

UK Armed Forces data: quality assurance of administrative data used in population statistics, Jan 2017

Report detailing the quality assurance undertaken on administrative data for Home Armed Forces data used within Population Statistics Division (PSD) publications.

Contact:
Pete Large
pop.info@ons.gsi.gov.uk

Release date:
20 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. [QA principles, standards and checks applied by data suppliers](#)
5. [Producer's QA investigations and documentation](#)
6. [Conclusion](#)

1 . Introduction

Each year, the Population Statistics Division (PSD) in the Office for National Statistics (ONS) requests data on numbers of UK Armed Forces from the Ministry of Defence (MoD). This data is used to help estimate the number and distribution of UK Regular Armed Forces (AF) personnel (and UK Regular AF and their dependants overseas who return to the UK) as part of the estimation of the usually resident population in local authorities in England and Wales. UK Regular AF are treated as a special population – termed the Home Armed Forces (HAF) – in the calculation of the estimates as the movements of military personnel are not captured by the usual data sources used to estimate international and internal migration.

To this end, data on numbers of UK Regular AF, and their dependants (for forces based overseas), are requested from Defence Statistics (Tri-Service) and British Forces Germany (BFG) via email from the PSD. Defence Statistics (Tri-Service) is a branch within the MoD. The collected data, which assigns people to an Armed Forces base, is transformed to data by local authority of residence (as required for the population estimates) using data from the 2011 Census.

This data feeds into 3 sets of population statistics that we publish – the standard mid-year population estimates for local authorities in England and Wales, the subnational projections for England, and the small area population estimates (for various types of area) for England and Wales. For readability, this report is framed in terms of the higher-profile local authority level estimates, with additional comments on the small area population estimates and subnational population projections included where appropriate.

This report describes the quality assurance of data from Defence Statistics (Tri-Service) and BFG, from its initial collection through to its inclusion in publications produced by our 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 QMI](#). A similar [small area population estimates QMI](#) is also available.

Population estimates are assessed as having a relatively high public-interest profile. However, UK Regular AF are a small part of the population in most areas and any likely 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. Furthermore, the data collected are judged to correspond well with the target concept of usual resident. We judge it unlikely that there would be a change in the information collected which would impact on the quality of the estimates.

Though uncertainty in the future location of the Armed Forces is an additional element of uncertainty in the subnational population projections, this uncertainty is simply uncertainty over the future rather than reflecting issues over the quality of this data source. We have therefore assessed this data source as having an A1 rating for the purpose of this quality assurance report.

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.gov.gsi.uk with your concerns.

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

- 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 Ministry of Defence's (MoD's) personnel data are derived from the department's Joint Personnel Administration (JPA) system, which is used for the administration of almost all Armed Forces (AF) personnel.

All personnel in the UK Regular AF must be recorded on JPA in order to receive their pay. Therefore, the overall strength figures are accurate. However, more detailed breakdowns relying on the information recorded for each individual can be less accurate due to variable quality of the data entered for these different fields. The variation in quality is partly due to important information required for managing individuals being recorded and updated centrally, whereas other information is left to the individual to complete through a self-service tool. There are also elements of late reporting though this is unlikely to be a material issue for the use of aggregate data in population estimates.

3 . Communication with data supply partners

A Data Confidentiality Agreement exists between the Ministry of Defence (MoD) and the Office for National Statistics (ONS). This agreement covers Armed Forces data for use in the mid-year population estimates.

The Data Confidentiality Agreement covers:

- data to be provided
- permitted use of the data
- permitted distribution of the data
- security

Staff in the Population Statistics division (PSD) contact Defence Statistics (Tri-Service) and British Forces Germany (BFG), via email, in the summer of each reference year to request the specified data. Variables requested are listed below:

Defence Statistics (Tri-Service)

- service – Army, Royal Navy and Royal Marines, or Royal Air Force
- continent, country, region or local authority
- assign type – Regular or Gurkha
- age
- sex
- count

Data from Defence Statistics (Tri-Service) is received in quarters:

- January
- April
- July
- October

It is the July quarter that is used in the production of the mid-year estimates. The July extract is used in the derivation of the population estimates as it contains the reference date of 30 June, but this extract is checked against those for other quarters to ensure that there are no implausible differences.

British Forces Germany (BFG) data is received for:

- military personnel by age and sex
- children by age and sex
- spouses by age and sex

The data we receive is stored in a set-up folder for the reference year. This is on a shared drive with access restricted to PSD staff. If any problems with the data are identified, Defence Statistics (Tri-Service) or BFG are contacted.

3.1 Engagement with users

PSD continually engages with users to understand how well outputs meet their requirements. PSD's user engagement activities include formal consultations on proposed changes to outputs, regular communication on plans through a quarterly newsletter, and external events open to all users. In addition, where evaluating changes to methods or sources has required specialist knowledge of local areas, PSD has organised Local Insight Reference Panels to elicit the views of relevant local authorities. From these activities, any issues relating to the sources and their fitness for the proposed use, will naturally come out. Issues restricted to one output will generally be addressed by the team responsible for that output while the Stakeholder Engagement team in PSD takes an overview of any issues with more general implications and ensures that this is considered in development of outputs across the division. It should be noted that users are more likely to comment on the overall methodology and the effect that it has on the final statistics than on a contributory data source.

Any issues around the quality of the statistics are described in the Quality and Methodology Information report accompanying each output. Issues around specific administrative data sources used in producing the statistics are considered in Quality Assurance of Administrative Data reports such as this.

When changes are proposed to methods (including changes in data sources being used in producing statistics) the ONS Population Methodology and Statistical Infrastructure Division will assess the resultant methods prior to implementation to assure that they are of sufficient statistical quality to meet user needs and are an improvement on the previous method. An independent evaluation by academic experts may also be undertaken, should methodological changes be extensive. The methods are also subject to scrutiny by the UK Statistics Authority as part of the National Statistics accreditation programme under Principle 4 of the Code of Practice for Official Statistics (sound methods and assured quality).

The Responsible Statistician is named for each release and contact details for them are provided, so should someone have concerns over the statistics they are able to communicate them with us. Methodology documents are published to enable users to provide scrutiny.

4 . QA 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 checks carried out by PSD upon receipt of the data are detailed in section 5.

Data collected from Defence Business Services, which manages the HR administrative system, is taken from the Joint Personnel Administration system as described in section 2.

Data is initially passed through Defence Statistics' Extract Transform and Load System to automatically validate the data against business rules. As a result of these checks, edits are made to the data to ensure the basic quality of the data and a series of derived fields are also calculated. The data are then passed to the 3 Defence Statistics Manpower branches who work with each of the single services, which then analyse the data to correct for known errors. Regular feedback ensures Defence Statistics staff are kept abreast of any changes or potential issues with the data and statistics, which is fed into the data validation and editing process. Only when it has passed these 2 stages is the data approved for supply to PSD.

Principle 4 in the National Statistics Code of Practice for Official Statistics is used to judge the quality of the data which is assessed using the Government Statistical Service's Quality, Methods and Harmonisation Tool and against European Statistical System's dimensions of quality. Defence Business Services are audited by the Defence Internal Audit organisation. The quality assurance carried out by Defence Statistics includes an extensive range of data quality checks.

5 . Producer's QA investigations and documentation

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

Checks are carried out on the components of population change (including the Home Armed Forces (HAF) component) to ensure that the data are accurate enough to be included in the mid-year estimates. Checks include, for example, that there are no significant changes between years for the number of UK Armed Forces (AF). Similar checks, at lower geographical levels, are carried out for the small area population estimates. Should the data have changed significantly, then several years will be investigated to see if it is reflective of a longer trend. If this does not prove to be the case then Defence Statistics (Tri-Service) or British Forces Germany (BFG) is contacted for confirmation and/or investigation. More detail on the checks for the Defence Statistics (Tri-Service) and BFG data is provided below.

Defence Statistics (Tri-Service)

The following checks are carried out in Excel spreadsheets on the Defence Statistics (Tri-Service) data:

- correct and complete data has been provided
- a total count of Armed Forces personnel for each extract (data for January, April, July and October) to ensure they are similar

If any problems are identified from the checks above, Defence Statistics (Tri-Service) are contacted.

BFG

The following checks are carried out in Excel spreadsheets on the BFG data:

- correct and complete data has been provided
- a total count of UK AF personnel stationed in Germany compared with previous years
- similar comparisons by age and sex
- a count of dependants by age and sex follow an expected distribution: partners aged 20 to 50 (mainly female) and children
- the total number of dependants is compared with previous years
- the ratio of dependants to overseas UK AF personnel is compared with previous years

If any problems are identified from the checks above, BFG are contacted.

6 . Conclusion

Population estimates are used to make decisions on resource allocation by central and local government and as such are assessed as having a relatively high public-interest profile. However, the Home Armed Forces are a small part of the population in most areas. Any substantial changes to the deployment of Armed Forces overseas resulting in large numbers of Armed Forces personnel leaving or returning to the UK, or the opening or closing of bases, could have an impact on the estimates at the local level around Armed Forces station locations. However, any likely inaccuracies in the data are judged likely to have only a 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).

Uncertainty about the future location of Armed Forces is an additional element of uncertainty in the subnational population projections but this uncertainty is simply uncertainty about the future rather than a weakness of the data source.

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 data from the administrative source used is judged to be likely to correspond reasonably well with the target concept of usual resident of the UK, though the data provided relates to workplace rather than local authority of residence. A formal agreement for provision of the data is in place, and the nature of the data collected is such that it is judged unlikely that there would be a material change in definitions or information collected which would impact on the quality of the estimates. As such we have deemed the UK Armed Forces data to have a low risk of quality concerns.

We have therefore assessed this data source as having an A1 rating for the purpose of this quality assurance report.