Population And Migration Statistics (PAMS) Committee (Scotland)

Improving migration estimates using patient registers

Introduction

1. The purpose of this paper is to demonstrate that our migration estimates can be improved by linking the Community Health Index(CHI) postcode (provided to National Records of Scotland (NRS) by NHS) to a demographic extract[1] of the National Health Service Central Register (NHSCR). This would allow more robust estimates of migration to be produced for different geographies, than the current system which relies on the (PNP[2] and SMIG[3]) migration files received from the NHSCR.

2. It has been our intention to move towards this change based on postcode rather than higher level health board data but the (currently on going) implementation of the April 2014 NHS Board area boundary changes has accelerated and necessitated this change.

The views of PAMS members on this methodology for the mid-2015 population estimates onwards are welcome.

3. The main advantages of using the NHSCR and CHI Postcode are that:

- Scottish health board posting codes and lower geographies can be derived from postcode,

- historic moves (i.e. moves that occurred on or before the 2011 Census) can be excluded (using the variables 'postingcodedate' and 'PCstartdate', i.e. health board posting start date and postcode start date - proxies for date of move) and

- age at mid-year can be calculated directly from date of birth (rather than relying on a calculation for which the code is not available).

4. The first point is particularly important given that from October 2014 to February/March 2015 the health board posting codes on the NHSCR extract (and consequently the PNP and SMIG files) correspond to a mix of pre-April 2006 and April 2014 NHS Board areas because of the implementation of the April 2014 NHS Board area boundary changes.

5. For more details see Annex A in the paper PAMS (14) 16.

Footnotes

1) A limited extract from the NHSCR which includes only key variables such as NHS number, date of birth, gender and health board posting.

2) The PNP file contains NHSCR changes between health boards, i.e. all moves between health boards by age and gender including cross-border flows. These moves are not linked to the CHI postcode data which is owned by the NHS.

3) The SMIG file contains additional NHSCR variables (such as date of birth) but only for moves to Scotland from the rest of the UK (excluding Northern Ireland).

Health board postings on the NHSCR extract

6. Health board postings on the NHSCR extract provide information on the NHS Board area in which a person who is registered with a General Practitioner (GP) is resident. The NHSCR extract also provides information on the date when the person registered with the GP which can be used as a proxy for the date that the person moved to the area.

7. In most cases new health board postings are created when a person:

- transfers from a GP in a Scottish NHS Board area to a GP in a different Scottish NHS Board area (within Scotland migrants),

- transfers from a GP in England & Wales or Northern Ireland, from overseas or from an armed forces GP to a GP in a Scottish NHS Board area (rest of UK/overseas/Armed Forces to Scotland migrants), or

- transfers from a GP in a Scottish NHS Board area to a GP in England & Wales or Northern Ireland, to overseas or to an armed forces GP (Scotland to rest of UK/overseas/Armed Forces migrants).

Postcodes from CHI

8. Postcodes used from the CHI system refer to the home postcode of the patient. Data is also provided on the date when the person informed their GP that they were resident at that postcode which, again, can be used as a proxy for the date that the person moved to the area.

9. Postcode data is shared by NHS National Services Scotland (NSS) with NRS NHSCR and under that agreement NHSCR staff hold copies of the CHI Postcode alongside the actual NHSCR which they may use in a limited way for example to help tracing patients for health board postings. The CHI postcode is not part of the NHSCR. The CHI Postcode data used by NRS is forwarded alongside the actual NHSCR data under a separate agreement with NHS and again may be used for statistical purposes limited by that agreement.

10. Further to this, NRS also receive a separate annual snapshot of CHI data including postcode as well as other data not linked to NHSCR directly from NHS NSS. This is currently used for annual sub-health board level migration estimates. Our next step is to reconcile this work with the data sourced from NHSCR linked to the CHI postcode described in this paper with a view (if possible) to moving to using a single source for all of this work.

11. We have recently undertaken a consultation including the proposal to add the CHI postcode to the NHSCR system for which NRS is responsible. This will greatly simplify our data management and make administration of the data and the associated agreements much more efficient.

12. New postcodes are created when a person informs their (existing or new) GP that they have changed postcode (within Scotland) or when a person first registers with a GP in Scotland. New postcodes are not created when a person moves out of Scotland or when a person returns to their previous postcode having moved out of Scotland.

13. In any given period there should be more new postcodes than new health board postings as new postcodes are created when people stay registered with a GP in the same NHS Board area when they move postcode (mostly within NHS Board area migrants) whereas new health board postings are not.

Method

14. We receive NHSCR and the linked CHI postcode data monthly. The NHSCR and CHI systems exchange data on a daily basis. Cross border flows between the constituent parts of the UK are agreed quarterly. Therefore, we need to process the data quarterly.

15. New health board postings[4] and postcodes[5] added in the current quarter are identified by comparing data at the end of the previous quarter with that at the end of the current quarter.

16. Evidence suggests that there is a two month lag between a person moving and that person registering with a GP, so for Q3 we compare data at the end of August and November, for Q4 we compare data at the end of November and February, for Q1 we compare data at the end of February and May and for Q2 we compare data at the end of May and August.

17. All new health board postings/postcodes added in the current quarter are compared with the previous health board posting/postcode (which could also have been added in the current quarter) to determine the origin and destination health board postings/postcodes.

18. In cases where a new health board posting/postcode has been added and there is no previous health board posting/postcode from within the UK then the person is assumed to be an immigrant from overseas registering for the first time with a GP in Scotland.

19. Most of the remaining cases correspond to new born babies first appearing on the NHSCR. These cases are matched with birth registrations data and in cases where the first postcode in the CHI data is different from the postcode at birth then the first postcode is replaced with the postcode at birth. The births component in the mid-year population estimates is derived from postcode at birth. This step ensures that the population estimates are not overestimated in certain areas because of differences between the first postcode in the CHI data and postcode at birth.

20. Cases where the origin and destination health board postings/postcodes are the same or the destination health board posting/postcode start date is on or before 27 March 2011 (the date of the 2011 Census) are discarded. The latter correspond to historic moves that occurred on or before the 2011 Census.

Footnotes

4) Health board postings corresponding to registration with a GP in a Scottish, English, Welsh or Northern Irish NHS Board area, registration with an Armed Forces GP or a move overseas are the only ones used to estimate migration. All other health board postings are discarded.

5) Non-Scottish postcodes are discarded.

21. The new postcodes are then split into those corresponding to moves within Scotland and those corresponding to moves to Scotland.

22. Within Scotland moves are estimated from only new postcodes corresponding to moves within Scotland. New health board postings corresponding to within Scotland moves are not used to estimate within Scotland moves.

23. Moves to/from Scotland are estimated by age, sex and origin/destination health board posting from new health board postings to/from Scotland. However, the destination/origin within Scotland is derived from postcode information.

24. For moves to Scotland for which there is a corresponding new postcode, the destination postcode is the destination postcode for the move to Scotland. If there is no corresponding new postcode then it is assumed that the person has returned to their previous postcode in Scotland and the most recent postcode corresponding to this person is the destination postcode for the move to Scotland.

25. Recall that new postcodes are not created for moves from Scotland. Therefore, for moves from Scotland the most recent postcode corresponding to this person is the origin postcode for the move from Scotland. Though new postcodes are not created for moves from Scotland, if a person moves within Scotland before moving from Scotland then the destination postcode for their move within Scotland is the origin postcode for their move from Scotland.

26. There are some issues with the postcode information: large user postcodes which cannot be mapped to small user postcodes, invalid postcodes and missing postcodes. Postcodes that cannot be mapped to small user postcodes, invalid or missing postcodes are imputed as part of the processing. The imputation methodology is detailed, and its effectiveness assessed, in Annex A.

27. At this stage postcodes are only imputed to allow us to determine the NHS Board area of origin or destination. We will continue to use postcode information directly from the CHI to estimate migration for lower geographies (e.g. Council areas and data zones) until further work can be done to assess the quality of the postcode information linked to NHSCR for estimating migration for lower geographies. In theory these should be very closely matched but there are a number of potential factors which could cause variation.

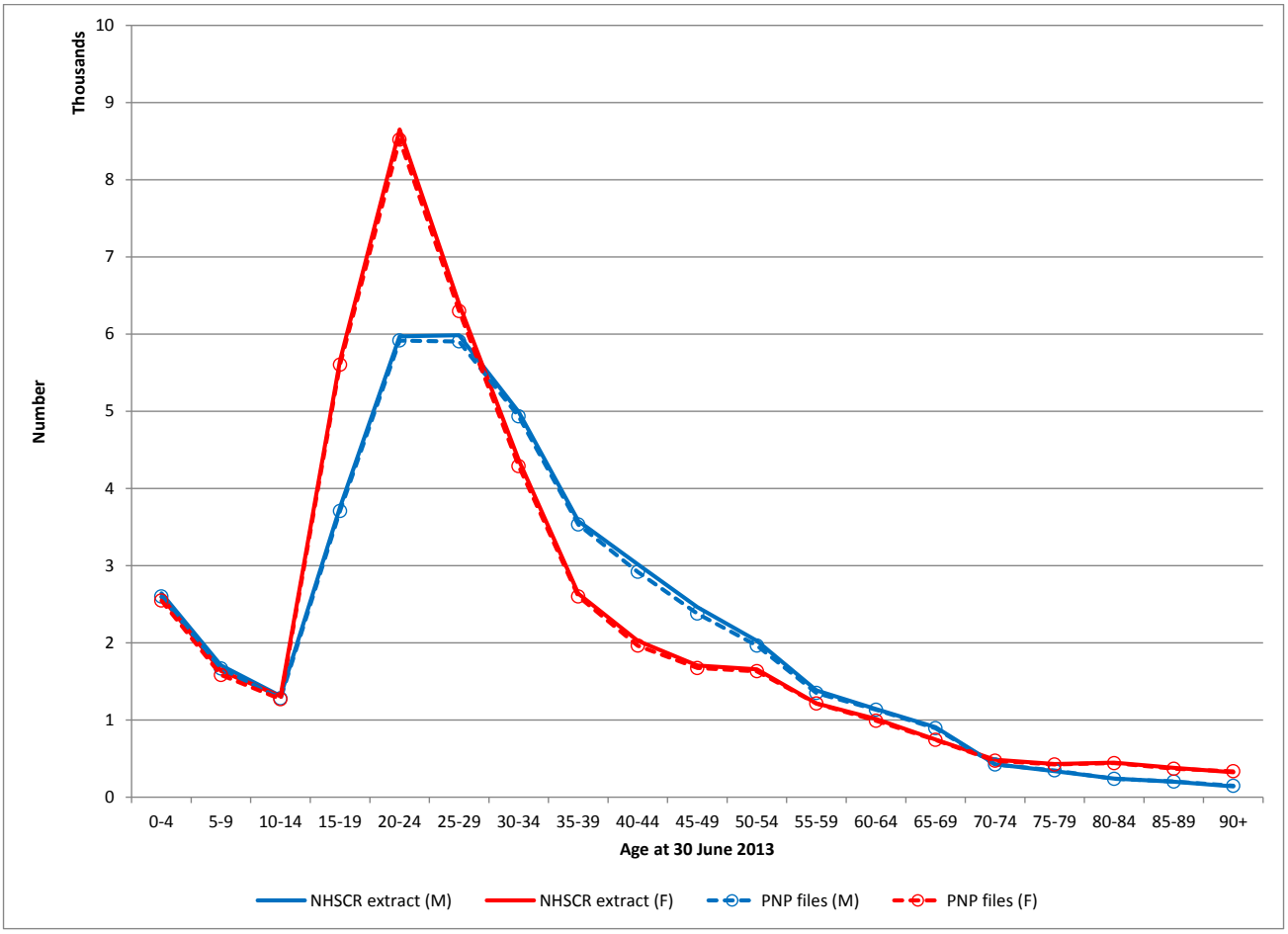
Within Scotland moves

Table 1: within Scotland moves between pre-April 2006 (15) NHS Board areas in the year to mid-2013 by quarter

Period	NHSCR Extract	PNP files	NHSCR Extract minus PNP files	
			n	%
2012 Q3	27,055	26,731	324	1.2
2012 Q4	18,282	17,740	542	3.1
2013 Q1	19,167	18,975	192	1.0
2013 Q2	21,352	21,016	336	1.6
All	**85,856**	**84,462**	**1,394**	**1.7**

28. Table 1 compares the total number of within Scotland moves between pre-April 2006 (15) NHS Board areas in the year to mid-2013 by quarter estimated using the NHSCR extract and the linked CHI postcodes (for brevity we use the term 'NHSCR extract' in the tables and charts in this section and below) and the PNP files. Table 1 shows that the NHSCR extract estimates 1.7 per cent more within Scotland moves than the PNP files.

Figure 1: within Scotland moves between pre-April 2006 (15) NHS Board areas in the year to mid-2013 by 5-year age group



29. Figure 1 compares the total number of within Scotland moves between pre-April 2006 (15) NHS Board areas in the year to mid-2013 by 5-year age group for males (M) and females (F) estimated using the NHSCR extract and the PNP files. Figure 1 shows that there is good agreement between the estimates for males and females in each age group.

Table 2: within Scotland moves between pre-April 2006 (15) NHS Board areas in the year to mid-2013 by NHS Board area

NHS Board area	IN				OUT				NET		
	NHSCR extract	PNP files	Diff[1]	Diff (%)	NHSCR extract	PNP files	Diff[1]	Diff (%)	NHSCR extract	PNP files	Diff[1]
Tayside	7,538	7,475	63	0.8	7,583	7,389	194	2.6	- 45	86	-131
Greater Glasgow	15,678	15,541	137	0.9	16,381	16,157	224	1.4	- 703	- 616	- 87
Grampian	8,335	8,283	52	0.6	7,245	7,168	77	1.1	1,090	1,115	- 25
Forth Valley	6,273	6,174	99	1.6	5,859	5,747	112	1.9	414	427	- 13
Dumfries & Galloway	1,748	1,734	14	0.8	1,954	1,928	26	1.3	- 206	- 194	- 12
Highland	4,041	3,974	67	1.7	4,475	4,402	73	1.7	- 434	- 428	- 6
Shetland	356	351	5	1.4	440	436	4	0.9	- 84	- 85	1
Orkney	446	429	17	4.0	423	409	14	3.4	23	20	3
Western Isles	503	499	4	0.8	571	572	- 1	-0.2	- 68	- 73	5
Fife	5,939	5,734	205	3.6	6,046	5,853	193	3.3	- 107	- 119	12
Ayrshire & Arran	5,182	5,107	75	1.5	5,066	5,008	58	1.2	116	99	17
Argyll & Clyde	6,936	6,835	101	1.5	7,378	7,298	80	1.1	- 442	- 463	21
Lothian	12,504	12,328	176	1.4	12,383	12,264	119	1.0	121	64	57
Borders	2,266	2,150	116	5.4	2,195	2,157	38	1.8	71	- 7	78
Lanarkshire	8,111	7,848	263	3.4	7,857	7,674	183	2.4	254	174	80
Total	**85,856**	**84,462**	**1,394**	**1.7**	**85,856**	**84,462**	**1,394**	**1.7**	**-**	**-**	**-**

1. NHSCR extract minus PNP files

30. Table 2 compares within Scotland moves between pre-April 2006 (15) NHS Board areas in the year to mid-2013 by NHS Board area (in-migration, out-migration and net migration) estimated using the NHSCR extract and the PNP files. Table 2 shows that the differences in net within Scotland migration (between NHS Board areas) estimated using the NHSCR extract and the PNP files range from -131 in NHS Tayside to +80 in NHS Lanarkshire.

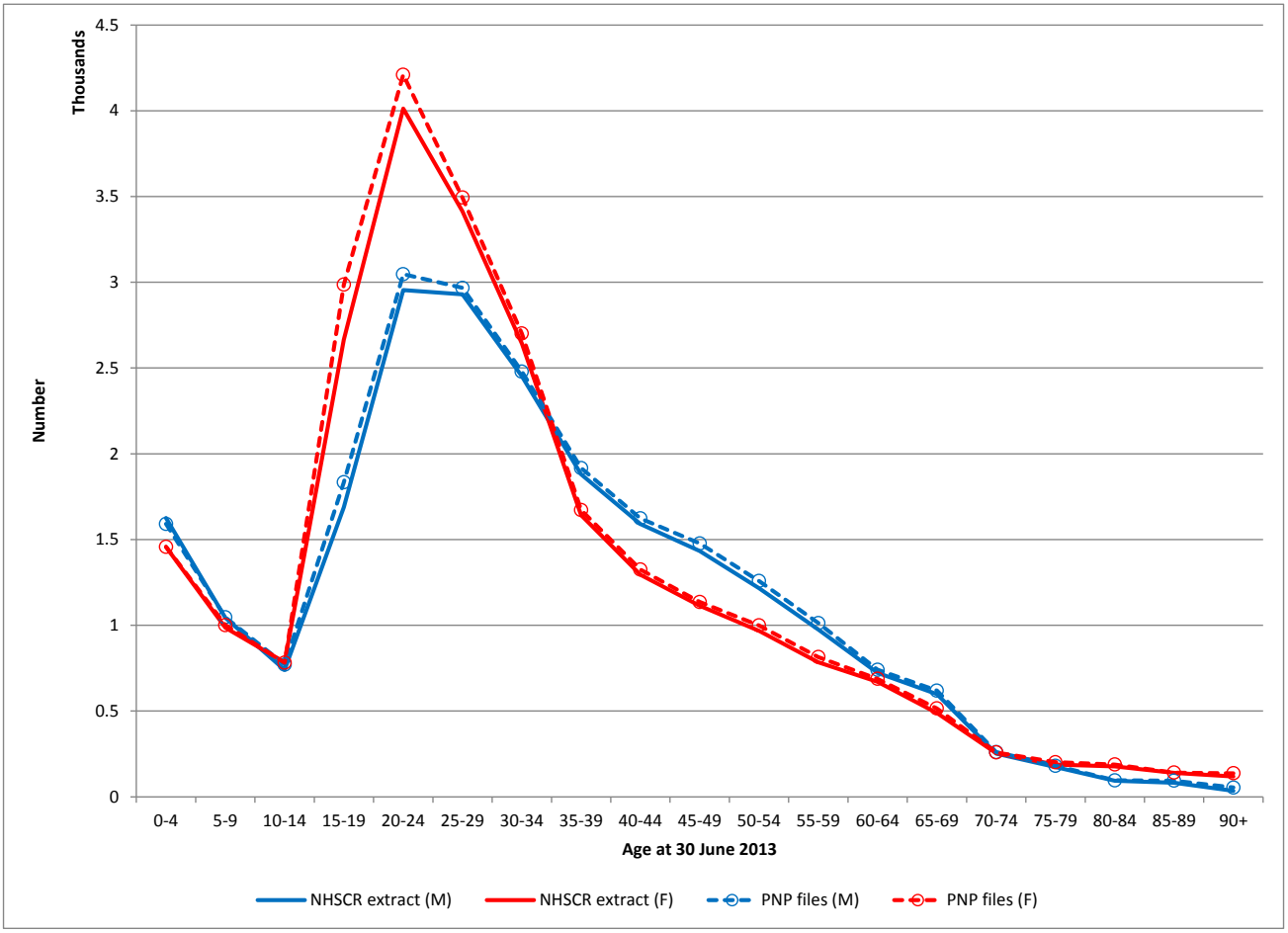
Moves to Scotland

Table 3: rest of UK to Scotland moves in the year to mid-2013 by quarter

Period	NHSCR Extract	PNP files	NHSCR Extract minus PNP files	
			n	%
2012 Q3	16,352	16,741	- 389	- 2.3
2012 Q4	9,542	9,597	- 55	- 0.6
2013 Q1	9,910	10,531	- 621	- 5.9
2013 Q2	10,483	10,909	- 426	- 3.9
All	**46,287**	**47,778**	**- 1,491**	**- 3.1**

31. Table 3 compares rest of UK (England, Wales and Northern Ireland) to Scotland moves between pre-April 2006 (15) NHS Board areas in the year to mid-2013 by quarter estimated using the NHSCR extract and the PNP files. Table 3 shows that the NHSCR extract estimates 3.1 per cent (approximately 1,490) fewer rest of UK to Scotland moves than the PNP files.

Figure 2: rest of UK to Scotland moves in the year to mid-2013 by 5-year age group



32. Figure 2 compares rest of UK to Scotland moves in the year to mid-2013 by 5-year age group for males (M) and females (F) estimated using the NHSCR extract and the PNP files. Figure 2 shows that there is good agreement between the estimates for males and females in most age groups.

33. The largest absolute differences for males and females are in the age group 15-19. The NHSCR extract estimates approximately 150 (8.0 per cent) and 320 (10.7 per cent) fewer male and female rest of UK to Scotland migrants aged 15-19, respectively.

Table 4: rest of UK to Scotland moves in the year to mid-2013 by NHS Board area

NHS Board area	NHSCR extract	PNP files	Diff[1]	Diff (%)
Fife	3,012	3,220	- 208	-6.5
Lothian	10,591	11,114	- 523	-4.7
Shetland	225	236	- 11	-4.7
Ayrshire & Arran	2,289	2,360	- 71	-3.0
Argyll & Clyde	2,724	2,806	- 82	-2.9
Western Isles	325	334	- 9	-2.7
Forth Valley	1,983	2,037	- 54	-2.7
Greater Glasgow	6,962	7,145	- 183	-2.6
Tayside	3,410	3,496	- 86	-2.5
Dumfries & Galloway	2,006	2,053	- 47	-2.3
Grampian	6,253	6,389	- 136	-2.1
Highland	2,455	2,491	- 36	-1.4
Lanarkshire	2,207	2,238	- 31	-1.4
Borders	1,560	1,576	- 16	-1.0
Orkney	285	283	2	0.7
Total	**46,287**	**47,778**	**- 1,491**	**-3.1**

1. NHSCR extract minus PNP files

34. Table 4 compares rest of UK to Scotland moves in the year to mid-2013 by NHS Board area estimated using the NHSCR extract and the PNP files. Table 4 shows that in general fewer rest of UK to Scotland migrants are estimated using the NHSCR extract than the PNP files. The largest percentage difference is in NHS Fife (-6.5 per cent) and the largest absolute difference is in NHS Lothian (-523).

35. Some of the difference between the estimates is due to historic moves coming through in the PNP files. NRS Population and Migration Statistics receive a third monthly (referred to in this paper as GROSMIG) file from NHSCR. This file has similar contents to the PNP files but contains health board posting start date which allows us to exclude historic moves. Excluding moves that occurred on or before the 2011 Census and comparing the estimate of rest of UK to Scotland migration from the GROSMIG files with the estimate from the NHSCR extract

- the difference between the estimates reduces from approximately 1,490 to 1,050 (3.1 per cent to 2.2 per cent),
- the differences in the age group 15-19 for males and females reduce from 150 and 320 to 20 and 30 (8.0 and 10.7 per cent to 1.3 and 1.0 per cent), respectively, and
- the percentage difference in NHS Fife reduces from 6.5 per cent to 1.6 per cent and the absolute difference in NHS Lothian reduces from -523 to -246.

36. This explains some but not all of the differences. We have investigated these differences extensively and have no reason to suspect that we are missing genuine rest of UK to Scotland moves. We have reproduced the PNP and GROSMIG files as closely as possible (following their specifications) and always find similar differences. Furthermore, working closely with colleagues from NHSCR, we have investigated

random samples of the health board postings we discard and have found no evidence to suggest that we are discarding genuine rest of UK to Scotland moves.

37. We will continue to investigate these differences but we ask PAMS members to approve the implementation of this method noting that we do not expect to find that the method is underestimating rest of UK to Scotland moves.

38. We will continue to use Office for National Statistics (ONS) Long-Term International Migration (LTIM) estimates to estimate (total) in-migration to Scotland from overseas and (total) out-migration from Scotland to overseas.

39. Currently single year of age, sex and geographical distributions for the LTIM estimates are derived from estimates in the PNP files, as described in Mid-Year Population Estimates for Scotland: Methodology Guide.

40. For the mid-2015 estimates onwards these distributions will be derived from estimates based on the methodology outlined in this paper.

41. Had this methodology been implemented in the year to mid-2013, the differences in the allocation of overseas in-migrants to pre-April 2006 (15) NHS Board areas would have been no more than +/- 0.1 percentage points.

Moves from Scotland

42. Currently our quarterly estimate of moves from Scotland to England & Wales from the PNP files are replaced with estimates (by single year of age, sex and origin and destination NHS Board area) provided by ONS (referred to in this paper as the EWMIG files) and our quarterly estimate of moves from Scotland to Northern Ireland from the PNP files is constrained to a total estimate provided by Northern Ireland Statistics and Research Agency (NISRA).

43. For the migration processing in the year to mid-2015, we will not be able to replace our estimate of moves from Scotland to England & Wales with the EWMIG files because the origin (Scottish) NHS Board areas will be a mix of pre-April 2006 and April 2014 NHS Board areas.

44. Therefore, for the migration processing in the year to mid-2015 onwards we propose to constrain our estimate of moves from Scotland to England & Wales derived from the NHSCR extract using the methodology described in this paper to the EWMIG files by 5-year age group (0-4, 5-9, …, 84-89 and 90+) and sex.

Figure 3: Scotland to England & Wales moves in the year to mid-2013 by single year of age

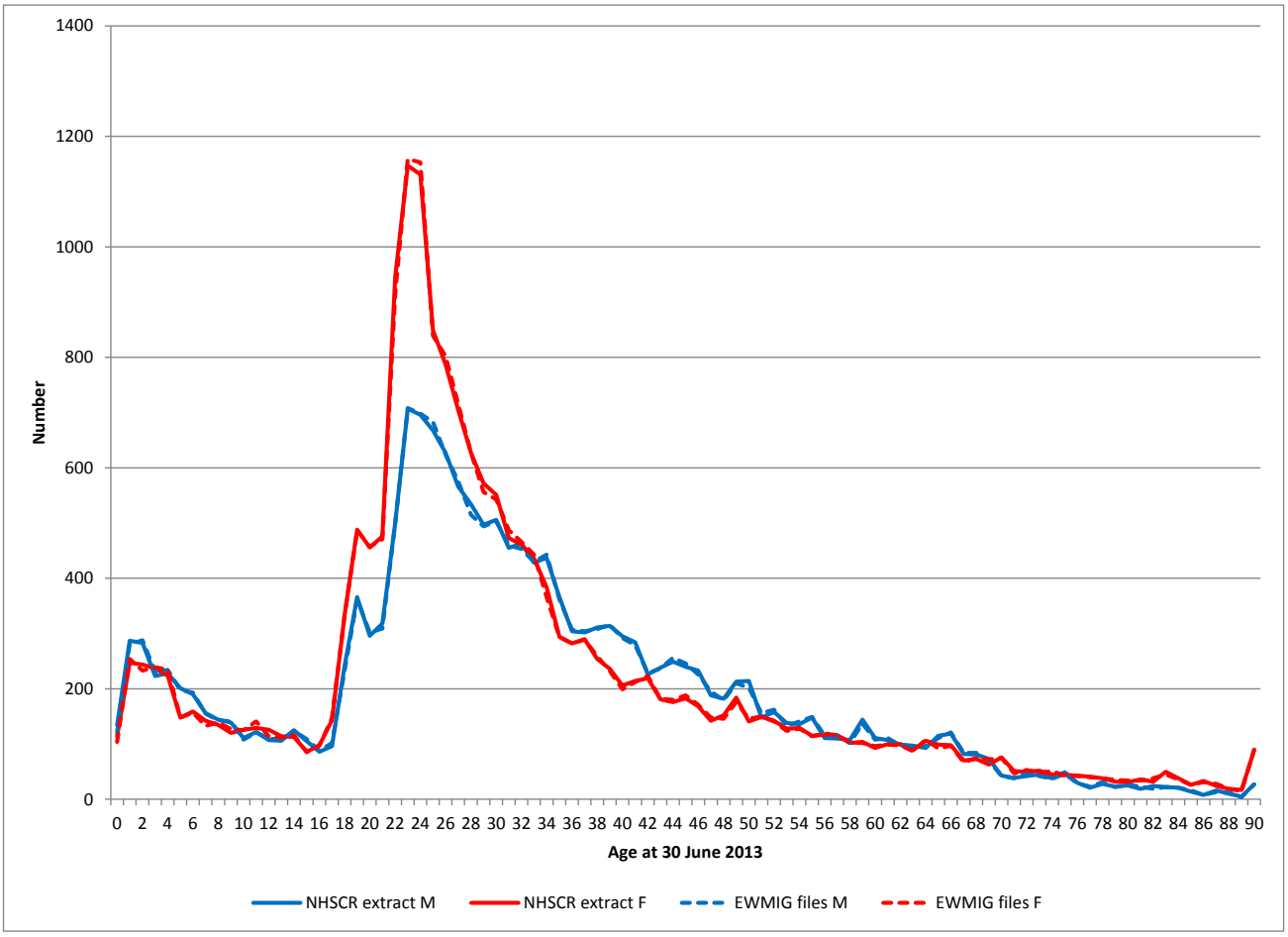


Table 5: Scotland to England & Wales moves in the year to mid-2013 by NHS Board area

NHS Board area	NHSCR extract	EWMIG files	Diff[1]	Diff (%)
Argyll & Clyde	2,194	2,239	- 45	-2.0
Dumfries & Galloway	1,392	1,417	- 25	-1.8
Fife	2,627	2,656	- 29	-1.1
Grampian	4,624	4,671	- 47	-1.0
Orkney	191	193	- 2	-1.0
Lanarkshire	1,840	1,856	- 16	-0.9
Tayside	2,789	2,809	- 20	-0.7
Forth Valley	1,544	1,551	- 7	-0.5
Shetland	201	202	- 1	-0.5
Ayrshire & Arran	1,759	1,753	6	0.3
Greater Glasgow	6,662	6,638	24	0.4
Lothian	8,906	8,825	81	0.9
Highland	1,698	1,669	29	1.7
Western Isles	212	206	6	2.9
Borders	1,102	1,056	46	4.4
Total	**37,741**	**37,741**	**-**	**0.0**

1. NHSCR extract minus EWMIG files.

45. Figure 3 compares Scotland to England & Wales moves in the year to mid-2013 by single year of age for males (M) and females (F) estimated using the NHSCR extract (constraining to the EWMIG files by 5-year age group and sex) and the EWMIG files. Figure 3 shows that there is good agreement between the estimates for males and females for all ages.

46. Table 5 compares Scotland to England & Wales moves in the year to mid-2013 by NHS Board area estimated using the NHSCR extract (constraining to the EWMIG files by 5-year age group and sex) and the EWMIG files. Table 5 shows that there is good agreement between the estimates in all NHS Board areas.

47. The views of PAMS members on this change to our methodology are welcome.

Conclusions

48. The results presented in this paper show that it is possible to use the NHSCR extract and the linked CHI postcodes to produce better quality quarterly migration data to replace the PNP and SMIG files.

49. The postcode information linked from CHI is of sufficient quality that it is possible to derive origin and destination NHS Board areas (when necessary) for all move types.

50. The views of PAMS members on this new methodology are welcome.

NRS: Population and Migration Statistics
5 May 2015

Annex A - Assessment of Imputation Methodology

Introduction

51. When using postcode information to produce migration data it is necessary to impute some of the postcode information. There are four reasons why a postcode may be invalid and have to be imputed:

- the postcode is a large user postcode that cannot be mapped to a small user postcode,
- the postcode is not recognised,
- there is a postcode record but the record is blank, or
- there is no postcode record (e.g. for moves to Scotland).

52. The purpose of this annex is to describe the methodology used to impute postcodes on the NHSCR and to assess its effectiveness.

Methodology

53. The migrants are split into three groups: within Scotland migrants, to Scotland migrants and from Scotland migrants.

54. Within Scotland migrants have origin and destination postcodes whereas to/from Scotland migrants have only a destination/origin postcode.

55. The within Scotland migrants are split into three groups:

- migrants with valid origin and destination postcodes (donors for within Scotland moves),
- migrants with a valid origin and an invalid destination postcode (type-1 recipients), and
- migrants with an invalid origin and a valid destination postcode (type-2 recipients).

56. A donor is selected for each type-1 recipient in stage 1 and a donor is selected for each type-2 recipient in stage 2. Type-1 recipients take their donor's destination postcode and type-2 recipients take their donor's origin postcode. The same donor can be selected more than once.

57. The donor for each recipient is selected randomly provided that they match with the recipient on certain variables. Donors are sought that match with the type-1/type-2 recipients on 5-year age group (0-4, 5-9, …, 85-89, 90+), gender,

- origin and destination postcode sector (part 1), else
- origin postcode district/sector and destination postcode sector/district (part 2), else
- origin postcode sector/district and destination postcode district/sector (part 3), else

- origin and destination postcode district (part 4), else
- origin postcode area/district and destination postcode district/area (part 5), else
- origin postcode district/area and destination postcode area/district (part 6), else
- origin and destination postcode area (part 7), else
- destination/origin postcode sector (part 8), else
- destination/origin postcode district (part 9), else
- destination/origin postcode area (part 10), else
- origin/destination postcode sector (part 11), else
- origin/destination postcode district (part 12), else
- origin/destination postcode area (part 13).

58. If no donors can be found following this then a donor that matches with the recipient on 5-year age group and gender only is selected (part 14), else a donor is selected randomly (part 15).

59. Postcode sectors, districts and areas are used to find donors so that, as much it is possible, some of the postcode information linked from the CHI is retained. For example, in cases where the postcode linked from the CHI is not recognised but can be decomposed into a valid postcode sector, district or area (perhaps because of an input error involving only the postcode unit, sector or district) then it is sensible to retain the valid postcode information if possible.

60. If a type-1 recipient has moved twice within Scotland in the same quarter then their invalid postcode will be imputed twice, once as destination in the stage 1 and once as an origin in stage 2. The donors selected in stage 1 and stage 2 will, in general, be different. To ensure that the same postcode is not imputed twice using different donors, the origin postcodes imputed in stage 2 that are also imputed in stage 1 as destination postcodes are overwritten with the imputation from stage 1. The exception to this rule are instances where the imputation in stage 1 occurred at a ‘later’ part (e.g. part 2 is ‘later’ than part 1) than the imputation in stage 2. In such cases the imputed destination postcode in stage 1 is overwritten by the imputed origin postcode from stage 2.

61. To and from Scotland migrants are split into four groups:

- to Scotland migrants with valid destination postcodes (donors for to Scotland moves),
- to Scotland migrants with invalid destination postcodes (type-3 recipients),
- from Scotland migrants with valid origin postcodes (donors for from Scotland moves), and
- from Scotland migrants with invalid origin postcodes (type-4 recipients).

62. A donor is selected for each type-3 recipient in stage 3 and a donor is selected for each type-4 recipient in stage 4. Type-3 recipients take their donor's destination postcode and type-4 recipients take their donor's origin postcode. Again the same donor can be selected more than once.

63. The donor for each recipient is selected randomly provided that they match with the recipient on certain variables. Donors are sought that match with the type-3/type-4 recipients on 5-year age group (0-4, 5-9, …, 85-89, 90+), gender,

- grouped origin/destination PNP code and destination/origin postcode sector (part 1), else
- grouped origin/destination PNP code and destination/origin postcode district (part 2), else
- grouped origin/destination PNP code and destination/origin postcode area (part 3), else
- grouped origin/destination PNP code (part 4).

64. If no donors can be found following this then a donor that matches with the type-3/type-4 recipient on grouped origin/destination PNP code only is selected (part 5), else a donor is selected randomly (part 6).

65. Where possible, grouped PNP codes are used to find donors as this ensures that the donors and recipients have made the same type of move, e.g. rest of the UK to Scotland, Scotland to Armed Forces et cetera.

Assessment

66. The effectiveness of the imputation methodology outlined above was assessed by removing migrants from the donor pool, adding them to the recipients, imputing their postcodes and then comparing their imputed postcode with their actual postcode. Donors were removed from the donor pool in such a way that the distribution of the recipients by 5-year age group and gender was representative of the real data.

67. The assessment was carried out for moves that occurred in Q3 2012 (i.e. moves that were added in the September, October or November 2012 NHSCR extracts).

Table 6: percentage of imputations with the correct NHS Board area by recipient type

Recipients	Total	Percentage (%) of imputations with the correct NHS Board area		
		Origin & Destination	Origin only	Destination only
Type-1 & 2	394	96.2	1.3	2.3
Type-1	2203	N/A	N/A	98.4
Type-2	1238	N/A	98.4	N/A
Type-3	452	N/A	N/A	96.5
Type-4	333	N/A	95.5	N/A

Table 7: percentage of imputations with the correct Council area by recipient type

Recipients	Total	Percentage (%) of imputations with the correct Council area		
		Origin & Destination	Origin only	Destination only
Type-1 & 2	394	88.3	4.8	5.3
Type-1	2203	N/A	N/A	94.5
Type-2	1238	N/A	93.9	N/A
Type-3	452	N/A	N/A	91.4
Type-4	333	N/A	83.8	N/A

Table 8: percentage of imputations with the correct data zone by recipient type

Recipients	Total	Percentage (%) of imputations with the correct data zone		
		Origin & Destination	Origin only	Destination only
Type-1 & 2	394	1.0	7.6	8.4
Type-1	2203	N/A	N/A	12.7
Type-2	1238	N/A	12.0	N/A
Type-3	452	N/A	N/A	23.0
Type-4	333	N/A	12.3	N/A

68. Tables 6, 7 and 8 show the percentage of imputations with the correct (pre-April 2006) NHS Board area, Council area and data zone, respectively, by recipient type.

Table 9: distance (km) between the imputed and actual postcode by recipient type

Recipients	Distance (km) between the imputed and actual postcodes					
	Origin			Destination		
	Mean	Maximum	Standard deviation	Mean	Maximum	Standard deviation
Type-1 & 2	4.1	107.5	10.1	3.9	95.0	10.4
Type-1	N/A	N/A	N/A	3.9	159.0	10.9
Type-2	4.1	146.8	12.3	N/A	N/A	N/A
Type-3	N/A	N/A	N/A	7.5	466.7	27.6
Type-4	11.5	415.7	31.2	N/A	N/A	N/A

69. Table 9 shows some summary statistics for the distance (in kilometres) between the imputed and actual postcodes by recipient type.

70. Tables 6, 7 and 8 show that for type-1 and type-2 recipients (i.e. migrants with invalid origin and destination postcodes) the imputed origin and destination postcodes correspond to the actual origin and destination NHS Board areas, Council areas and data zones in 96.2, 88.3 and 1.0 per cent of cases, respectively. A ‘success rate’ (the percentage of correctly imputed higher areas) in excess of 95 per cent for NHS Board areas is very encouraging. As expected the success rate decreases for lower geographies but is still in excess of 85 per cent for Council areas.

71. Tables 6, 7 and 8 show that for type-1 (type-2) recipients the imputed destination (origin) postcodes correspond to the actual destination (origin) NHS Board areas, Council areas and data zones in 98.4, 94.5 and 12.7 (98.4, 93.9 and 12.0) per cent of cases, respectively. Furthermore, Table 9 shows that the mean distance between the imputed and actual postcode is 3.9 kilometres for type-1 recipients and 4.1 kilometres for type-2 recipients.

72. Similar observations to those for type-1 and type-2 recipients can be made for type-3 and type-4 recipients. The success rates for NHS Board and Council areas are lower for type-3 and type-4 recipients than they are for type-1 and type-2 recipients, respectively. However, the success rate for data zones is higher for type-3 (type-4) recipients than for type-1 (type-2) recipients. This may be a consequence of the fact that migrants arriving in (or leaving) Scotland (e.g. students or Armed Forces personnel) tend to migrate to similar areas (e.g. areas with student halls of residence or Armed Forces bases), increasing the chances of selecting a donor moving to the same area as the recipient.