing Area
SME	Subject Matter Expert
SVD	Serious Violence Duty
SNA	Strategic Needs Assessment
WMP	West Midlands Police

Data Science Terminology	
ALGO-CARE	All projects have used the ALGO-CARE to consider ethical implications: Advisory, Lawful, Granularity, Ownership, Challenge, Accuracy, Responsible, Explainable
EDA	Exploratory Data Analysis