Summary Quality Report for Internal Migration

1 Introduction

This report is part of a rolling programme of quality reports being introduced by the Office for National Statistics (ONS). The full programme of work being carried out on Statistical Quality[1] is available on the National Statistics website. Summary Quality Reports are overview notes which pull together key qualitative information on the various dimensions of quality as well as providing a summary of methods used to compile the output.

Internal migration estimates are used by government, academia, the media, special interest groups and the general public. Internal migration is also a key component of local population change and is used in the production of National Statistics Population Estimates and Projections[2].

2 Summary of Quality

2.1 Relevance

The degree to which the statistical product meets user needs for both coverage and content.

ONS produces estimates of internal migration within the UK at the country level and at more detailed levels, namely for Government Office Regions (GOR), Local Authorities (LA) and former health authorities level for England and Wales. These estimates currently represent the best available source of information on movements within the UK and have a direct impact on population estimates for LAs and other sub-national geographies. Moves between England and Wales and from Scotland and Northern Ireland to England and Wales are produced from England and Wales patient register sources. Moves from England and Wales to Scotland and Northern Ireland are produced by the respective receiving country and provided to ONS for inclusion in published estimates. ONS Internal Migration datasets[3] with estimates at the local and former health authority level are available on an annual basis and at country and GOR level (for England and Wales) on a quarterly basis, the latter published in Population Trends[4]. Gender and age group breakdowns of UK flows are also published for the UK constituent countries and at the LA level for England and Wales and these are also published in the ONS Internal Migration datasets[3]. More specialised ad hoc origin/destination tables may be available on request from the Migration Statistics Unit[5].

Because there is no single system to record population movements within the UK, internal migration estimates must be derived from proxy sources. Two administrative data sources are used, namely the National Health Service Central Register (NHSCR) and the Patient Register Data System (PRDS). Both of these sources have different properties in terms of coverage and content but used together they represent the best available.

The NHSCR receives notification when a patient in England and Wales transfers to a new NHS doctor within a different health authority boundary. There is a similar system in both Scotland and Northern Ireland. Weekly counts of these re-registrations are used by ONS as proxy indicators for movements around the UK. Estimates derived from the NHSCR are considered to give the most comprehensive coverage of the population and represent the most reliable indicator of internal migration within the UK. However, they are only based at the broad geographic level of the former health authorities.

Therefore, the PRDS data source is used to estimate internal migration at smaller geographical levels. This system holds lists of all patients and their respective residential addresses registered with each NHS general practice. ONS receives an annual download of each list allowing a comparison of patient's postcode between consecutive years. Migration estimates between postcode areas can subsequently be estimated and geographically aggregated as required. However, as this is an annual count, only one move per patient per year can be captured and some types of moves may not be captured at all, since a patient must be present in each consecutive download (see section 2.2).

Combining the relative strengths of the two sources, ONS is able to produce products that detail estimates of internal migration by age and gender at the local authority level accounting for within and between year moves.

2.2 Accuracy

The closeness between an estimated result and the (unknown) true value.

Despite some limitations (see below), research has shown that there is no one other available source that has as good coverage and quality as the NHS administrative data (NHSCR and PRDS) for estimating internal migration within the UK. Detail of the research can be found in the following papers: A Review of Migration Data Sources[6] and The Determinants of Migration Flows in England: A Review of Existing Data and Evidence[7].

Data from the two sources undergo regular, systematic quality assurance to ensure consistency and plausibility over time. Estimates are compared against time series data and outliers are examined in greater depth by comparing the age and sex distributions against the equivalent distributions for previous time periods. Other administrative data sources are also used to validate the quality assurance process such as the Census in England and Wales[8], Labour Force Survey[9], Higher Education Statistics Agency[10], Universities and Colleges Admission Service[11] and Communities and Local Government Housing Data[12].

Research has shown that for males and females, with the exception of young males (ages 16-36), migration estimates for England and Wales from the patient register data are comparable with those derived from the Census. Further information about this can be found in the Family Health Service Area (FHSA) Implementation Project Report.

Limitations of the data sources

Overall, the accuracy of NHS based data depends on patients re-registering with a new doctor when they change residence or informing their current GP of any change of residential address. It is known that re-registration patterns vary by sex and age group, this is dealt with in more detail in the following reports: FHSA Implementation Project Report, Report on Research into Revising Internal Migration Estimates[13] and the Update Following Further Investigation into Revising Internal Migration Estimates[14]. Research has been undertaken to assess the feasibility and benefits of adjusting the estimates to account for variations but no suitable data source currently exists that would facilitate any improvement to the current methodology.

Data from the NHSCR system are limited to movements between the former health authorities. These former health authorities no longer exist as administrative entities but continue to be used due to technical constraints of the NHSCR processing system. PRDS data are also inadequate as a stand alone source as a number of within year moves are not captured, including:

1. migrant babies aged less than one year old
2. new non-birth registrations e.g. ex-armed forces personnel and international in-migrants who join the NHS and then move within the same year
3. people who move during one year but then leave the NHS register before the end of the second year e.g. the deceased, new armed forces personnel and international out-migrants

Limitation of estimates derived from a combination of NHSCR and PRDS are:

1. there is variation in the delay between a person moving and registering with a new doctor
2. some moves may not result in a GP re-registration and therefore will not be recorded

3. individuals may move and re-register more than once in a single year but remain within the same health authority

2.3 Timeliness and Punctuality

Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the time lag between the actual and planned dates of publication.

NHSCR quarterly rolling year estimates are usually released nine months after the end of the migration period. The Release Calendar[15] provides information on forthcoming dates of release.

Mid-year estimates are usually released 14 months after the end of the migration period and are timed to coincide with the publication of the mid-year population estimates.

2.4 Accessibility and Clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format(s) in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.

Mid-year estimates at geographical levels down to LA, derived from NHSCR constrained PRDS data, are available from 1999 onwards. Tables with breakdowns by age and sex are published annually on the National Statistics website[16]. Confidentiality issues prevent the publication of estimates for areas smaller than the local authority.

Quarterly rolling year migration estimates at geographical levels down to health authority, are available from 1975 onwards. FHSA level migration estimates are available from 1975-1998. Interim FHSA/former health authority coded estimates are available for 1999-2000. Former health authority-level estimates are available from 2001 onwards.

Summary data are published annually in a report in the Autumn Edition of Population Trends[4] and included in Key Population and Vital Statistics[17], Social Trends[18] and Regional Trends[19].

2.5 Comparability

The degree to which data can be compared over time and domain.

Data Collection

From April 1975 to March 1984 migration estimates were based on a 10% systematic sample of the NHSCR re-registrations. Since April 1984 a 100% count has been taken, and each person's age, sex, old former health authority and new former health authority are extracted.

Until 1999, ONS used information from electoral registers and the most recent census to apportion the NHSCR inflows and outflows between the constituent local authorities. From 1999, mid-year estimates at local authority level have been generated from the PRDS data constrained to NHSCR data.

No revisions have been made to the pre-1999 estimates and therefore the current data series is not directly comparable with pre-1999 data at local authority level.

Changes in Migration Statistics Geography

For ONS migration purposes, the United Kingdom comprises England, Wales, Scotland and Northern Ireland.

Until 1998, NHSCR data were based around FHSA. These areas were equivalent to shire counties, metropolitan districts and groupings of London boroughs. In 1996, the geography of health service areas changed from FHSAs to Health Authority (HA) areas. This included some boundary changes, as well as some merging and splitting of different FHSAs to create the new HA areas. Thus data for 1999 and 2000 were based on interim codes comprising a mixture of old FHSA codes and the new HA codes. Since 2001, NHSCR data have been based entirely on these health authority areas. Further restructuring of the NHS geographies

to Strategic Health Authorities (SHAs) since 2001 have not been incorporated into the NHSCR data but the HAs are now referred to as former HAs.

Prior to re-organisation of health authority databases from FHSAs to HAs in 1998 some NHSCR database boundaries were realigned. This resulted in a small number of transfers of patients between databases to fit the new boundaries. For the most part, this movement was done outside the NHSCR system and therefore had no effect on migration data. However, a small number were transferred within the system which resulted in the generation of small numbers of spurious migrants. ONS were advised of the adjustments made by the Department of Health and appropriate amendments were made to the migration estimates for that period.

More information on all of the changes in migration statistics can be found at the English Health Geography [20] page.

Estimation of Migration Date
Prior to 1991, a 3-month time lag was assumed between a person moving and their re-registration by NHSCR. Since computerisation of the NHSCR in 1991, date of acceptance by the receiving HA has been available and therefore a single month's lag is considered more accurate.

NHSCR vs PRDS Constrained to NHSCR
There are two key differences between the published migration estimates from the two data systems.

1. Frequency: NHSCR data are collected weekly, processed and delivered quarterly on a rolling year basis. The PRDS data are produced annually from a mid-year download of records.

2. Age: Quarterly estimates from the NHSCR data indicate age at migration. Migration estimates derived from the patient registers, represent age at mid-year. This is done to support the compilation of mid-year population estimates where age at mid-year is required to estimate a rolled forward population estimate. The differences between NHSCR and PRDS age calculations potentially mean a migrant can appear up to one year older on the PRDS data than on the NHSCR data.

2.6 Coherence

The degree to which data that are derived from different sources or methods, but which refer to the same phenomenon, are similar.

The Census in England and Wales[8] gathers data on population moves through the reporting of address changes between the census year and one year earlier. For males and females, with the exception of young males, migration estimates for England and Wales from the PRDS data are comparable with those derived from the census (see section 2.2). However, internal migration estimates from PRDS are more timely than those of Census.

Data on moves from England and Wales to Northern Ireland and Scotland are provided by Northern Ireland Statistics and Research Agency[21] (NISRA) and the General Register Office for Scotland[22] (GROS) respectively. Invariably, the number of migrants moving to Northern Ireland as recorded by the NHSCR is different from the number of moves to Northern Ireland notified to ONS by NISRA. The consensus is that the country receiving the migrants will have a more accurate count than the country of origin. Consequently, ONS apportions the NHSCR data to take account of the differences in the number of moves recorded by both NHSCR and NISRA. These apportionments are applied to published tables.

3 Summary of Methods Used to Compile the Output

NB: All references to HA in this section refer to former HA's.

Producing Quarterly Rolling Year Estimates

Records of NHS patient moves are extracted from the NHSCR and sent weekly to ONS for processing. Each record contains the old HA, new HA, patient date of birth, and date of acceptance by the new HA. Thus, only patient moves that involve a change in GP and a move from one HA to another are received by ONS. The data are used to produce summary counts of moves into and out of geographical areas by age and sex.

Migration records are also created when patients join the armed forces, emigrate or immigrate from outside of the UK. These records are not included in the internal migration estimates but are accounted for elsewhere in the population estimates. Other records that are also excluded from migrations estimates are moves to long-stay psychiatric hospitals and imprisonment.

Producing Annual Mid-year Estimates

Every NHS authority holds a record of NHS patients registered with GPs in their area comprising NHS number, gender, date of birth, date of acceptance and patient's address. ONS receive an annual mid-year download from each authority's register which, when combined, forms a total patient register for England and Wales.

Comparing registers from consecutive years, linked by NHS number, ONS identifies all individuals whose postcode differs between the two years.

These counts are then constrained to NHSCR HA level data to produce internal migration estimates (see below).

Constraining the PRDS Data to NHSCR Data for Between HA Moves

Constraining PRDS data to NHSCR data, by matching on origin and destination HA, preserves the relationship between origin and destination HA in the constrained dataset.

1. For ages *one and over*

   Calibration factors are created for moves between HA's for all combinations of HA origin/ destination as follows.

   $$\frac{\text{number of migrants in the NHSCR data}}{\text{number of migrants in the PRDS data}}$$

   The appropriate calibration factor is applied to each PRDS between-HA total flow to obtain mid-year flow totals at HA level that agree with the NHSCR data.

2. For ages *under one*

   The PRDS data do not capture any migrants aged *under one*.

   Number of migrants at postcode level for age *under one* is estimated from NHSCR counts of migrants age *under one* using the postcode based distribution of migrants *age one*.

Constraining the PRDS Data to NHSCR Data for Within HA Moves

1. For ages *one and over*

   Single calibration factors for moves *within* each HA are created by taking an average of the two calibration factors created during the constraining process for each HA (into and out of) for between HA moves at *age one and over* (see above).

   The appropriate average calibration factor is then applied to each within HA total flow to obtain within-HA level mid-year flows constrained to NHSCR data.

2. For age *under one*

   PRDS data are constrained by the following ratio.

$$\frac{\text{total number of moves } \textit{within} \text{ HA age } 1^{c}}{\text{total number of moves } \textit{between} \text{ HA age } 1^{a}} = \frac{\text{total number of moves } \textit{within} \text{ HA age } 0^{d}}{\text{total number of moves } \textit{between} \text{ HA age } 0^{b}}$$

   Where totals number of moves[a] are constrained PRDS values for *age one*.

   Total number of moves[b] are NHSCR values for *age under one.*

   Total number of moves[c] are constrained PRDS values for *age one within* HA.

   Total number of moves[d] are Total number of moves[c] multiplied by the following ratio.

$$\frac{\text{total number of moves } \textit{between} \text{ HA age } 0}{\text{total number of moves } \textit{between} \text{ HA age } 1}$$

Rounding

Mid-year estimates published on the ONS website are rounded to the nearest 10 and row and column totals are rounded to the nearest 100.

Statistical Disclosure Control

Statistical disclosure control methodology is also applied to data. This ensures that information attributable to an individual organisation is not disclosed in any publication. The National Statistics Code of Practice, and specifically in the Protocol on Data Access and Confidentiality set out principles for how we protect data from being disclosed. The Protocol includes a guarantee to survey respondents that "no statistics will be produced that are likely to identify an individual unless specifically agreed with them". More information can be found on the ONS Statistical Disclosure Control Methodology[23] page.

4 References

	Title of Reference	Website location
1	Statistical Quality Programme	http://www.statistics.gov.uk/about/data/methodology/quality/default.asp
2	National Statistics Population Estimates and Projections	http://www.statistics.gov.uk/CCI/nscl.asp?ID=7595
3	ONS Internal Migration datasets	http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=7070&Pos=&ColRank=1&Rank=128
4	Population Trends	http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=6303&Pos=&ColRank=1&Rank=422
5	Migration Statistics Unit	migstatsunit@ons.gsi.gov.uk
6	A Review of Migration Data Sources. *Population Trends 66, pp47-47. (1991)*	Office for National Statistics (ONS), Quarterly, ISBN=9780230526136, ISSN=03074463, On-line edition, First/News/Press Release, Hardcopy Publication.
7	The Determinants of Migration Flows in England: A Review of Existing Data and Evidence	http://www.communities.gov.uk/archived/general-content/housing/225030/determinants/
8	Census in England and Wales	http://www.statistics.gov.uk/census/default.asp
9	Labour Force Survey	http://www.statistics.gov.uk/STATBASE/Source.asp?vlnk=358
10	Higher Education Statistics Agency	http://www.hesa.ac.uk/
11	Universities and Colleges Admission Service	http://www.ucas.com/figures/index.html
12	Communities and Local Government housing data	http://www.communities.gov.uk/housing/housingresearch/housingstatistics/housingstatisticsby/locallevelstatistics/
13	Report on Research into Revising Internal Migration Estimates	http://www.statistics.gov.uk/downloads/theme_population/Revising_Internal_Migration_Estimates.pdf
14	Update Following Further Investigation into Revising Internal Migration Estimates	www.statistics.gov.uk/downloads/theme_population/Internal_Migration_Est_Report.pdf
15	Release Calendar	www.statistics.gov.uk/ReleaseCalendar/currentreleases.asp
16	National Statistics website	http://www.statistics.gov.uk/
17	Key Population and Vital Statistics	http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=539&Pos=&ColRank=1&Rank=272
18	Social Trends	http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=5748&Pos=&ColRank=1&Rank=422
19	Regional Trends	http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=14356&Pos=&ColRank=1&Rank=422
20	English Health Geography	http://www.statistics.gov.uk/geography/england_health.asp
21	Northern Ireland Statistics and Research Agency	www.nisra.gov.uk/
22	General Register Office for Scotland	www.gro-scotland.gov.uk/
23	Statistical Disclosure Control Methodology	http://www.statistics.gov.uk/about/data/methodology/general_methodology/sdc.asp

Date of publication: December 2007