Disclosure statement

Harassment (Protection from Harassment Act 1997, s.2) and stalking (Protection from Harassment Act 1997, s.2A) data

Background

This bespoke data collection exercise was conducted to gather detailed information from sentencers in magistrates' courts in England and Wales about how they sentenced offences of harassment (s.2) and stalking (s.2A). The data collection was undertaken pre and post guideline to support the evaluation of the Intimidatory offences definitive guidelines. The pre guideline data collection ran between 1 November 2017 and 30 March 2018 and the post guideline data collection ran between 23 April 2019 and 30 September 2019.These collections were mainly administered online via a survey hosted on the Sentencing Council's website.

The surveys collected information including the date of sentence and court location, some demographic details of the offenders, the relevant culpability and harm factors taken into account when each adult offender was sentenced, the sentence starting point, and the outcome imposed after any reduction for guilty plea. The potential disclosure issues surrounding these data have been carefully considered and the subsequent actions that were taken to ensure that the offender's confidentiality has been protected are outlined in this document.

Assessment of risk

In accordance with the Government Statistical Service (GSS) guidance, an assessment of the risk of disclosure was undertaken in order to determine which features of the data may help to protect individuals' identities and which aspects of the data may lead to disclosure risks. This is to help prevent identification of individuals and/or prevent additional knowledge of any previously unknown characteristics for those offenders sentenced for harassment (s.2) and stalking (s.2A) offences during the time periods of the data collection.

The following information was used to assess the level of risk:

- the sample size of the survey
- the response rate of the survey
- the likelihood that the data could be linked or matched to other data sources
- the nature and level of detail of the information collected

Applying disclosure control

The following changes were made to the data to reduce the chance of identification of individuals and thus maintain disclosure control. This included aggregation or banding of variables, where appropriate, in order to reduce unique counts and/or the removal of any variables that could make it easier to identify an individual:

- removal of court location and sentencer type
- removal of date the form was completed (a potential proxy for sentencing date)
- removal of free text fields where the raw data were deemed too disclosive and were not deemed high enough quality to clean and recode
- aggregation of low volume counts within certain variables into bands where appropriate
- removal of any factors with counts fewer than 10

Comments on these data

- The data collection covered all magistrates' courts in England and Wales, but the response rates were estimated at only 13 per cent for harassment and 14 per cent for stalking for the pre guideline data, and 21 per cent for harassment and 27 per cent for stalking for the post guideline data. This means that a substantial proportion of offenders sentenced for these offences during the period of the pre or post guideline data collections did not have a form filled in for them and are not captured in the data. Therefore, if someone knew some details about an offender sentenced for these offences during the period of the data collection and found a unique record in the data that seemed to match that individual, they still could not be certain that the record in the data relates to that specific person. As such, the Council considers that the data are sufficiently anonymous and feels that a higher level of disclosure control would be detrimental to the usefulness of the data.
- The likelihood has been judged as very low that individuals can be identified in the published data through linking to another dataset, given the lack of other publicly available data sources of the same information, and that the sample cannot be replicated.
- The names of the courts and date that the survey was filled out in the pre and post guideline periods have been removed and, as such, it is not possible to

identify exactly where the sentencing took place or when any particular form was filled in.

- Where the counts of records were extremely low (fewer than 10), these variables have been removed from the dataset to reduce the risk of positive identification (see the metadata file for further information).
- The age of the offender and the length of any custodial sentence were both collected in a banded format, which does limit the usefulness of the data but reduces the risk of disclosure.
- Sentencers were asked about the 'single most important factor' affecting their sentence, and also if there was anything else they would like to us about how they sentenced their offence, which were both free text fields. These included very specific details about the offender, offence, location and other details that may have risked being disclosive, so these variables have not been published.
- It is acknowledged that the cases about which these data relate would have been heard in open court. Although this may increase the risk of identification due to some of the defendant's details being more accessible, the impact is anticipated to be negligible due to the factors listed above. However, it also means that the impact of any potential disclosure issues is minimal as much of this information is already in the public domain.

If users would like any further details of the disclosure controls applied to the data, please contact the Analysis and Research team at Research@sentencingcouncil.gov.uk.