

# Prisoners data: quality assurance of administrative data used in population statistics, Jan 2017

The quality assurance undertaken on administrative data for prisoner data used within Population Statistics Division (PSD) publications.

Contact:  
Pete Large  
[pop.info@ons.gsi.gov.uk](mailto:pop.info@ons.gsi.gov.uk)

Release date:  
27 January 2017

Next release:  
To be announced

## Table of contents

1. [Introduction](#)
2. [Operational context and administrative data collection](#)
3. [Communication with data supply partners](#)
4. [Quality assurance principles, standards and checks applied by data suppliers](#)
5. [Producer's quality assurance investigations and documentation](#)
6. [Conclusion](#)

# 1 . Introduction

Each year data on the number of prisoners held in prison establishments in England and Wales are requested from the Ministry of Justice (MoJ) to help estimate the number and distribution of people resident in prison as part of the estimation of the usually resident population. Prisoners are treated as a special population in the population estimates, as it is assumed that movements of people into and out of prisons are not captured by the usual data sources used to estimate internal migration and international migration. For the purposes of the population estimates, a person is regarded as usually resident in a prison if they have been sentenced to serve 6 months or more.

This data feeds into 2 sets of population statistics we publish – the standard mid-year population estimates for local authorities in England and Wales and the Small Area Population Estimates (for various types of area) for England and Wales. To aid readability, this report is framed in terms of the higher profile local authority level estimates, with additional comments on small area population estimates included where appropriate.

This report covers the process from data collection through to the population estimates produced by Population Statistics Division (PSD). It identifies potential risks in data quality and accuracy as well as details of how those risks are mitigated.

This report does not aim to report on the whole of the mid-year population estimate processing or the quality assurance relating to the processing of the other components used in its production. Further information relating to the quality of the mid-year population estimates can be found in [Population Estimates Quality Tools](#) and in the [Annual Mid-Year Population Estimates quality and methodology information \(QMI\) document](#) on our website. A similar [Small Area Population Estimates QMI](#) is also available.

Population estimates are assessed as having a relatively high public interest profile. However, prisoners are a very small part of the population and any inaccuracies in the data would have a similarly small impact on the quality of the population estimates, particularly in respect of the uses typically made of the estimates. We have therefore judged that the appropriate level of assurance for this data source is A1.

The A1 rating means that a basic level of assurance is required for these sources and this document will provide information to meet this level of assurance. If you feel that this document does not adequately provide this assurance then please contact [pop.info@ons.gsi.gov.uk](mailto:pop.info@ons.gsi.gov.uk) with your concerns.

The toolkit outlines 4 areas for assurance; the rest of this document will be split into these areas. The areas for assurance are:

- operational context and administrative data collection
- communication with data supply partners
- Quality Assurance (QA) principles, standards and checks applied by data suppliers
- producer's QA investigations and documentation

## 2 . Operational context and administrative data collection

The prison population is formed of four main custody categories:

Remand: Prisoners in custody on remand are those awaiting commencement or continuation of trial prior to verdict. It also includes prisoners who have been convicted but are still waiting to be sentenced.

Sentenced: Those held in custody as a result of receiving a sentence in a criminal court. Persons committed for default of a fine are normally included in this group.

Recall: Those held in custody for breaching the terms of their licence conditions following release into the community.

Non-criminal: Those held for civil offences or under the Immigration Act. A civil non-criminal prisoner is someone who is in prison because of a non-criminal matter, for example, non-payment of council tax or contempt of court.

There are 2 different types of sentence length information available for sentenced prisoners in the population.

The judicially imposed sentence length is the sentence length given at court. The effective sentence length is the judicially imposed sentence length adjusted for any time already spent on remand, tagged bail or unlawfully at large. Population Statistics Division use the judicially imposed sentence length in the calculation of population estimates.

Prison establishments record details for individual inmates on the prison IT system known as Prison National Offender Management Information System (p-NOMIS). Information recorded includes details such as date of birth, sex, religion, nationality, ethnic origin, custody type, offence, reception and release dates and, for sentenced prisoners, sentence length. The data from individual prison establishments then feeds through to a central computer database, called the Inmate Information System (IIS), from which data extracts are used to produce various analyses of prison population.

On 30 June 2015, the data extracts used to produce statistics on the prison population transitioned to a new extract which extracts information from the Prison-NOMIS system directly and without needing to be processed by the Inmate Information System (IIS).

### **3 . Communication with data supply partners**

A Data Confidentiality Agreement exists between the Ministry of Justice (MoJ) and Office for National Statistics (ONS) This agreement covers prisoners' data for use in the mid-year population estimates.

The Data Confidentiality Agreement covers the:

- data required
- data access period
- purpose for which the data are provided
- products and publications in which the data will be used
- minimum information needed
- matching or linking
- lawful use of the data
- arrangement when period of access expires
- security of the data
- breach and dispute proceedings
- approval and signatories

We send a standard data request to MoJ around July each year via email. The data requested are:

1. prisoners held in England and Wales sentenced to six months or more as at 30 June of the reference year by: single year of age; sex; prison code; and name of prison (the specification of sentences of 6 months or more allows the data collected to correspond to the target concept of usual residence)
2. details of prisons or wings that have opened and prisons that have closed in the 12 months to 30 June of the reference year

The requested data are supplied by MoJ in a spreadsheet via email and stored on a restricted access drive. Access is limited to those Population Statistics Division (PSD) staff who are involved in the production process.

Any errors found or any queries from users following publication of the estimates are fed-back to MoJ via email for resolution.

The confidentiality of the data is preserved in outputs and publications.

### **3.1 Engagement with users**

PSD continually engages with users, through a variety of means, to understand how our outputs are meeting their requirements. Feedback provided tends to relate to the overall statistical methodology and the impact on the final statistics, rather than to any individual data source. To date no specific feedback on the use of this data source has been provided.

## **4 . Quality assurance principles, standards and checks applied by data suppliers**

This section details the checks and standards applied to the data prior to receipt by Population Statistics Division (PSD).

The Ministry of Justice (MoJ) carries out detailed checks including:

- checks for missing values and investigate any inconsistencies (for example a male in a female prison, or a juvenile in an adult prison)
- check of the custody type
- checks to previous quarters/years for any inconsistencies in the component
- verifying that the total prison population in the data matches the actual total number of prisoners in the estate
- checks that data being extracted from a new source is consistent with the original
- checks for males younger than 15 (who should have been recorded at a juvenile offenders institution)
- checks for females younger than 18 (who should have been recorded at a juvenile offenders institution)
- checks that the prison establishment is active

## **5 . Producer's quality assurance investigations and documentation**

This section details the checks and standards applied to the data for prisoners after receipt by Population Statistics Division (PSD). The checks carried out prior to receipt of the data are detailed in Section 4.

Checks are carried out by PSD once the prisoners dataset has been loaded onto the restricted access drive. The checks are designed to identify any obvious errors which have occurred in the size and content of the dataset as well as confirm that the data is plausible and the datasets are fit for purpose and can be processed in line with the agreed methodology.

The data received from Ministry of Justice (MoJ) are checked to make sure that the file received only contains prisoners in England and Wales on the spreadsheet with a sentence of 6 months or more by prison, age and sex. Checks are carried out on the data received to make sure codes exist for each prison (if not we input any missing codes); for any new prisons identified (their codes are added to the file); and to make sure that the file contains all the variables needed for processing.

Current prison totals are then compared to previous years' data to make sure there are no large or unexpected differences between the years. If the number has changed significantly, for example if there has been an increase in males and females where the percentage difference is over 25%, sometimes an explanation will have been offered on the raw data file (for example, that a new wing has been opened). Where possible these differences are checked against the capacity of the prison. If there are any unexplainable differences these are queried with our prison contact to resolve the issue and await explanations before processing.

If any errors in the data were to avoid detection through the quality assurance process, these would be expected to have small impacts on the population estimates of a limited number of areas and to be unlikely to affect decisions based on the estimates.

## **6 . Conclusion**

## 6.1 Strengths and Limitations

Overall, this data source is judged to be of adequate quality for the use to which it is being put in the population estimates:

- coverage is comprehensive, with a requirement for all prisoners to be included
- it is expected to be sufficiently up-to-date with no material impacts from any delays in updating the data source to reflect change

However, there are some acknowledged limitations in the data:

- the data has been collected for administrative purposes and is not necessarily consistent with estimates of the numbers of prisoners derived from the 2011 Census which are used in the base population estimates for 2011
- no information is available on place of previous (or next) usual residence which would allow a fully accurate and consistent accounting of moves between local authorities
- early release of prisoners may give rise to over count

## 6.2 Justification of rating

The local authority level population estimates are assessed as having a relatively high public interest profile. However, prisoners are a very small part of the population and any inaccuracies in the data would have a similarly small impact on the quality of the population estimates, particularly in respect of the uses typically made of the estimates (for example in planning housing provision or demand for local services). Whilst this population subgroup might be a larger part of the estimate for a smaller area (published as part of the small area population estimates) these estimates are judged to have a much lower profile of public interest. The administrative data supplied corresponds well to the target concept of usual residence. The quality assurance process followed at each stage of the data collection, data supply and processing is judged to be of high quality. We have therefore judged that the appropriate level of assurance for this data source is A1 - basic assurance.