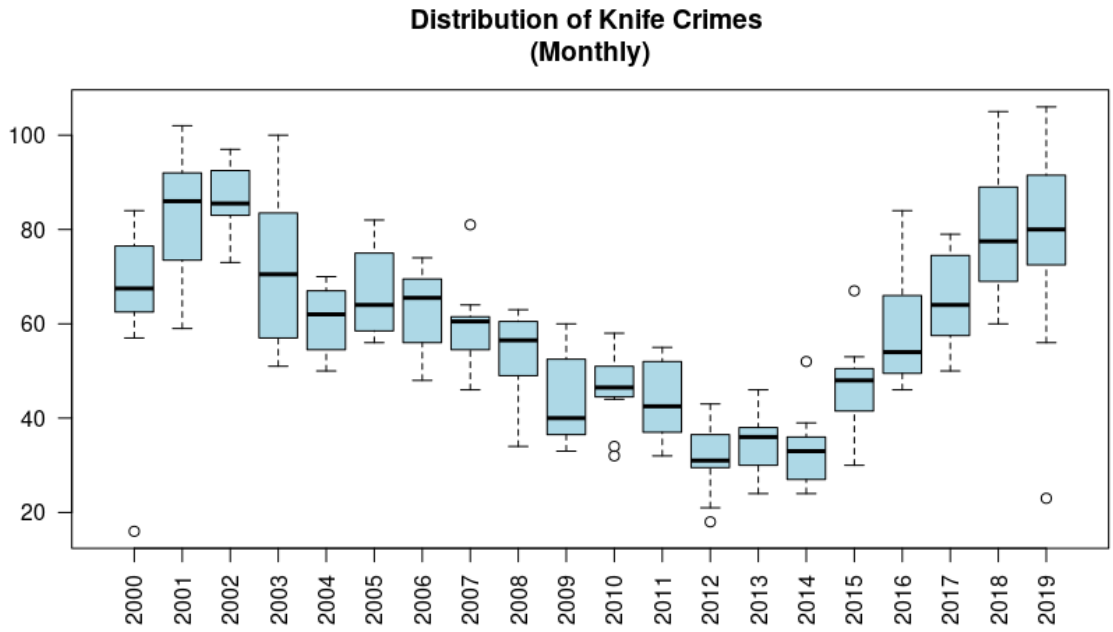
Ethics Committee Briefing Note

Project Reference: DAL_2019_0012_Knife Crime

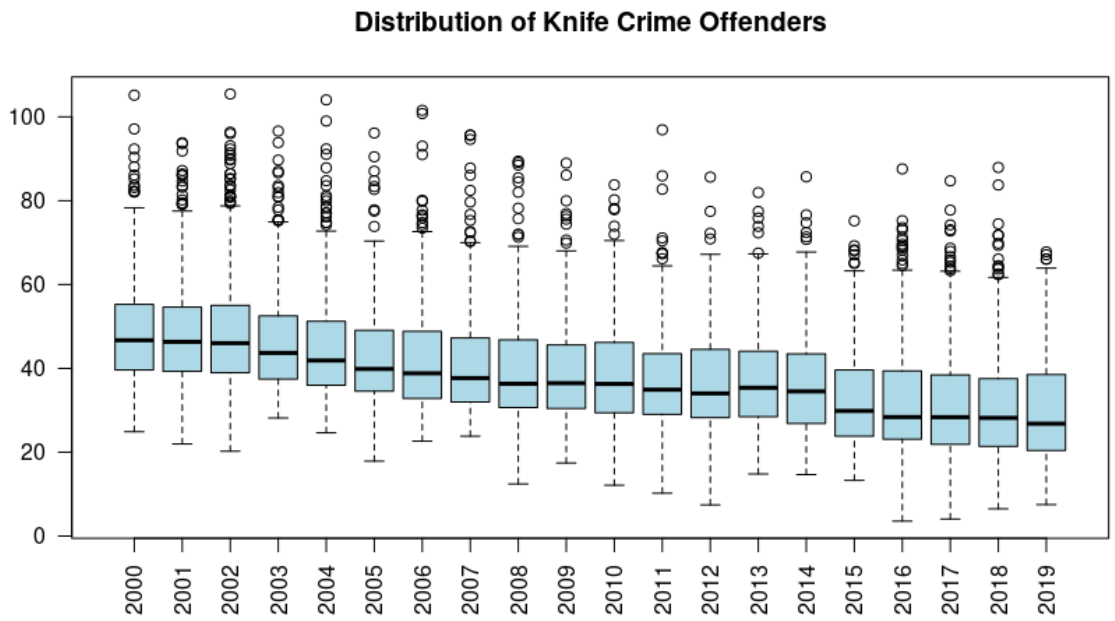
Purpose of data analysis:

Knife crime within the West Midlands has seen an increase over recent years from a low in 2012 to levels that were last seen in the early 2000s;



Indeed, the mean number of monthly knife crimes (where they were used causing injury) in 2019 represents an increase of 148% over the mean number of monthly crimes in 2012.

At the same time, the average age of those committing offences has also fallen:



This shows that knife crime is increasingly affecting younger members of the community.

The purpose of this project would be to develop a statistical model to predict the likely levels of knife crime over time and space within the WMP area. At present, this is envisaged to provide forecasts over a 4 weekly period allowing for potential responses to be planned, e.g. through project Guardian.

Source of analytical question / hypotheses to be examined:

The business question was identified via discussions with officers associated with project Guardian.

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
☑ West Midlands
☑ Other

The unit of analysis will be crimes where a knife was used causing injury within the WMP area.

Reliability of data:

An extensive exploratory data analysis (EDA) phase will be undertaken to examine the extent of any data quality issues. Discussions with subject matter experts (SMEs) will be undertaken both to capture any extraneous requirements and to sense check the analyses.

Sample or entirety:

If sample: Not applicable.

Method of sampling: Not applicable.

Method of choosing sample size: Not applicable.

Sample size: Not applicable.

Type of analysis:

☐ Exploratory
☐ Explanatory
☑ Predictive
☐ Optimisation

Proposed methodology:

It is apparent that there are patterns to knife crimes – used causing injury (hereinafter knife crimes), both over time and over space. Therefore at present it is envisaged that a spatio-temporal model would likely provide the best means of prediction, potentially also with a separate univariate time series model. For the purposes of spatial prediction, it is also envisaged that the WMP area would be broken into small spatial units in order to cater for analyses and predictions over areas that are distinct and small enough to allow operational activities to be planned.

Will the project eventually be automated:

☑ Yes
☐ No

Means of evaluation:

ALGO-CARE considerations:

Advisory:

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

Predictions arising from the model would be used to assess levels of risk and potential levels of harm and threat. This would feed into operational planning, particularly for project Guardian.

Does a human officer retain decision-making discretion?

The predictions would be for informational purposes only to feed into decisions that would ultimately be made by officers following operational protocols.

Lawful:

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

Given the rise in knife crime over the last few years this has fed into the priorities for WMP moving forward, indeed, allied to this is the goal of WMP of reducing serious violence (affecting young people) within WMP's Improvement Plan 2019-20 (which is to help WMP achieve the goals of its longer-term Ambition Plan), with a target of a 10% reduction in (non-DA) serious violence where either victim or offender are under 25 years old.

This project would feed into this aim.

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

Whilst data regarding individuals will be processed, this would be to produce aggregated data (counts of crimes over certain time periods and in various locations) as the basic unit of analysis would be crimes and their location. Therefore there would essentially be no interference with the privacy of individuals.

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

There are currently no means of predicting knife crime within WMP. Such a tool would enable a better picture of risk and potential requirements for resource allocation to be provided to WMP.

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 are from WMP systems and collected to enable their normal day-to-day operations.

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 it's purpose and the nature of the data processed?

Given the overall aim of predicting the number and locations of knife crime, using crime data aggregated over time and to bespoke spatial units would enable better decision making within WMP.

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

bias?

Using aggregated counts would best suit the aims of the project, but would not involve any other form of categorisation (given that crimes would be the unit of analysis).

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

The project 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 knife crime 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 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?

Currently envisaged as circa every four weeks.

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 would own the algorithm and 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 model 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?

Any model be developed and productionised, checks will be made as to its accuracy on an on-going basis (overall accuracy, sensitivity, specificity, AUC, etc.) 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?

The 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?

Not yet known, however any model developed would aim to maximise specificity whilst trying to gain as high a sensitivity as possible.

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

Currently the broad approach has been identified due to the nature of the business question and the data available.

What are the (potential) consequences of inaccurate forecasts?

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?

For any model that was productionised, it's accuracy would be assessed on an on-going basis via measuring its accuracy (sensitivity, specificity and AUC) as well as producing histograms of counts / 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) and lead to inefficient decision making and resource deployment.

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 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, when 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?

Enabling the reduction of knife crime 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 produced for any model arising from this project.