

Quality Indicators for Sub-National Mid-Year Population Estimates in Scotland

Interim Report

1. Introduction and Purpose

- 1.1 This paper presents a discussion of existing and possible quality indicators and measures for mid-year population estimates at local authority level in Scotland. Mid-year population estimates currently undergo a number of quality checks at local authority level before they are published. But there are no fixed procedures in place to identify and adjust any anomalous estimates. Decisions on any adjustments to improve the quality of the population estimates are left to the discretion of the staff in Population and Migration Statistics branch who are responsible for the production of the estimates and undertaking the quality checks. This paper will look at more objective methods of improving the quality of mid-year population estimates using available data.
- 1.2 The Office for National Statistics (ONS) is developing methods of improving the quality of mid-year population estimates for England and Wales under the Migration Statistics Improvement Programme¹ (MSIP). Some of the work being undertaken by ONS is described in this paper, along with the possibility of adapting it for use in Scotland.
- 1.3 One relatively straightforward method of determining the quality of a population estimate is to apply a plausibility range to each estimate. These plausibility ranges consist of lower and upper limits for each estimate, based on information from a number of different sources such as administrative data. Any estimate that lies outside the plausibility range can be easily identified and investigated for possible adjustment. [Section 3](#) discusses the use of plausibility ranges to improve the quality of mid-year population estimates.
- 1.4 Another indicator of the quality of the population estimates could be achieved by calculating a measure of uncertainty for each estimate. This is a more complex and long-term piece of work and would involve calculating the uncertainty in each component of population change. In effect, this is more or less the equivalent of measuring the uncertainty in migration estimates (internal and international), since this is the component of change for which most uncertainty applies. [Section 4](#) looks at the work that the Office for National Statistics (ONS) is proposing to do on this subject.
- 1.5 [Section 5](#) gives a brief summary of the quality assurance work being done on the 2011 Census population estimates. The quality assurance work for the 2011 Census will be done at local authority and data zone levels. High quality census data at these levels will help ensure a similar quality for mid-year estimates at local authority level for subsequent years.

¹ <http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/mig-stats-improve-prog/index.html>

2. Background

- 2.1 The General Register Office for Scotland (GROS) produces annual mid-year population estimates for Scotland and its administrative areas (councils and NHS board areas). These estimates are based on the most recent population census and are derived using the standard demographic cohort-component method. This method 'ages on' the population from the previous year and applies natural change (births minus deaths) and migration estimates.
- 2.2 The main area of uncertainty in the process is the estimation of migration (internal and international). Migration is derived from a variety of sources, such as GP registers and the International Passenger Survey. For a variety of reasons, each of the sources can potentially introduce inaccuracies into the population estimates. As we get further away from the census these inaccuracies could be compounded.
- 2.3 As part of the process of producing the mid-year population estimates various quality assurance checks are carried out on the population estimates for council and NHS board areas. These include:
 - Comparing population estimates with previous years.
 - Comparing migration estimates with previous years.
 - Comparing sex ratios with previous years.
 - Comparing armed forces counts with previous years.
 - Comparing asylum seekers counts with previous years.
 - Comparing household sizes with previous years.
 - Checking population estimates for children against child benefit and school census data.
 - Checking population estimates for adults against electoral register statistics.
 - Checking population estimates for pensioners against data from the older persons database.
 - Checking migration estimates against National Insurance registration data.
- 2.4 These checks can help identify possible inaccuracies in the population estimates at council and NHS board area level. However, there are some issues with the current quality assurance checking:
 - There are no guidelines that specify when a population estimate should be adjusted to improve the quality of the data. Decisions are left to the discretion of those responsible for producing the estimates.
 - Some of the checking relies on the data from the previous year being accurate.
 - The administrative sources used to do the checking will usually not give complete coverage of the population (or sub-group of the population), and in some cases the coverage may be substantially below the total population.
 - Some administrative data sources may not be available at the time that the mid-year estimates are being produced.

3. Plausibility Ranges

- 3.1 As part of the MSIP Phase 2 work to improve mid-year population estimates, ONS is planning to reconcile the estimates with administrative sources. This work will deliver a plausibility range for each estimate and a strategy for applying contingency adjustments to estimates that lie outside the plausibility range. The range is determined using data from administrative sources to set upper and lower limits for each population estimate.
- 3.2 Plausibility ranges would be fairly straightforward to set up, assuming at least two sources of suitable administrative data are available for each age and sex.
- 3.3 A suitable choice for the upper limit of the range would be the GP register data. The number of Currently Registered Patients (CRP) is established from the Community Health Index (CHI) extract that is provided to GROS each year. The CRP is known to overestimate the size of the population, so we might expect that population estimates would be lower than the corresponding count from the CRP. However, it should be remembered that the CRP figures and the mid-year population estimates are not independent as the CRP is used in determining the distribution of migrants across councils. Other administrative data sources are more than likely to underestimate the size of the population and would be suitable for the lower limit of the range.
- 3.4 Figures 3.1 and 3.2 below show examples of how the 2009 mid-year estimates could be reconciled with administrative data sources by setting appropriate plausibility ranges. In the first example, the Mid-Year Estimates (MYE) are compared with Child Benefit data (CB) and the Currently Registered Patient count (CRP). It would appear that the MYE may be too high for ages 0 to 9 and we may need to apply an adjustment. In the second example, the mid-year estimates (MYE) are compared with data from the Older Persons Database (SOPD) and the currently registered patient count (CRP). Here, the MYE appears to over-estimate the population for most ages, but the plausibility range is very narrow because the SOPD and CRP values are very close to each other.

Figure 3.1: Example of council area plausibility ranges, 2009, ages 0-15, using CB and CRP

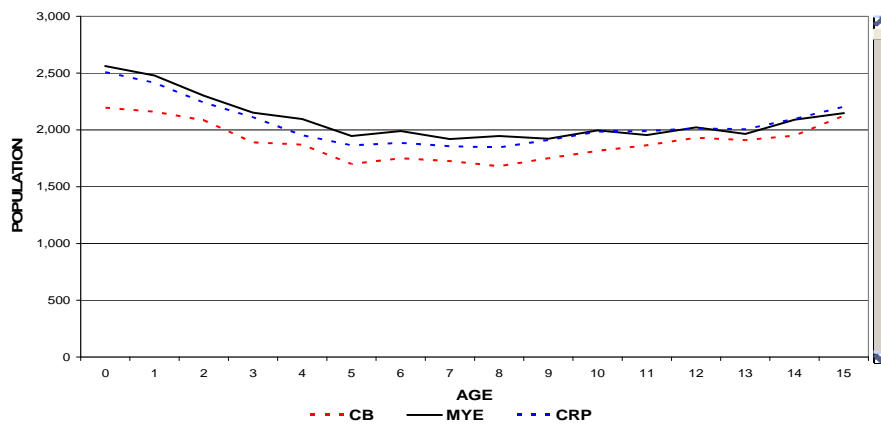
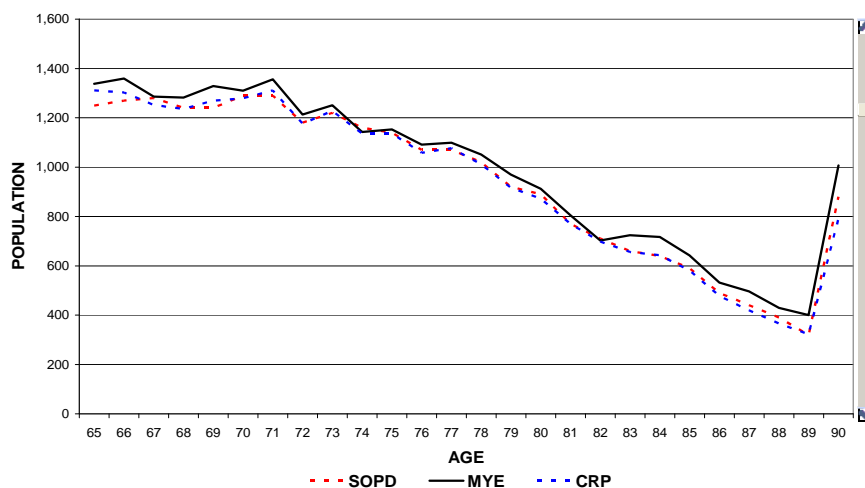


Figure 3.2: Example of council area plausibility ranges, 2009, ages 65+, using SOPD and CRP



3.5 While it is relatively straightforward to set up the plausibility ranges for the mid-year population estimates, it may be more difficult to establish a strategy for deciding when and how to apply contingency adjustments to the population estimates. Some issues that need to be addressed include:

- The administrative data used for the lower/upper limit may not always be less than/more than that for the upper/lower limit.
- The plausibility range may be very narrow (Figure 3.2 above) or very broad.
- Should the administrative data used for the upper and lower limits be adjusted in some cases? For example, in Figure 3.2 above, adding 5% or 10% to the upper limit and subtracting the same amount from the lower limit for each age might give a more practical plausibility range.
- The mid-year population estimates and the administrative sources used to set the plausibility range may not be independent.
- The administrative data used to construct the plausibility range may not be available in time for the mid-year estimates.

- Should contingency adjustments be made in all cases where a population estimate is outside the plausibility range?
- If a contingency adjustment is made to an estimate should a balancing adjustment be made elsewhere within the local authority to retain the total population?
- How will contingency adjustments be reported in the MYE publication?
- How will contingency adjustments be applied at small area level?

3.6 The ONS strategy for making contingency adjustments to population estimates using plausibility ranges is currently being considered by their methodology division. It is expected that the ONS plans will be finalised in early 2011.

4. Measure of uncertainty

4.1 The mid-year population estimates for the council and NHS board areas of Scotland are subject to a degree of uncertainty. In particular, there is uncertainty related to the migration estimates, arising from the data sources used to calculate this component of population change. It may be possible to produce an approximate measure of the uncertainty associated with each component of change. However, this is not an easy thing to do and not something that is done in any other countries as far as we are aware.

4.2 ONS has established a project with the aim of producing an approximate measure of the uncertainty of mid-year population estimates at local authority level in England and Wales. An [Interim Report](#) looked at measuring uncertainty in internal migration. The report identified six key areas that affect the quality of the population estimates:

- Evidence of differential or longer time lags when re-registering with a GP.
- The impact of differences between the 2001 population estimates and the 2001 population base in the NHS data.
- GP registers not capturing those who move during the year, where they are not registered at the two mid-year points, e.g. 0 year olds.
- Uncertainty arising from constraining GP register data to NHSCR figures.
- Errors in estimates of migration to the rest of UK and allocation to local authorities.
- Double counting of moves of school boarders.

4.3 A later ONS paper, '[Quality Measures for Population Estimates](#)', describes a possible long-term approach to measuring the uncertainty in population estimates. The paper explains the proposed approach as follows:

'... the methodological approach being taken in the UK is to map out and describe the source of uncertainty associated with each of the data sources and methods used, estimate the uncertainty for each quality issue and derive an error distribution for each component of change for each year at local authority level. Values can then be randomly generated (or simulated) using the error distributions and combined into an overall estimate of error. This simulation is then repeated a large number of times to produce a range of composite error estimates for each local authority.'

- 4.4 ONS has concluded that the simulation-based methodology described in the paper is a feasible long-term approach to producing measures of uncertainty. However, the production of composite error measures is technically difficult and will take some time to complete.
- 4.5 It is probable that the approach being developed by ONS could be adapted for use in Scotland, but at the moment it is too early to say whether GROS will be able to produce measures of uncertainty for local authority mid-year population estimates. We will need to wait until ONS publishes the results of their research before we know whether we will be able to use the same methodology.

5. 2011 Census data quality assurance

- 5.1 As part of the downstream processing for the 2011 Census, a number of data quality checks will be carried out on the data collected. The main requirement of the census data quality assurance strategy that is relevant to the mid-year estimates is
'to quality assure the Census data against other sources, involving both quality assurance of variable distributions and demographic characteristics'.
The 2011 Census quality assurance work will be done at local authority and data zone level. It is expected that some of this work could be applied at local authority level for the mid-year population estimates.
- 5.2 A list of checks and comparators (administrative and other data sources) has been compiled for the census data quality assurance. Some of these may not be suitable indicators of the quality of mid-year population estimates – for example, those data sources that are used in the production of the mid-year estimates. Also, there may be data sources that we want to use in the mid-year estimates quality checks that are not used in the census quality assurance.
- 5.3 A Data Quality Management System (DQMS) is being developed to automate much of the census quality assurance work. The system is an Excel/VB package that will identify any unusual results by setting tolerance ranges for the data, using data from other sources as the basis for the tolerances. It is likely that the DQMS package can be adapted to set the plausibility ranges for the mid-year estimate checks.
- 5.4 The work done on the quality assurance of the census data will help us to understand the limitations of the administrative and other data sources and to identify which can be used as quality indicators of the mid-year population estimates.

6. Next steps

- 6.1 This is an early version of a report to be completed and circulated by the end of March 2011. The final version will have more details of the methodology that ONS will use for setting up plausibility ranges and calculating measures of uncertainty. Their proposals are currently with the ONS Methodology team for comment and they expect to get feedback in January 2011.

- 6.2 The work being done by ONS is important in the development of quality indicators for mid-year estimates. GROS needs to be aware of the progress that ONS are making with the development of a methodology for measuring the quality of population estimates.
- 6.3 The downstream processing data quality assurance plans for the 2011 Census will be helpful in the development of quality indicators for mid-year estimates. In particular, we can identify the census checks and comparators that can be used for checking the quality of the mid-year estimates, and the DQMS software for setting tolerance ranges can be investigated for possible use.

Annex – Office for national Statistics (ONS) Links and Papers

Improving Migration and Population Statistics (IMPS) and the Migration Statistics Improvement Programme website.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/index.html>

ONS (2010). Improving Migration and Population Statistics: Quality Measures for Population Estimates.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/mig-stats-improve-prog/comm-stakeholders/improvements-2008-pop-est/key-documents/quality-measures-for-population-estimates.pdf>

ONS (2009). Measuring Uncertainty in the Local Authority Population Estimates – Interim Report Focusing on Internal Migration.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/historical-updates-reports/updates-reports-09/measuring-uncertainty-in-the-pop-estimates---november-2009.pdf>

ONS (2010). Migration Statistics Improvement Programme Annual Overview for 2009/10.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/current-updates-reports/migration-statistics-improvement-programme-annual-overview-2009-10.pdf>

Population Trends 141, Autumn 2010: An analysis of patient register data in the Longitudinal Study – what does it tell us about the quality of the data? Smallwood S and Lynch K.

<http://www.statistics.gov.uk/populationtrends/downloads/poptrends141web.pdf>

ONS (2009). Quality Assurance Strategy for Implementation of Migration and Population Statistics Improvements – 2008 Round.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/mig-stats-improve-prog/comm-stakeholders/reference-groups/quality-assurance-strategy.pdf>

ONS (2009). Interim Report on the Potential Use of Department for Work and Pensions Data to Improve Population and Migration Statistics.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/interim-report-on-the-potential-use-of-dwp-data-to-improve-population-and-migration-statistics.pdf>

ONS (2010). Improving Migration and Population Statistics – Overview of the Package of Improvements.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/mig-stats-improve-prog/comm-stakeholders/improvements-2008-pop-est/key-documents/migration-improvements-overview-of-package.pdf>

ONS (2009). A summary of administrative data sources and their potential to inform statistics on migration and population.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/historical-updates-reports/updates-reports-09/summary-of-admin-data.pdf>

ONS (2009). Feasibility report on the potential use of record level Migrant Workers Scan data to improve migration and population statistics

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/historical-updates-reports/updates-reports-09/initial-feasibility-report---october-2009.pdf>

ONS (2009). Use of School Census data to Improve Population and Migration Statistics.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/historical-updates-reports/updates-reports-09/research-paper-on-the-use-of-school-census-data-to-improve-population-statistics---october-2009.pdf>

ONS (2009). Understanding and measuring uncertainty associated with the mid-year population estimates.

<http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/current-updates-reports/interim-report-on-quality-measures-for-population-estimates.pdf>