

Quality Assurance of Administrative Data used in population statistics: National Health Service Central Register

Published on 24 November 2022



Contents

1. Overview of National Health Service Central Register.....	3
2. Data source information	3
3. Operational context and administrative data collection – How is the data extracted and processed?.....	5
4. Communication with data supply partners	6
5. Quality assurance principles, standards and checks applied by data suppliers...	6
6. Quality assurance investigations carried out by the NRS Population and Migration Statistics team.....	7
7. Strengths and Limitations of the data source	8
8. Risk/Profile Matrix.....	8
9. Useful links and contact details.....	10

1. Overview of National Health Service Central Register

Data supplier:	National Health Service Central Register (NHSCR)
Supplier info:	<p>The NHSCR is an electronic demographic database of all people born in Scotland, those who died in Scotland, and those who have ever registered with a GP in Scotland. The NHSCR Team is a branch of National Records of Scotland (NRS) and is responsible for maintaining the register.</p> <p>NRS is a Non-Ministerial Department of the Scottish Government. The purpose of NRS is to collect, preserve and produce information about Scotland's people and history and make it available to inform current and future generations.</p>
Data type (counts or unit records)	Anonymised unit records
Data content:	Anonymised extract of live NHSCR system. Variables include demographic information and NHS administrative information.
Supply schedule:	NHSCR data is extracted monthly and processed quarterly.
Use of data:	<p>Mid-year population estimates</p> <p>Small area population estimates</p> <p>National population projections</p> <p>Sub-national population projections</p> <p>Population estimates of the very old (Centenarians)</p> <p>Migration Statistics</p>
Matrix Score	A1 – Basic Assurance
Last review date	24 November 2022

2. Data source information

The NHSCR (National Health Service Central Register) contains basic demographic details of everyone born in Scotland, everyone who died in Scotland, plus anyone else who is (or has been) on the list of a general medical practitioner in Scotland. It is held as an electronic database. The main purpose of the Register is to allow the smooth transfer of medical records of patients who move between Health Board areas (or across borders within the UK). It is also used by the NRS Population and Migration Statistics (PAMS) Branch, for creating estimates of population and migration.

Two patient registers are used in the creation of population estimates; the National Health Service Central Register (NHSCR) and the Community Health Index (CHI). The NHSCR for Scotland records the movement of patients to and from NHS Board

areas within Scotland. The CHI is a register of all NHS patients in Scotland. Both datasets record new registrations from the rest of the UK and from outside the UK.

How is NHSCR data used to estimate migration within Scotland?

National Records of Scotland (NRS) use the NHSCR to estimate moves between NHS Board areas, which are considered the best estimates of internal migration within Scotland. Unlike the NHSCR, the CHI contains more detailed information which allows for the distribution of migration to smaller geographies including council areas and data zones.

[NRS QAAD for CHI in Scotland](#)

Quarterly extracts of the NHSCR are compared and records with different postcodes or different NHS Board codes are retained, processed and aggregated for the reference year. These form the control for the total migration flows within Scotland.

To estimate migration for smaller geographies, extracts of CHI data are compared on two occasions one year apart. These estimates are then constrained to the figures at NHS Board level which come from the NHSCR.

How is NHSCR data used to estimate migration between UK Countries?

For moves between Scotland and the rest of the UK, the receiving country is deemed to have the best data. This is due to a known issue where some people do not de-register with their GP surgery when moving (so can be seen as passive) and registration requesting access to services is an active event. The other UK countries then agree to the totals to ensure consistency.

Therefore, NRS determine the inflows from the rest of the UK to Scottish NHS Board areas based on registration data from the NHSCR, while the Office for National Statistics (ONS) and Northern Ireland Statistics and Research Agency (NISRA) determine the outflows from Scotland. A summary is shown in the matrix below.

Flows are sent in the form of data on single year of age, sex, Previous Location ID, New Location ID, and Quarter/Year. Reports on the quality of similar administrative dataset used across the UK are available from:

- [ONS QAAD for the Patient Register in England and Wales \(Dec 2016\)](#)
- [NISRA Quality report on data quality for Northern Ireland \(June 2016\)](#)

How is NHSCR data used to estimate migration from outside the UK?

When a person from outside the UK first registers with an NHS doctor in Scotland, an indicator is used to denote this. GP registration data is then collected by the NHSCR and the CHI. This indicator is used as the basis to extract data for age, sex and area distribution within Scotland for migration from countries outside the UK.

NHSCR records hold limited data on international outflows from NHS Board areas to countries outside the UK, unless people inform their GP that they are leaving the UK. International outflows are allocated to NHS Board areas based on an average proportion of :

- International inflows from two years prior to the reference date

- Outflows to the rest of the UK
- Population share of each NHS Board area

The analysis produces gross numbers of people who have moved within, into and out of Scotland in the course of the year to 30 June (mid-year) of the reference period. These are then used to calculate the population for an area based on changes in births, deaths and migration.

3. Operational context and administrative data collection – How is the data extracted and processed?

The NHSCR is confidential, accessible only to a list of bodies approved by the Scottish Parliament and set out in the Local Electoral Administration and Registration Services (Scotland) Act 2006 (LEARS Act). The act and its associated schedules set out the information that the Registrar General can hold in the NHSCR and who this information can be shared with. The information provided to these bodies is the minimum necessary for them to do their work.

A snapshot of the NHSCR database is taken by Atos (NHS National Services Scotland [NHS NSS] IT suppliers) at the end of each month.

Each month Atos provide a cut-down (and anonymised) version of the full NHSCR extract for the latest month to the NRS Population and Migration Statistics (PAMS) Branch. This extract is made up of three datasets – Person Information, Health Board Information and Postcode Information, each with its unique ID and with a PersonID common to each dataset for linking. IDs are allocated to new records sequentially as they are created.

- Health Board Table is used to identify the new NHS Board and when it happened.
- Person Table is used to aggregate demographic information on individuals and identifies new immigrants. It also identifies deaths; these records are excluded.
- Postcode Table is used to link the postcode details to the person.

Each month NRS IT access the data through Atos via Secure File Transfer Protocol (SFTP) server, download, process and anonymise it. NRS IT save the files onto a shared secure drive for use by the Population and Migration Statistics team. These files are processed quarterly to produce summarised figures by age and sex for all moves within, as well as to and from, Scotland. Extracts at the end of one quarter are compared to extracts at the end of the previous quarter, and new Health Board/postcodes are identified and previous Health Board/postcodes determined. Records relating to new-borns are removed (as they are new to the database, but have not moved and will be counted in the births) and age at mid-year is calculated. Once all four quarters are available and processed, annual totals for all destinations by single year of age, sex and origin are calculated. As the receiving country's estimates are assumed to be most reliable, NRS use the data received from ONS for moves to England and Wales, and NISRA for moves to Northern Ireland to constrain

the estimates from the NHSCR to England, Wales and Northern Ireland, by five year age group and sex.

A series of quality checks are made to ensure the data fits with normal patterns. Any anomalies are referred back to the NHSCR Team.

NHSCR moves are used as the control at Health Board level. The CHI extract, which provides a postcode level set (by single year of age and sex) of all postcode changes in the reference year, is then used to allocate moves down to Data Zone level.

A lag of two months is assumed between when a person moves and registers the move with a GP. Therefore, the four quarters ending August of the reference year make up the annual mid-year figures used in the population estimates.

4. Communication with data supply partners

Governance of the NHSCR is a function of the NHSCR Stakeholder Reference Group which meets twice a year to provide strategic advice on the operation, future development and external relations of the NHSCR in support of the NHS in Scotland, medical researchers and other users. Membership includes NHSCR, NRS (Director, Statistical Services), Practitioner Services Division (Part of NHS NSS), Scottish Government (E-health), the Improvement Service, University of Aberdeen (Information Governance Advice) and the CHI Advisory Board.

Extracts of NHSCR data are produced by Atos, the IT supplier for NHSCR and supplied to the Population and Migration Statistics team in NRS via approved NRS data transfer procedures.

The PAMS team are in regular contact with the NHSCR team and discuss any changes of issues that might affect the data. These close working relationships ensure that issues and planned changes to data collection can be considered before the data is used to produce population estimates.

The [Population and Migration Statistics Committee \(Scotland\)](#) meets twice a year. This group includes representatives from local authorities and other organisations with an interest in demographic statistics. NRS engage and seek feedback on statistical outputs, changes in methodology and data sources at these meetings.

5. Quality assurance principles, standards and checks applied by data suppliers

When NHSCR receives data extracts they are automatically matched to the database. Checks on variable values are coded into the matching process. Any records that do not match are verified manually by NHSCR staff and enquiries are sent to the data providers if anomalies are present.

The data entered by NHSCR staff is regularly scrutinised. Supervisors check 5% of the work undertaken by staff to identify any potential training issues. These records are randomly selected based on subject matter, taking into account new areas of

work, trends or concerns previously identified. This also helps the NHSCR to meet its service level agreement with the Scottish Government, NHS National Services Scotland which requires an accuracy level of 97%, which is currently being achieved.

As well as this, the NHSCR team undertake a variety of data quality initiatives on an annual/bi-annual basis where staff investigate the population of different variables in the register and correct duplicates. These initiatives are carried out relatively frequently as they target areas of known concern and the findings are generally kept internal to the NHSCR team. These data quality initiatives include:

- investigating records where no death has been recorded for a person aged over 110 years old. In the majority of cases a death is traced (these are usually deaths that were missed at the time, usually from the 1970s or 1980s before the NHSCR was computerised) and the record is updated to reflect this.
- checking records where the NHS Board variable is blank. This allows us to be confident that all records have a valid posting attached (i.e. a valid NHS Board or special code). Where no posting exists it is usually for people who were born in Scotland but never registered with a Scottish NHS GP.
- populating records that do not have a Community Health Index (CHI) number - either with the CHI number if one exists, or with a flag to show that there is not a CHI number for that record. [Information on CHI numbers.](#)

NHSCR also collects feedback from users of the NHSCR extracts where anomalies are identified and investigates these anomalies so a resolution or explanation can be found.

6. Quality assurance investigations carried out by the NRS Population and Migration Statistics team

The data are checked for changes to the previous month and how they compare to the same month in previous years. Data are then checked for missing values using a series of programs. The outputs include:

- the number of NHSCR person, posting and postcode records by month
- the percentage of person records with missing values by variable and month
- the percentage of posting records with missing values by variable and month
- the percentage of postcode records with missing values by variable and month
- the number of postcode records by source ('1' = NHSCR, '2' = Local Authority, '3' = CHI) and month

If there are any issues with the extract, PAMS contact NHSCR to find out if the work that they have done on the NHSCR database may have affected the extract.

7. Strengths and Limitations of the data source

Strengths	Limitations
<ul style="list-style-type: none"> NHSCR is a comprehensive source of record level data that covers the vast majority of the population. 	<ol style="list-style-type: none"> It does not pick up people who leave the UK (unless they informed their GP), leading to some inflation in the register. For this reason additional sources are used to calculate international outflows. <ul style="list-style-type: none"> Moves within Scotland cannot be picked up until the patient registers with a new GP. As a result, some people will be recorded in the wrong area. This is particularly an issue among younger adult males. There will be a lag in recent migrants into Scotland appearing on the NHSCR as they will only appear when registering with a GP. A two month lag is assumed between a move and registration of the move with a GP. There is a delay in new born babies appearing in NHSCR with a postcode (and posting) until they are registered with a GP.

8. Risk/Profile Matrix

This section contains a risk/profile matrix for the NHSCR. The matrix reflects the levels of risk of data quality concerns and the public interest profile of the statistics. These have been determined by a review undertaken by the NRS Population and Migration Statistics team using the information contained within the UK Statistics Authority's [Administrative Data Quality Assurance Toolkit](#). For the use of data for this project, the cell highlighted is appropriate:

Level of risk of quality concerns	Public interest profile		
	Low	Medium	High
Low	Statistics of low quality concern and low public interest. [A1]	Statistics of low quality concern and medium public interest. [A1/A2]	Statistics of low quality concern and high public interest. [A1/A2]
Medium	Statistics of medium data quality concern and low public interest. [A1/A2]	Statistics of medium quality concern and medium public interest. [A2]	Statistics of medium quality concern and high public interest. [A2/A3]
High	Statistics of high data quality concern and low public interest. [A1/A2/A3]	Statistics of high quality concern and medium public interest. [A3]	Statistics of high quality concern and high public interest. [A3]

Justification for Matrix Score

The public interest profile has been set to “High” because population and migration statistics are used for the following purposes:

- Resource allocation by central and local government
- Planning of services such as education and health
- Informing local and national policy
- Modelling the economy
- These statistics are incorporated into most other government statistics as denominators, including health statistics.
- Population estimates are also used to measure progress of the population indicator in Scotland’s National Performance Framework

The risk of quality concerns has been set to “low” for the following reasons:

- The service level agreement to have at least 97% accuracy ensures a high level of quality and low risk.
- The risk is further reduced as the NHSCR team have a variety of data quality initiatives that are undertaken on a regular basis to mitigate these data quality issues.
- The NHSCR team and the PAMS team both fall in the Statistical Services division of NRS and both report to the same Director. This means that there is an increased awareness of issues and the impact this could have on the other party. We can therefore be confident that we will be made aware of any changes that would have an impact on how this data is used.

- List inflation is a known issue and therefore can be accounted for when using the data.

9. Useful links and contact details

You can access detailed statistics that we have produced from the [Statistics](#) section of our website. Scottish Census statistics are available on the [Scotland's Census](#) website.

We also provide information about [future publications](#) on our website. If you would like us to tell you about future statistical publications, you can register your interest on the Scottish Government [ScotStat website](#).

You can also follow us on twitter [@NatRecordsScot](#)

Enquiries and suggestions

Please contact our Statistics Customer Services if you need any further information. Email: statisticscustomerservices@nrscotland.gov.uk

© Crown Copyright

You may use or re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. Further information is available within the [Copyright and Disclaimer section](#) of the National Records of Scotland website.