



General Register Office  
*for*  
SCOTLAND  
*information about Scotland's people*



THE NATIONAL  
ARCHIVES OF SCOTLAND

## USING NHSCR FOR STATISTICAL PURPOSES

July 2011

## NATIONAL RECORDS OF SCOTLAND

preserving the past; recording the present; informing the future

1. This paper updates the Board on progress made by National Records of Scotland (NRS) with the use of NHSCR data for statistical purposes. It is for information.
2. Between February 2009 and March 2011 we made seven data extracts for statistical purposes. Each of these was analysed to identify the nature and extent of anomalies in the data. The results of these analyses showed the continuous improvement over the period in the data quality, due to the data cleaning and improvement which is being continuously undertaken by the NHSCR staff.
3. The overall count of the population registered on the NHSCR continues to exceed our population estimates for Scotland as a whole by around 20%. This is partly due the presence of duplicate records, partly to deaths not being registered on the database and partly to people leaving Scotland (especially if they go abroad) without deregistering from their GPs. Investigations have shown that the first of these is not a significant factor and that the rate of duplication is well under one half of one percent. The effect of the second should decrease over time with improvements in the way that death registrations are handled. The third however is likely to be an ongoing problem as it is difficult to see any action which NHSCR could take to address it.
4. One important application of the extract is in improving the quality of migration statistics. It is now possible to compile migration matrices at Health Board level within Scotland using the extract though, for figures relating to moves to and from other parts of the UK, it is necessary to wait for two months to be sure that the moved have been entered and that the data are accurate. For moves from one Health Board to another within Scotland, four weeks is probably sufficient. The next step is to produce migration figures at local authority level using postcodes and work done on this so far suggests that it will be possible to do this to an acceptable degree of accuracy despite the existence of a number of postcodes which cannot be validated and allocated to Local Authorities. We have not completed the work required to fully assess the impact of changing the current migration estimation methodology by using the extract and this will be necessary before any such move is made. To enable such assessment we have introduced additional regular monthly deliveries of reduced anonymised extracts specifically for migration purposes. This arrangement is now in operation and is working satisfactorily.
5. Some record linkage work between NHSCR and other sources was also undertaken. One of these explored the GP registration patterns for students from Scotland enrolled for Higher Education (HE) courses in England and Wales. We used data collected by Higher Education Statistical Agency (HESA) which contains the term-time postcode for students domiciled in Scotland but studying at institutions in England or Wales. By relating the current term time postcode to the year of study in each case, it was possible to investigate the extent to which, and the speed with which, students register with GPs in the towns and cities in which they are studying. The results indicate that for a sizeable minority the NHSCR does not capture the residential move in the first year of study. This proportion gradually declines over the full course of study. A presentation on the results from this work is planned for the Population and Migration Statistics (PAMS) users conference which NRS is holding in October.

6. We have continued to progress quality improvement work for the Scottish Longitudinal Study (SLS). The SLS uses the NHSCR as a 'spine' to link census and civil registration records for a sample of the Scottish population. The extract has been used to identify groups of individuals known to have been missed from the SLS so that their details can be added to make the sample more complete and up to date.
7. The most important use of the extract in the coming years will be the development of a population spine as part of the Beyond 2011 project in Scotland. An important part of this will be the linkage of the 2011 Census in Scotland to the extract as it reflected the population of Scotland on Census Day. Plans are currently being formulated to undertake this which will be a substantial computing exercise but which, if successfully completed, would be important both in the development of census alternatives and in the quality assurance of the 2011 Census.
8. The security arrangements for transferring the data to Ladywell House are under constant review and were upgraded in March 2011 by moving to a more secure storage medium for the transfer. They will be upgraded again in the near future by adopting Secure File Transfer Protocol for electronic transfer of the data to the servers at Ladywell House. This will remove the current requirement for a courier to travel to the ATOS Origin site in Livingston and reduce security risks.
9. Our work over the last two years on the NHSCR extract has been very encouraging and has shown that the NHSCR has great potential, probably unmatched by any other public data set, to play a fundamental part in the development of a more efficient system for the production of population statistics in Scotland. We see huge potential for work on population statistics using this data and our plans for future application continue to develop.