Ethics Committee Briefing Note

Project Reference: DAL_2019_0003_Disproportionality

Purpose of data analysis:

The issue of disproportionality within policing and involvement within the criminal justice system more generally has been a concern within the UK for a long time and has come to the fore in recent years culminating in the Lammy Review 2017. Much of the work involving assessing the presence of disproportionality has centred on arrest and incarceration rates and stop and search. These issues are complicated by characteristics of localities and wider socio-economic patterns present within them.

The broad aim for this project is to examine WMP data for the presence of disproportionality by ethnicity and gender.

Source of analytical question / hypotheses to be examined:

The business question was posed by the FET.

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

Reliability of data:

The data are sourced from WMP systems, notably crimes (data regarding crimes committed by individuals), eSearch (the stop and search data), ICIS (custody records), Use of force (a system whereby officers report any instances where some degree of force (greater than hand-cuffing)has had to be used and TP010 (traffic stops) data.

Sample or entirety:

Entirety.

If sample:

Not applicable.

Method of sampling:

Not applicable.

Method of choosing sample size:

Not applicable.

Sample size:

No applicable.

Type of analysis:

☑ Exploratory
☐ Explanatory
☐ Predictive
☐ Optimisation

Proposed methodology:

It is intended to go through the different systems and calculate risk ratios in order to identify where, in what ways and to what extent any disproportionality exists. The denominators for the risk ratio calculations will be formed from (a) the general population statistics as gained from the Census 2011, (b) the population statistics for the 10 most deprived wards (as a rough means of taking socio-economic factors into account) and the populations as present within the system in question.

In some circumstances simple logistic regression will be used in order to ascertain the odds of some event occurring associated with certain features (enabling, for example, an isolation of the odds of an event occurring given a certain ethnicity whilst also taking age into account). This obviates, where possible, the need for analyses to be split over a number of different dimensions.

Will the project eventually be automated:

☐ Yes
☑ No

Means of evaluation:

Not applicable.

ALGO-CARE considerations:

Advisory:

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

The results from the analyses will be used to inform WMP as to the potential extent and nature of

disproportionality.

Does a human officer retain decision-making discretion?

Not applicable.

Lawful:

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

It is a requirement that data regarding sensitive attributes of individuals are collected where possible in order that track can be kept of disproportionality, etc. The suggested analyses will enable an identification of the extent and nature of any disproportionality and as such could feed into any applicable policy development.

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

The analysis (in aggregate) of these sensitive attributes is necessary to answer the research question.

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

Whilst similar analyses have been (and are) undertaken and reported for certain systems' data, it has not been previously possible to undertake such explorations of this amount of data across this number of systems. The analyses are necessary to understand the current situation.

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 gathered as part of the normal day-to-day operations of WMP and are required to be collected as part of national requirements.

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?

Not applicable.

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

Whilst the underlying data are not categorised, reporting will be categorised according to ethnicity / gender as this is necessary in order to undertake the envisaged analyses. The potential for paradoxes arising from any such categorisation will be kept under review during the analysis.

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

As an exploratory project any issues with data quality will become apparent during its undertaking. It is envisioned that having a resulting idea of the extent of any disproportionality will be of benefit as this could help inform policy development.

Is the provenance and quality of the data sufficiently sound?

Excepting for any omissions, the provenance and quality of the data enable the analyses to be undertaken in a sound manner.

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.

Ownership:

Who owns the algorithm and the data analysed?

The data are wholly owned by WMP.

Does WMP need rights to access, use and amend the source code and data?

Not applicable.

Are there any contractual or other restrictions which might limit accountability or evaluation?

Not applicable.

How is the operation of the algorithm kept secure?

The data and the analyses are contained wholly within the WMP Hadoop system and the security measures employed therein. The type of analyses do not lend themselves to on-going operation (in an automated fashion).

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 use of risk ratios and, in some circumstances logistic regressions, are appropriate for exploring this area and would aid with any policy development.

Can the accuracy of the algorithm be validated periodically?

Not applicable.

Can the percentage of false positives / negatives be justified?

Not applicable.

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

The methods chosen are the most appropriate for the exploratory research question.

What are the (potential) consequences of inaccurate forecasts?

Not applicable.

Does this represent an acceptable risk?

Not applicable.

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

Not applicable.

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

Inaccuracies in the data would essentially lead to inaccurate results and so the veracity of the data is a factor to be examined during the project (checking for anomalies, discussions with SMEs, etc.).

Use of out of date data (even if accurate for the time) could essentially skew results applicable today by reflecting patterns that were present in the (relatively speaking) distant past. To avoid this issue, only data within specific time periods are utilised (which are different for the different systems primarily in relation to when the system came into use and the amount of data available).

Responsible:

Would the operation of the algorithm be considered fair?

As the analyses do not involve an algorithm in the sense of machine learning pattern recognition there is no differentiation in application across groups.

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

The analyses to be used are simple and are not to be employed on an on-going basis.

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

Given the desire for equity in treatment underlying the exploratory research question, any findings could be useful for the future approach of WMP and therefore also likely in the public interest.

Explainable:

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

The nature of the analyses is such that the features of interest constitute the methods and therefore provide the information regarding features used, etc.