

# Prisoners data, quality assurance of administrative data used in population statistics, May 2019

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

Contact:  
Joanna Harkrader  
pop.info@ons.gov.uk  
+44 (0)1329 444661

Release date:  
16 May 2019

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

## What are prisoners data?

Prisoners data relate to people who have been convicted of a crime by a court of law and sentenced to a jail term. Within population statistics we regard prisoners as being those who have been sentenced to serve a period of six months or more in jail as being in line with the usual resident population basis that we use.

## How are prisoners data used by the ONS in population statistics?

Each year data on the number of prisoners held in prison establishments in England and Wales are requested from the Ministry of Justice (MoJ). These help estimate the number and distribution of people resident in prison as part of the estimation of the usually resident population and help build a picture of how many households there could be in future years in England over a 25-year period.

The Office for National Statistics (ONS) receives two data supplies from MoJ, one is a bespoke output covering age, sex and location of prison, the other is a published output but does not contain age.

## Mid-year population estimates

Population estimates for England and Wales include all prisoners serving a jail sentence in England and Wales.

Prisoners, who conform to our usually resident population criteria, are treated as a special population in the mid-year population estimates, as movements of people between prisons are not captured by the usual data sources used to estimate internal migration. This is because prisoners are treated under a separate health system and prison transfers will not appear as a GP re-registration on the data source used for migration estimation. Using this data source ensures that the prisoner population can be adequately captured.

The MoJ supply data on the number of people resident in prisons in England and Wales of that reference year. This is by prison location, sex and age. For the purposes of creating population estimates a person is regarded as usually resident in a prison if they have been sentenced to serve a period of six months or more in jail.

Change in the prisoner population between the two mid-year points is estimated by subtracting the previous mid-year's estimated prisoner population from the total population estimate for the current mid-year and then adding in the current year's estimated prisoner population, by area, sex and age. The prisoner population distribution can fluctuate widely between mid-year points due to operational needs.

## Household projections

Household projections for England provide projections of the population living in households in future years based on current information.

Prisoners are classed as people living in communal establishments and as such do not conform to the definition of people living in households. Prisoners data are used to make an adjustment to population estimates and projections to better reflect the population of concern when calculating household projections. Corrections for other communal establishments are made using census data.

It should be noted that we do not include people held at immigration removal centres (IRCs) as they are non-criminal detainees and therefore not included in the definition of those held for six months or more.

We collate published prison population data from MoJ about the number of prisoners in each prison in England by sex, grouping data for individual prisons together by local authority.

To calculate the proportion of the population in each local authority by sex, serving sentences of six months or more, we apply the distribution from published data about the number of prisoners by sentence length for England and Wales to the local authority data (data for England only are not published). Finally, we apply the England and Wales-level age distribution of prisoners to the local authority data, to provide an age breakdown at local authority level.

This report covers the processes involved from data collection through to the population estimates and household projections produced by the Centre for Ageing and Demography. 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 or household projections processing or the quality assurance relating to the processing of the other components used in their production.

This report is intended to supplement existing documentation, such as:

- [Mid-year population estimates Quality and Methodology Information \(QMI\)](#)
- [Mid-year population estimates methodology](#)
- [Mid-year population estimates mid-2017](#)
- [Small area population estimates QMI](#)
- [Small area population estimates methodology](#)
- [Small area population estimates mid-2017](#)
- [Household projections QMI](#)
- [Household projections methodology](#)
- [Household projections mid-2016 based](#)

## Strengths and limitations

Overall, this data source is judged to be of suitable and sufficient quality for the use to which it is being put within the population estimates and household projections methodologies.

There are several strengths in using prisoner data for these purposes:

- data coverage is comprehensive, with a requirement for all prisoners who meet our usual residence criteria to be included
- these data provide valuable insight into the changing size and structure of the prison population between censuses
- for logistical reasons, the data on prisoners are kept up-to-date, meaning that the data are accurate and do not suffer from lag effects seen on some administrative sources
- the data cover a special population of residents of communal establishments, who may otherwise be unaccounted for in household projections
- data accuracy by geographical area means we can reflect timely information about prisons and prison wings that have opened and closed throughout our population statistics

However, there are some acknowledged limitations when applying the data to our uses:

- mid-year population estimates are re-based for census years using data from the census; prisoner information is taken from the census for those years, the most recent of which was 2011; the data provided here have been collected for administrative purposes and are provided on a mid-year basis, as such they are not necessarily consistent with the data used 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 areas when entering or leaving prison
- early release of prisoners may give rise to over-count
- the data used for household projections are not broken down by the age groups we use and publish by, therefore we need to model these to gain an understanding of data relating to age
- prisoners data cannot be used to distinguish prisoners who leave the prison system and those who have died; all outflows are assumed to be leavers and therefore local authorities containing prisons may sometimes have residents who are double-counted as they are deemed to have left for another area and are counted as an out-migrant but are not given a place of arrival; they are then removed again when their death is registered

## What is QAAD?

The Quality Assurance of Administrative Data (QAAD) is a framework created by the UK Statistics Authority. The framework is a way of assessing the quality of data inputs to ensure that they are of suitable and sufficient quality for use in National Statistics methodologies. It is a way of ensuring that the final statistics are built on a foundation of robust data.

QAAD reports are required to demonstrate the quality of the data from the point the data are passed from an individual to an organisation, through to the point that the data are used in the production processes of National Statistics. The relative strengths and weaknesses of the data for the purpose to which they are being applied are also highlighted.

Within the UK Statistics Authority's [QAAD – Setting the Standard \(PDF, 298KB\)](#) documentation it states, "The need for investigation and documentation increases at each level of assurance."

The [QAAD Toolkit](#) sets out four levels for the quality assurance that may be required of a dataset:

- A0 – no assurance
- A1 – basic assurance
- A2 – enhanced assurance
- A3 – comprehensive assurance

The UK Statistics Authority states that the A0 level is not compliant with the [Code of Practice for Statistics](#). The assessment of the assurance level is in turn based on a combination of assessments of data quality risk and public interest. The toolkit sets out the level of assurances required as follows:

- A1: basic assurance – the statistical producer has reviewed and published a summary of the administrative data quality assurance (QA) arrangements
- A2: enhanced assurance – the statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance
- A3: comprehensive assurance – the statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit and published detailed documentation about the assurance and audit

Within the UK Statistics Authority [QAAD – Setting the standard \(PDF, 298KB\)](#) documentation it states:

“Quality assurance of administrative data is more than simply checking that the figures add up. It is an ongoing, iterative process to assess the data’s fitness to serve their purpose. It covers the entire statistical production process and involves monitoring data quality over time and reporting on variations in that quality. Post collection quality assurance methods, such as data validation, are an important part of the quality assurance process, but can be of limited value if the underlying data are of poor quality. The Authority encourages the application of critical judgment of the underlying data from administrative systems before the data are extracted for supply into the statistical production process. As with survey data, producers need to: investigate the administrative data to identify errors, uncertainty and potential bias in the data; make efforts to understand why these errors occur and to manage or, if possible, eliminate them; and communicate to users how these could affect the statistics and their use.”

The toolkit outlines four areas of assurance; the rest of this report can be split into these areas. The areas for assurance are:

- operational context and administrative data collection
- communication with data supply partners
- quality assurance principles, standards and checks applied by data suppliers
- producer’s quality assurance investigations and documentation

## **Assessment of quality assurance level**

Our assessment was carried out using the [QAAD Toolkit \(PDF, 195KB\)](#). This assesses an administrative data source in terms of the risk to data quality and its onward use in statistics as well as the profile of the statistics produced from the source. The matrix approach to assessment advised by the UK Statistics Authority has two components: separate assessments of the public profile of our statistics (low, medium, high) and data quality concerns about the inputs for our statistics (low, medium, high).

Quality assurance of our final statistics and the methods used to create our statistics are not part of this assessment but are subject to scrutiny elsewhere as part of our National Statistics accreditation.

The outcome of our assessment then determines the level of assurance and documentation required to inform people about the quality assurance arrangements in place for the administrative systems from which our statistics are sourced. The results of those assessments are an A1 rating.

The A1 rating means that a basic level of assurance is required for these sources and this report will provide information to meet this level of assurance.

Table 1: Summary of quality assurance level for prisoners data

<b>Use</b>	<b>Risk of quality concern</b>	<b>Public profile</b>	<b>Assurance level</b>
Overall assessment			A1 - basic
Population Estimates Unit (PEU):	Low	High	A1 – basic
• mid-year estimates	Low	Low	A1 – basic
• small area population estimates			
Household projections	Low	High	A1 - basic

Source: Office for National Statistics

Prisoner data used within our outputs have been given an A1 rating based on risk profile scoring. This is a basic level of assurance. Prisoner data are considered a low concern in terms of quality (risk) and are used in statistics of high public interest (profile).

The level of risk of data quality concerns is low because:

- prisoners are a very small part of the population
- any inaccuracies in the data would have a similarly small impact on the quality of the population estimates or household projections
- 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

If you feel this report does not adequately provide this assurance then please contact [pop.info@ons.gov.uk](mailto:pop.info@ons.gov.uk) with your concerns.

## 2 . Operational context and administrative data collection

### Custody categories

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 that 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.

## **Sentence length**

There are two 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

## **Recording of information**

Prison establishments record details for individual inmates on the prison IT system known as Prison National Offender Management Information System (Prison-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.

Information related to name, sex, date of birth and sentence are taken from the court warrant, which provides the legal authority to detain that person. Court warrants obtain personal information from the plea from the individual submitted ahead of their trial.

Other information collected will be self-declared by the individual when processed at the prison. If during the information collection process, the individual has official identification documentation on their person, such as a passport, then this is used to verify the recorded information. Subsequent variables needed to inform statistics, including age and duration of sentence are derived from these. These data are subject to a number of management checks to ensure their accuracy on the system.

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 IIS.

### 3 . Communication with data supply partners

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

The MoU covers:

- data required
- data access period
- purpose for which the data are provided
- products and publications
- 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 the MoJ around July each year via email in order to secure data that are not already published for use in population estimates. The data requested are:

- 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 six months or more allows the data collected to correspond to the target concept of usual residence)
- 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 is supplied by the MoJ in an Excel spreadsheet via email and loaded into a secure working environment. Any errors found are fed-back to the MoJ via email for resolution. The confidentiality of the data is preserved in outputs and publications.



The use of prisoners data for household projections was discussed with the MoJ prior to application in the method. While the data supplied for population estimates may be of higher quality, given the use of the data, the MoJ specifically requested we use the published data as part of statistical disclosure control. This is because the prisoners data supplied for population estimates are an unpublished dataset and all of the other sources for household projections are published, as such it would be possible to determine the population estimate prisoners data supply through disclosure by differencing.

## Engagement with users

The ONS 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 these data sources have 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 the Centre for Ageing and Demography.

### Data quality arrangements for the MoJ

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

- checks for missing values and investigation of any inconsistencies (for example, a male in a female prison, or a juvenile in an adult prison)
- checks the custody type
- checks previous quarters or 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 years (who should have been recorded at a juvenile offenders institution)
- checks for females younger than 18 years (who should have been recorded at a juvenile offenders institution)
- checks that the prison establishment is active

Alongside these checks, the MoJ have set out their [quality strategy principles and processes \(PDF, 358KB\)](#) to ensure that the statistics they produce are of high quality. It also outlines the steps they will undertake to meet these requirements, which include providing information on data quality issues faced. This ensures that they are being as transparent as possible.

The following is an example of data quality issues identified and resolved by the MoJ. The National Offender Management Service began the roll-out of a new case management system for prisons. During this roll-out, data collection issues emerged, which affected the supply of data for statistical purposes (July 2009 to February 2010). This affected information on sentence length and offence group, and meant that this information was not available during this period. They successfully resolved the issue in March 2010.

Other data quality work conducted in the summer of 2015 led to improvement in the methods for identifying sentence information for indeterminate sentences. This resulted in a decrease in the number of offenders classified as “tariff not known”.

The MoJ has found that the use of the Prison-NOMIS system has resulted in improvements in data quality and more detailed information about the prison population becoming available, this includes more accurate sentence length information.

The MoJ do note that the figures in the supplied prison tables have been drawn from administrative IT systems which, as with any large-scale recordings system, are subject to possible errors with data entry and processing, but this has, where possible, been kept to a minimum.

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

This section details the checks and standards applied to the data for prisoners after receipt by the Centre for Ageing and Demography. The checks carried out prior to receipt of the data are detailed in Section 4.

### **Population Estimates Unit (PEU)**

The checks carried out on the prisoners data are designed to:

- identify any obvious errors that have occurred in the size and content of the dataset
- confirm that the data are plausible
- confirm that the datasets are fit for purpose
- confirm that the data can be processed in line with the agreed methodology

The data received from the Ministry of Justice (MoJ) is checked to make sure that the file received only contains prisoners in England and Wales on the spreadsheet with a sentence of six months or more by prison, age and sex. Checks are carried out on to prison codes to identify whether any new prisons have been added. A check is also carried out to make sure that the file contains all the variables needed for processing.

Current prison totals, supplied by prison location, are then compared with data from previous years 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 differences that we are unable to determine an acceptable reason for, these are queried with our prison data contact, who provides an explanation before we carry out statistical 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.

## Household projections

Offender Management Statistics (quarterly) publications are received from the MoJ via their website. It should be noted that checks are carried out by the data supplier prior to them being published on their website. Some checks are carried out to ensure that the data taken from the website match figures held by the Office for National Statistics (ONS).

The processing of the prison population data for household projections is dual-run by two people, to ensure that the same results are achieved after each step of the processing.

At the end of this, further quality assurance checks are completed including sense checks against published prison figures by local authority to see if they match up. Any differences are investigated to establish the cause. If the issue is found to be with the supplied data then the MoJ is contacted to provide clarification or to re-supply the data.

Advice was taken from the MoJ about which categories of prisoners should be removed from the data, in order to produce figures for prisoners serving sentences of six months or more. The reason why these prisoners are taken out within the processing stage is due to the processes only counting people who serve sentences lasting six months or more. If they are serving less than six months then they are not included as part of the communal establishment population.

## 6 . Conclusion

The prisoner data used in the production of population estimates and household projections are deemed to be low risk in terms of data quality concerns. The data also have a number of strengths to justify inclusion in these statistical methods. As such, they are deemed to carry an A1 assurance rating, using the QAAD toolkit, meaning that a basic level of assurance is provided.

Prisoner data are used in population estimates to account for changes in the prison population, particularly the geographical distribution, which is affected by prison transfers.

Prisoner data are used in household projections to make an adjustment to the total population, so that the resulting data better reflects the population living in households to project from as prisons are deemed to be a communal establishment.

Prisoner data are subject to a number of checks both before and after they are provided to us to use in the previously described ways to ensure that the individual data supply is fit for use in statistical production processes.