Ethics Committee Briefing Note

Project Reference: DAL_2019_0009_Community Tensions & Violence

Purpose of data analysis:

The last few years (notably since early 2015) have seen an increase in the number of violent incidents, especially in the use of knives. At the same time, the average age of nominals involved in violent incidents has been decreasing. This is a situation that has been seen throughout the country. Indeed the Government announced extra funds to be made available to Police forces to help tackle violent crime amongst young people and knife crime issues. WMP currently has a strategic objective to reduce serious violence amongst under 25 year olds by 10%.

Anecdotal suggests that intelligence reports can be related to incidents of violence at some later point in time. This project aims to ascertain if there is evidence of intelligence reports relating to community tensions at one point in time that can be related to incidents of violence at a later point in time. If this is found to be the case, then the project would aim to develop a model that uses intelligence reports of community tensions to estimate the likelihood of ensuing violence (i.e. can intelligence reports be used to make predictions of violent incidents at later points in time).

As a part of this, the project would aim to:

1. Delineate a means of identifying reports of community tensions within locales / within different communities (not necessarily geographic).

2. Create a means of scoring the severity of any such community tensions.

3. Identify those acts of violence associated with intelligence reports of violence that were received prior to the violence.

4. Ascertain the average time between reports of tensions and the resulting crime(s).

5. Ascertain whether the severity of those tensions increase towards the eventual crime(s).

6. Develop a means of estimating the probability of violence arising from intelligence reports of community tensions.

Source of analytical question / hypotheses to be examined:

The business question was posed by the Intelligence Department.

Data to be used:

Level of analysis:

☑ Individual
   Individuals aggregated?
   ☑ Yes
   ☐ No
☑Specific Area:

☐ Output Areas
☐ Super Output Areas - Lower
☐ Super Output Areas - Mid
☐ Wards
☐ Districts
☑ Other

☐ West Midlands
☐ Other

Reliability of data:

An extensive exploratory data analysis (EDA) phase will be undertaken to examine the extent of any data quality issues.

Sample or entirety:

If sample: N/A

Method of sampling: N/A

Method of choosing sample size: N/A

Sample size: N/A

Type of analysis:

☑ Exploratory
☐ Explanatory
☑ Predictive
☐ Optimisation

Proposed methodology:

Using an initial sift of the intelligence reports under the heading of community tensions, map these in relation to crimes over time. This covers the period 2013 to the present. Through specific examples, isolate crimes within various radii of the locations mentioned in the intel. reports over various time periods and assess whether there appear to be links. If there is evidence that there are links, develop a model to predict any likely violent incidents following reports of community tensions. More explicit methods are not as yet known, however this project will require an extensive exploratory data analysis (EDA) stage. This will be assessed following review of the EDA stage.

Will the project eventually be automated:

☑ Yes
☐ No

If links between intel. reports and episodes of violence are established and if a model can be built that uses intel. reports to predict violent incidents, then this model would eventually be productionised.

Means of evaluation:

If a predictive model can be built, this would be tested on a separate test dataset (separate from the data used to train a model) and beta testing would be entered into prior to roll out in order to assess accuracy (accuracy, sensitivity, specificity, precision and AUC) on completely new data.

ALGO-CARE considerations:

Advisory:

If applicable, are the outputs from the algorithm to be used in an advisory capacity?

Findings from the initial EDA would be provided to subject matter experts (SMEs) for information / comment. If it is considered that a predictive model could be built, the eventual results from the model would be fed into the wider business, particularly the violence reduction project Guardian. It is likely that the outputs of any such model would essentially be estimates of the probability of a violent incident in specific locations.

Does a human officer retain decision-making discretion?

Any decisions arising from the provision of any information resulting from this project would be subject to decision making by officers.

Lawful:

What is the policing purpose justifying the use of the algorithm (means and ends)?

The reduction of violence is a key aim of WMP and a tool that could be invaluable in this regard could feed into operational decisions to help achieve this aim.

Is the potential interference with the privacy of individuals necessary and proportionate for legitimate policing purposes?

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.

In what way will the tool improve the current system and is this demonstrable?

There is no current system relating reports of community tension to future potential episodes of violence. If such a system can be developed this would enable improved decision making in attempting to reduce violence.

Are the data processed by the algorithm lawfully obtained, processed and retained, according to a genuine necessity with a rational connection to a policing aim?

The data to be analysed are collected as part of WMP's duty to collect and assess intelligence and to investigate crimes and as such have been collected in the appropriate manner for the appropriate operational purposes.

Is the operation of the tool compliant with national guidance?

The analyses proposed would accord with DCMS Data Ethics Framework 2018.

Granularity:

Does the algorithm make suggestions at a sufficient level of detail given its purpose and the nature of the data processed?

It would be the aim of any model that is developed to provide predictions in such a manner as to be of use to WMP decision making – most likely at some small spatial scale. Details are currently unknown given the thorough EDA required to make an assessment as to the possibility of building a model.

Are data categorised to avoid broad-brush grouping and results and therefore issues of potential bias?

Any categorisation that is required will be undertaken so as to optimise the performance of any model and the decisions to be made therefrom. This would include processes to identify the presence of any bias and to ensure that no bias is built into the analyses / any resulting model.

Do the potential benefits outweigh any data quality uncertainties or gaps?

If potential violence could be identified via intel. reports of community tensions, then it is possible that processes could be put in train to try and reduce any such tensions as well as knowing where to allocate resources to try and offset incidents of violence. Should violence be reduced this would constitute large savings. Based on costings from the Home Office, incidents of most serious violence amongst those aged 25 years and under within the WMP area would have led to a total unit cost of circa £37.6m in 2017. Preventing 1% of incidents would equate to circa £376,000. Given that any predictive model would be tested for accuracy, if a reasonable degree of accuracy can be gained, it is likely that uncertainties within the data would be more than made up for by the potential benefits.

Is the provenance and quality of the data sufficiently sound?

The data have been gathered during the day-to-day investigative work of WMP and do enable analyses of the type envisioned for this project.

If applicable, how often are the data to be refreshed?

The underlying data are refreshed a number of times per day. Should a predictive model be productionised it is likely that it would be run once per day.

If the tool takes a precautionary approach in setting trade-offs, what are the justifications for the approach taken?

Ultimately any model developed would aim to maximise specificity whilst trying to gain as high a sensitivity as possible. This approach would mean that we could best allocate WMP resources whilst ensuring a minimisation of false positives.

Ownership:

Who owns the algorithm and the data analysed?

WMP once developed.

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?

Any model developed from this project would be kept wholly within the WMP computing environment.

Challenge:

What are the post-implementation oversight and audit mechanisms, e.g. to identify any bias?

Should any model be developed that is then productionised, checks will be made as to its accuracy on an on-going basis as well as any consistent patterns that may represent biases.

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?

Any resulting model would aim to produce information to aid decision making within WMP which would be in line with its aims and policies.

Can the accuracy of the algorithm be validated periodically?

The productionisation of any model resulting from the project would include checking its accuracy on an on-going basis.

Can the percentage of false positives / negatives be justified?

Currently unknown, however the sensitivity and specificity would be balanced so as to maximise its use for decision making whilst not leading to waste of WMP resources.

How was the method chosen as opposed to other available methods?

Not currently answerable – this will be dependent upon the EDA stage and its findings.

What are the (potential) consequences of inaccurate forecasts?

If a model could be developed, the main issues arising from inaccurate forecasts would be (a) potential for actions for WMP that may not be necessary (including in particular locations) and (b) WMP resources being allocated ineffectively.

Does this represent an acceptable risk?

Any model arising from this project would seek to balance the advantages against the risks arising from inaccurate predictions partly via balancing the model's sensitivity and specificity and partly through assessing the types of decisions for which any such model would be effective for and any actions that would arise from these decisions. This would be subject to periodic review.

How are the results checked for accuracy and how is historic accuracy fed back into the algorithm for the future?

If any model were to be productionised, accuracy would be assessed on an on-going basis via measuring its accuracy (sensitivity, specificity and AUC) as well as producing histograms of estimated probabilities so that any degradation of the model could be tracked and the model rebuilt if necessary.

How would inaccurate or out-of-date data affect the result?

This is partly dependent on the nature of any model should one be capable of being built. Generally inaccurate or out-of-date data could detrimentally impact on the model's performance (in terms of accuracy), however currently the potential for any such issues is unknown.

Responsible:

Would the operation of the algorithm be considered fair?

During the development of any model, the presence of any biases in the underlying data or for predictions to produce biases would be fully examined and mitigated if the potential for them was present.

Is the use of the algorithm transparent (taking account of the context of its use), accountable and placed under review?

The details of any model arising from this project would be provided and, as mentioned above, if productionised there would be on-going checks as to model performance.

Would it be considered to be used in the public interest and to be ethical?

The reduction of violence would be the aim of the project and given the cost to society of violence, enabling more effective means of its reduction would likely be seen as positive.

Explainable:

Is information available about the algorithm / decision-making rules and the impact of each feature?

Such information would be available for any model arising from this project.