POPULATION AND MIGRATION STATISTICS COMMITTEE (SCOTLAND)

SUPPORTING INFORMATION -Validation of Proposed Methodology

This paper is a supplement to PAMS (03) 10, and describes both the proposals for 2002, and the validation of the methodology using CHI extracts for 2000 and 2001. It is not our intention for this paper to be discussed at the PAMS meeting; it is presented as background information only.

STEP 1 - validate postcodes and match CHI extracts

Each CHI extract contains around 9 million records, and so filters must be applied to get rid of around 4 million unwanted records. The filters that are applied to the CHI extracts are slightly different from those applied in SAPE production, because we are not necessarily interested in just those people who are 'currently registered'. We can gain useful information about the migration of people who move then die before the date of the second extract, and on people who move out of Scotland. In order to capture these people, the filters applied to the two extracts are as follows:

CHI Extract 1	CHI Extract 2	Comments
Exclude all records with a deletion indicator	Exclude all records with a deletion indicator	The deletion indicator refers to records that have been deleted because they have been entered onto the database in error. The records are not completely deleted but archived and are inaccessible to most users.
Exclude all records without a GP Practice code	Exclude all records without a GP Practice code	Records without a GP practice code are considered not to be registered with a GP. For example, some hospitals may enter patients on the CHI for local administrative purposes but they are not registered with a GP and are therefore excluded from the 'currently registered' population.
Exclude all records where the date of transfer out is before 30 September 01, if date of transfer in is blank or earlier than date of transfer out	Exclude all records where the date of transfer out is before 30 September 01, if date of transfer in is blank or earlier than date of transfer out	The date of transfer out refers to the patients *most recent transfer*, so if a person's record has a date of transfer out before a date of transfer in, then he is currently registered with a GP – eg, a person leaves home to work in England, and then returns home a few years later. His record will have both a transfer in and out, but by looking at the dates it is possible to determine the correct location of the patient at the time of the extract.
Exclude all records where the date of death is not blank		Excludes patients who have died before the date of the first extract, but not those who might have moved then died before the date of the second extract.
Exclude all records where the CHI status is not blank or a 'C'	Exclude all records where the CHI status is not blank or a 'C'	Excludes all of a patient's previous records. For instance when a patient registers with a new GP in another area, a new record is created and this indicator will identify the current record as 'live' and the previous record as 'historical'.
	Exclude records where age is –1	Because age is calculated as at three months before the date the extract was taken, babies born between the reference date and the extract date of extract 2 will have an age of -1, and should be removed.

Postcodes on the two filtered CHI extracts are validated using a lookup file derived from the GROS postcode index version 02/2. Both extracts are mapped to the same lookup table, so that when postcodes are compared to identify migrants, we are comparing like with like, and are not generating false migrants in cases where a postcode has been introduced, changed or deleted. Area codes are then added to the filtered CHI extracts, and the two extracts are matched by encrypted CHI number. This gives each record an origin postcode field (location in mid 2001), and a destination postcode field (location in mid 2002), and enables the identification of migrants, as described in step 2.

STEP 2 - Identify potential migrants in CHI data

In the 2000 MYE, the CHI was used in an ad hoc way to estimate within-Scotland, council level migration. However, the CHI was not used to estimate the migration of zero year olds, of people who died before the second extract, or of people entering or leaving Scotland. This section describes how the CHI can be used to capture these migrants. Methods were tested using CHI extracts for 2000 and 2001.

Zero year old migration

A dataset was provided by Vital Events Branch (VEB) of births between 1 July 2000 and 30 June 2001. From the matched CHI data, we extracted the set of records that were only present on the second extract and where the age was zero. This dataset was matched to the VEB births data by NHS number, and where a match was found, the postcode of birth was copied into the CHI 'PC0' field, ie. postcode at 1 July 2000. Of the 52,686 births registrations and 45,622 CHI records, 42,993 could be matched. Of the matched records almost 7,000 were migrants.

The CHI figure is significantly lower than the VEB births data because we assume that it takes three months to register with a GP after moving, and so we use a CHI extract from end-September to derive the currently registered population as at end June. This means that if a person is born between July and September 2000, and also registers with a GP by end September 2000, he will appear on both CHI extracts, and so won't be included in this sub-set of unmatched 0 year old records.

PAMS is asked to agree that from 2002 onward 0 year old migrants be identified by matching the CHI data to VEB births data

Migrant deaths

When a person dies, his record is not immediately removed from the CHI, so by including people who died between 1 July 2000 and 30 June 2001, we can determine the location of the person at the date of death, and whether he had migrated between 1 July 2000 and the date of death.

PAMS is asked to agree that from 2002 onward migration of people who die between the dates of two CHI extracts be estimated by including people on the second of the two extracts whose date of death is after 30 June of the previous year.

Transfers on and off the CHI

Each patient's CHI record has a field for information on transfers onto the CHI/into the CHI consortium and off the CHI/out of the CHI consortium. When a patient, say, moves from Scotland to England and re-registers with a new GP, his CHI record is updated with the date of transfer, transfer type and area code of the destination English health authority. Similarly, if a patient moves from England to Scotland, if it is his first registration with a Scottish GP, a new record will be created, and the 'transfer in' date and origin code recorded. If he has an existing CHI record, his record will be updated.

The CHI data received by GROS holds information on the patient's most recent transfer in and transfer out. The 'transfer in' field is an 11 character text string of the form yyyymmdd followed by a 3 character old area code. The 'transfer out' field is a 12 character text string of the form yyyymmdd followed by a 1 character transfer type, and a 3 character new area code. The type of 'transfer in' was initially not thought to be of use for estimating migration, so it was not requested by ISD[1]. However, it may be of use in determining which transfers onto the CHI are migration moves, in cases where no origin area code exists, so we wish to request this code for future extracts.

Four different flows can be identified from the CHI transfers information – flows to and from England & Wales, Northern Ireland, the Armed Forces and overseas. The working group paper SAPE 02(07) presented an initial analysis of the use of CHI transfer indicators to estimate these flows, comparing CHI transfers (mid 2000 – April 29 2001) with flows derived from NHSCR data. The results were encouraging, but there were some large differences in the flows captured by the CHI and the NHSCR, particularly for Armed Forces and Northern Ireland moves. The working group was keen to see this work repeated using mid 2000 - mid 2001 data to ensure that any discrepancies were not due to the fact that only 10 months of data were being analysed, and to allow further analysis to investigate the reasons for the discrepancies. Summarised results of this work are shown in Appendix B, and it was found that generally, the CHI and NHSCR flows are more similar for mid00-mid01 than for mid00-April01, although the two sets of results are broadly comparable:

Flows between Scotland and the Armed Forces

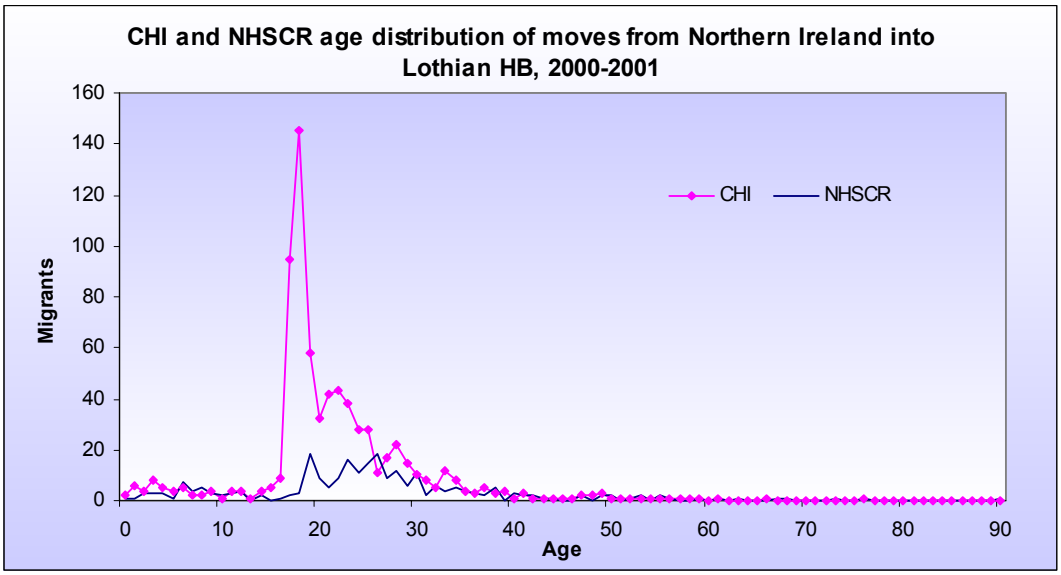
Flows to and from the Armed Forces appear to be recorded differently by the CHI and NHSCR. The CHI can and does distinguish between Armed Forces personnel and Armed Forces dependents, although the coding can be inconsistent. For example, where the reason for transfer indicates an enlistment with the Armed Forces, but the 'new area' indicates a civilian move to Northern Ireland. Or, the reason for transfer indicates a civilian move from elsewhere in the UK, but the new area is "AF" – a move to the Armed Forces, or "SMO" – an Armed Forces dependent. In these cases it is difficult to decide whether the move is an Armed Forces move or not. On the NHSCR, the figures for discharges from the Armed Forces appear to include dependents, though the figures for moves to the Armed Forces don't. Because of these difficulties, we do not propose to use the CHI Armed Forces data alone to distribute NHSCR Armed Forces flows down to Council level.

Flows between Scotland and Northern Ireland

CHI and NHSCR flows from Scotland to Northern Ireland were very similar in both sets of results, though slightly more so for the mid 2001 data. However, in both cases, there were large discrepancies between the CHI and NHSCR for moves *to* Scotland *from* Northern Ireland, particularly in Tayside, Lothian, and Forth Valley. The discrepancies in Tayside and Lothian can perhaps be partially explained by examining the ages and destination postcodes of these migrants. Large numbers of 18 and 19 year olds are recorded on the CHI as moves into student areas and halls of residence during October 2000 and late September 2001, and it is possible that registering with a university GP is part of the matriculation process. These people are therefore being recorded on the CHI a lot sooner than 3 months after the date of migration. However, these people appear not to be recorded by the NHSCR (See figure 1 for moves from Northern Ireland to Lothian).

[1] All transfers onto the CHI that could involve a migration move are given the same code "T = Transfer on from another consortium or from anywhere not using CHI."

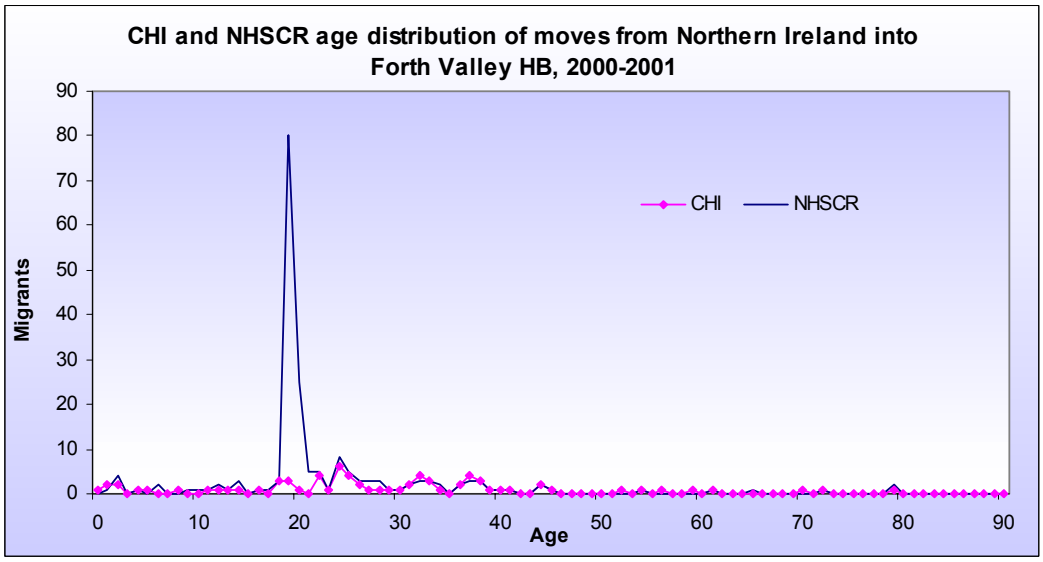
Figure 1



Normally there is only a minimal delay between the CHI being notified of a new registration and the NHSCR being updated. However, a lengthier delay, estimated to be up to a month, does occur during the autumn with the registration of students. The delay originating from student movements means that the movement is recorded on the NHSCR about a month after the movement is recorded on the CHI. However, any significant imbalance caused by the differences in timing should balance out from one year to the next. If this is a timing issue, then it could indicate that the delay between the CHI and NHSCR being updated is not consistent from year to year, or that the 2001 CHI extract was taken at a later date than the 2000 extract. It is important to ensure that the extracts are taken at the same time each year to ensure that this does not happen.

There are similar discrepancies between CHI and NHSCR moves from Northern Ireland into Forth Valley, but in this case the CHI rather than the NHSCR misses moves of 19 and 20 year olds, as illustrated in figure 2.

Figure 2

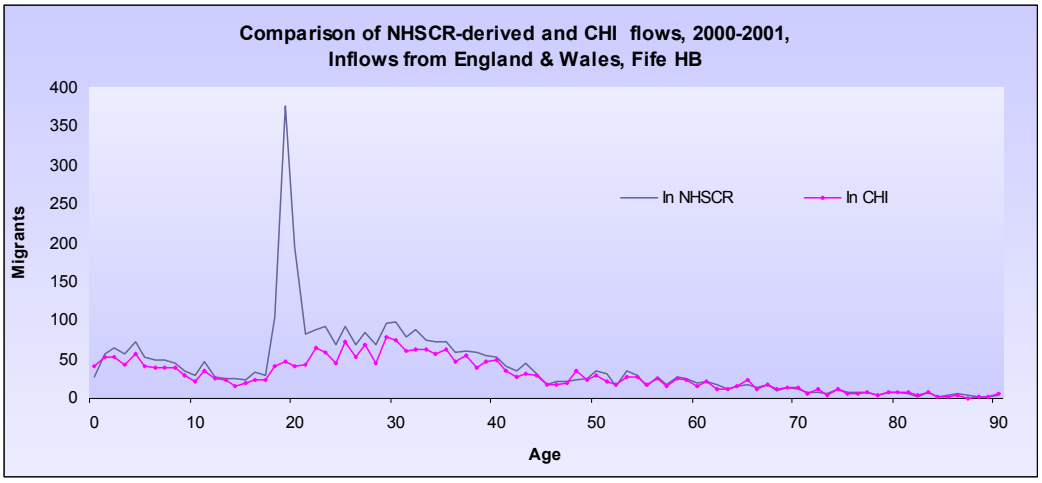


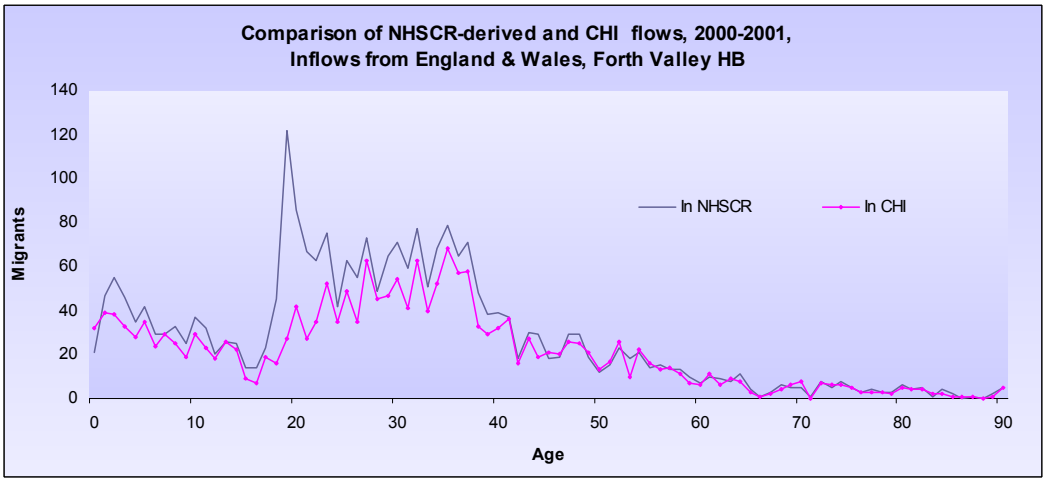
Again, it is possible that this is caused by differences in the timing of the recording of moves by the CHI and NHSCR, or differences in the timing of the two CHI extracts. If this was the case, we might expect to see a large peak of CHI 18 year olds in the 2002 migration data. It will be interesting to see if this is the case. Alternatively, the differences between the CHI and NHSCR in Forth Valley could be caused by the coding practices of the CHI consortium. CHI migration flows into Fife show an under-recording of 18-19 year olds similar to Forth Valley's, though to a lesser extent. The fact that these two health boards are part of the same CHI consortium, might lend weight to this theory. Whatever the cause, further work will be required to untangle the reasons for the discrepancies before we can use this data alone to produce migration estimates for moves between Scotland's councils and Northern Ireland.

Flows between Scotland and England & Wales

Agreement between the CHI and NHSCR is particularly good for moves between Scotland and England & Wales for most health boards. Although the CHI slightly under-records moves in both directions, the age/sex distributions are very similar, and would be suitable for creating council level estimates of flows to and from England & Wales. However, as with the inflows from Northern Ireland, the exceptions are Forth Valley and Fife. The CHI inflows from England & Wales are missing the peak of 18-19 year olds that are present in the NHSCR data – see figure 3. If we can find a reason for the discrepancy, we can either choose to ignore it, or can make an appropriate adjustment, and then use this flow to produce estimates of migration between Scotland's councils and England & Wales. Alternatively, we could group the Northern Ireland and England & Wales CHI flows to create council level estimates of moves between Scotland and the rest of the UK.

Figure 3





Flows between Scotland and Overseas

Both the CHI and the NHSCR overestimate immigration and underestimate emigration. The overestimate of immigration is due to definitional differences between the CHI/NHSCR, and the International Passenger Survey (IPS), which is used to create Scotland-level migration. An immigrant is defined by the IPS as someone who enters the country with the intention of staying for at least 12 months, whereas anyone who registers with a GP after entering the country from abroad is captured by the CHI and NHSCR as an immigrant. The underestimate of emigration is caused by there being no requirement to notify one's GP before moving abroad.

For both inflows and outflows the CHI records slightly fewer moves than the NHSCR, though the age/sex patterns are generally very similar. Again, moves into Fife and Forth Valley are missing 18-19 year olds. However, the most noteworthy discrepancy between the CHI and NHSCR inflows occurs in Grampian, as follows:

When a person transfers onto the CHI, his record is updated with the place he transferred in from. If he is an international immigrant, the code is 'IMM', a move from England & Wales is given a Health Authority 'Q code', a move from Northern Ireland is coded 'NI', and a return from the Armed Forces is 'AF' or 'SMO'. In the majority of cases this code is present - for transfers in between mid 2000 and mid 2001, it was present in 93% of records. However, it is not a mandatory field, and differing coding practices mean that the missing codes are not evenly distributed between Health Boards. For mid00 – mid01 almost 60% of the moves with missing codes were moves into Grampian, and a further 20% were moves into Tayside (see figure 4).

Figure 4

Destination HB	'Transfers In' with missing origin	As % of all transfers in
Highland	315	9.3%
Grampian	2932	30.0%
Tayside	994	20.0%
Fife	59	1.5%
Lothian	105	0.6%
Borders	6	0.3%
Forth Valley	33	1.2%
Argyll & Clyde	149	3.5%
Greater Glasgow	111	0.8%
Lanarkshire	33	1.1%
Ayrshire & Arran	88	2.8%
Dumfries & Galloway	40	1.6%
Orkney	32	12.1%
Shetland	14	4.7%
Western Isles	2	0.9%
Scotland	**4913**	**6.8%**

Through comparison with NHSCR-derived flows for mid00-mid01, it seems likely that the flows into Grampian without an origin code are international immigrants, and those into Tayside and perhaps Highland are cross border migrants from England & Wales. Figures 5 and 6 are comparisons of 3 Health Board inflows from England and Wales and overseas respectively. The three flows shown are:

1. The NHSCR-derived inflow (GROS)
2. The CHI inflow (CHI)
3. The CHI inflow plus CHI records where the area of origin is unknown (CHI with null area records)

Figure 5

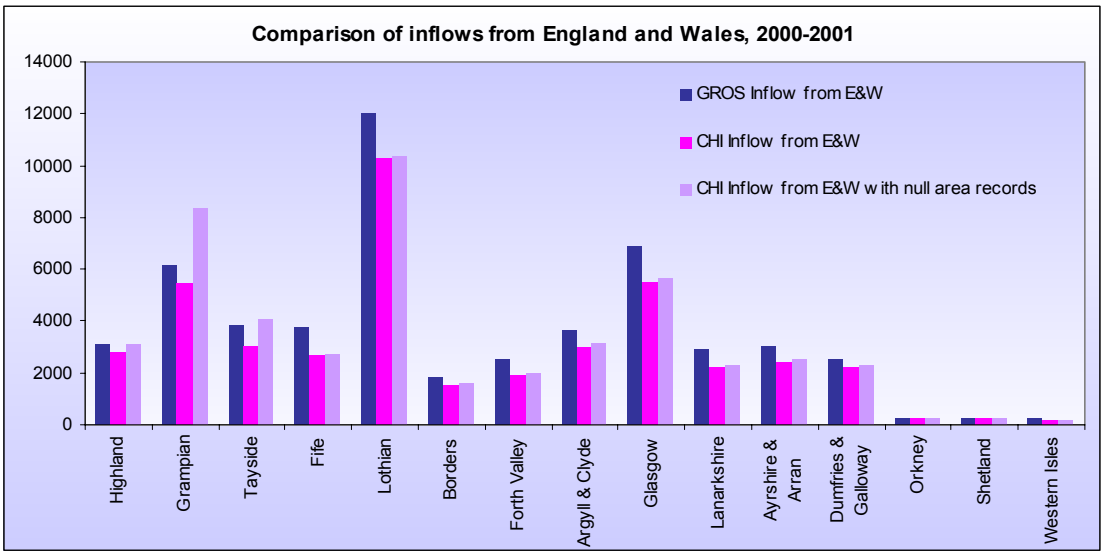


Figure 6

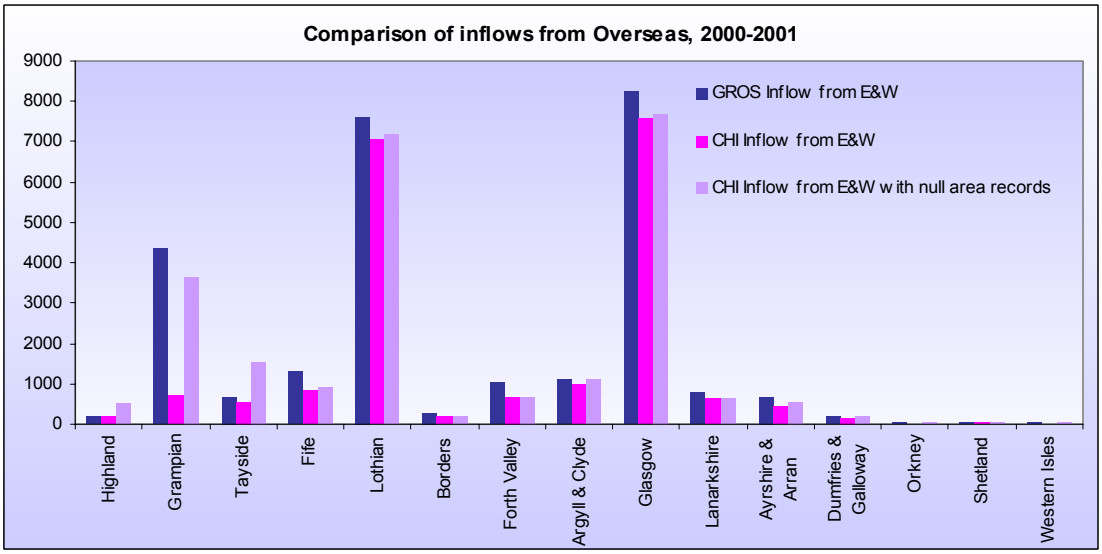
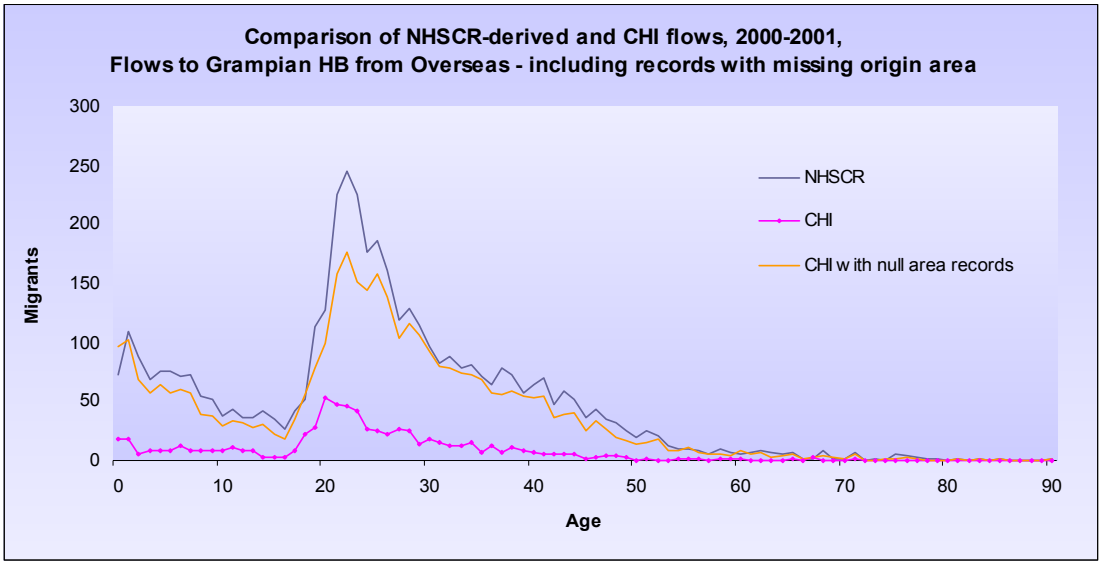


Figure 7 further illustrates the improvement in the overseas CHI inflow to Grampian HB when these 'null area' records are included. It is therefore proposed that in the 2002 migration data, these records with a missing origin code be included in the overseas inflow to Grampian, and that further work be undertaken to allow us make a decision about the inclusion of such moves for the other Health Boards.

Figure 7



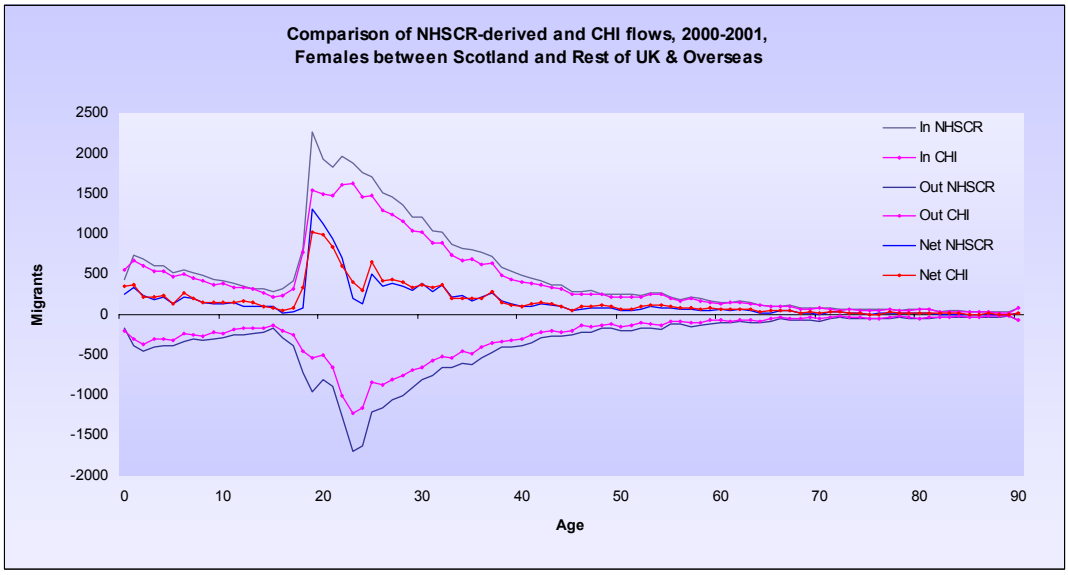
Using CHI transfers to and from Scotland in the 2001 Migration data

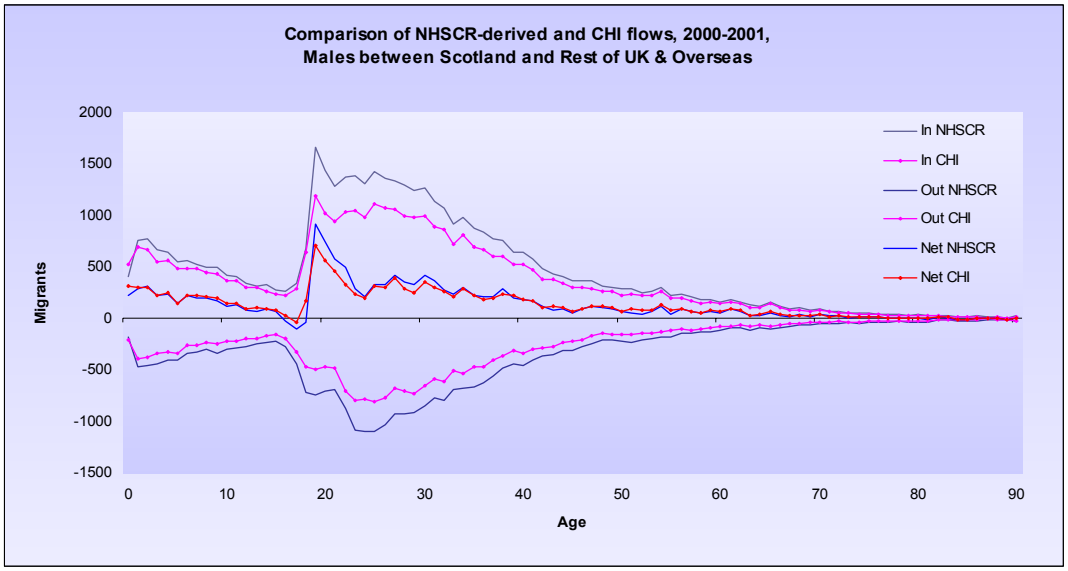
Because of a) the differences in the CHI and NHSCR flows for moves to and from the Armed Forces, b) the discrepancies between flows in Fife and Forth Valley, and c) the HB distribution of 'transfers in' with missing area codes, for 2001, no attempt was made to separate the CHI transfers in and out of Scotland into their constituent flows. Figure 8 is a comparison of the CHI and NHSCR combined flows between Scotland and the 'rest of the world'. Although both the inflows and outflows on the CHI are lower than the NHSCR, the net flows are similar, as are the age/sex distributions, and in the 2001 migration data, these combined CHI flows were used to apportion the equivalent NHSCR-derived Health Board flows to councils.

In the 2001 CHI migration data, a 'transfer in' was defined as all transfers in between 1 October 2000 and 30 September 2001, except those transfers from another CHI consortium in Scotland, which are denoted by a single character HB code. For mid 2000 – mid 2001, this gave an inflow of 74,693 cross border and international migrants, and returns to the NHS from the Armed Forces.

Transfers off the CHI are harder to count, because unless the patient is claiming state benefits or notifies his doctor, there is no way for international emigrants to be identified. This means there is a much larger proportion of transfers off the CHI with a missing destination area, than transfers onto the CHI with a missing origin area. Also, there are substantial numbers of records marked as untraced transfers out of a practice district, or HB area, which may or may not be a move outside Scotland. For these reasons, only those transfers where the area of destination is present were included in the cross border and international out-migration flows. For mid 2000 – mid 2001, this yielded 44,939 CHI migrants.

Figure 8





PAMS is asked to agree that information on transfers on and off the CHI be used to estimate migration into and out of Scotland as follows:

- ***From 2002, all transfers onto the CHI, except those from another CHI consortium in Scotland will be counted as moves into Scotland from outside Scotland***
- ***Further work should be undertaken to investigate the reasons for the observed discrepancies between the CHI and NHSCR data, particularly in Fife and Forth Valley Health Boards, with a view to producing separate migration flows for moves to/from the rest of the UK and moves to/from overseas.***
- ***From 2002, only those transfers off the CHI with a destination area code that represents a move outside Scotland will be counted as migration moves.***

STEP 3 - Impute missing postcodes and area codes

An imputation methodology was proposed in the working group paper SAPE 02(04), and was approved subject to validation. A method has since been developed and validated and was used in the production of 2001 migration data. We propose to carry on using a version of this method in the production of future migration estimates.

The methodology uses a donor imputation process to impute missing postcodes in the CHI migration data. Donor imputation assumes that individuals with similar characteristics will live (and migrate to) similar geographic locations. Therefore, for each record with a missing postcode, a match is sought from a set of complete records with, say, the same age, sex, origin and destination as the recipient record. It follows that the larger the number of variables that are available to match against, the smaller the number of potential donors will be for each record. However, as the number of potential donors decreases, so the accuracy of the imputation increases. This method starts by trying to match 'dirty' records against as many variables as possible from the donor dataset. Then, if a match is not found, the matching criteria are gradually relaxed, and the matching process repeated as many times as is necessary to find a suitable donor. This ensures that the imputation is as accurate as it can be.

To test the methodology, a cut down version of the matched CHI data was used. A random sample of size 500,000 was taken from the set of complete mid00 – mid01 matched CHI records. From this, postcodes were randomly blanked out to give the same distribution of missing postcodes in the sample as in the real data. New postcodes were imputed, and the original postcode compared with the imputed postcode. This enabled us to determine the matching hierarchy that gave the greatest accuracy, and a 5-stage process was developed, as illustrated in Appendix C.

Results of Imputation Testing

Our sample contained 500,000 CHI records. Of these, the postcode information was incomplete in 9129 records (1.83%). 483,694 of the 500,000 records had both an origin and a destination within Scotland. Of these, 1411 records had a missing origin postcode, 2151 a missing destination code, and 5335 had both postcodes missing. For the purposes of migration we are more interested in records with only one postcode missing, for logically, these people are more likely to be migrants - i.e. for a person's CHI record to have a valid postcode in one year, but not in the other, the record must have been updated, and one reason for a record to be updated is a change of address.

Table 1 shows the imputation accuracy achieved by each of the 5 stages of imputation, and compares this with results achieved using the existing imputation methodology, which matches against current HB, GP practice, age and sex only. More detailed results by health board and age are shown in Appendix D.

Table 1

Imputation stage	Number of records	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Missing origin postcode - within Scotland	1,411	100	100	96	51	98	69	52	38
Missing destination postcode - within Scotland	2,151	100	98	97	53	98	74	57	40
Both missing - within Scotland (origin PC)	5,335	100	96	96	26	99	64	35	2
Both missing - within Scotland (destination PC)	5,335	100	96	96	26	99	65	36	2
Missing destination postcode - Transfers In	158	99	94	93	30	99	63	43	13
Missing origin postcode - Transfers Out	74	100	92	92	39	99	51	32	7
Original method used in 1999 and 2000 small area population estimates	8,906	100	96	96	25	99	62	33	2

These results show that for those records where both origin and destination postcodes are missing, the new method is comparable with the old (indeed, these two methods are very similar). However, where only one of the two postcodes is missing, the new method is considerably more accurate than the old, with over half of the records being allocated a correct postcode sector and/or ward, and more than a third assigned a correct unit postcode.

When the imputation of postcodes on the CHI data was first investigated in mid 2000, it was suggested that even where an incorrect postcode had been imputed, it was likely that the distance between the imputed and original postcode would be small. This work wasn't undertaken at that time, but has now been completed using our sample data.

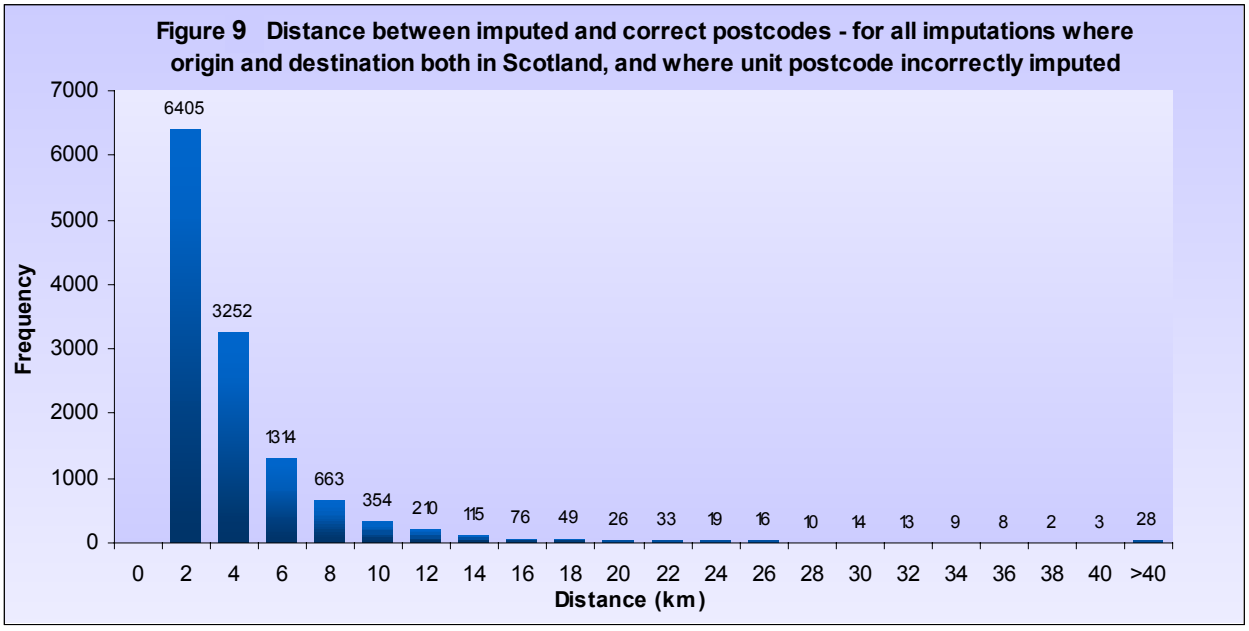
For each imputation type, the grid references of the centre point of the original correct postcode and the imputed postcode were derived. Using the following formula, the distance in metres between the two grid references was calculated:

$$\text{distance} = \sqrt{((\text{East}_i - \text{East}_o)^2 + (\text{North}_i - \text{North}_o)^2)}$$

where East_i = Grid reference Easting of imputed postcode
East_o = Grid reference Easting of original, correct postcode
North_i = Grid reference Northing of imputed postcode
North_o = Grid reference Northing of original, correct postcode

The results, summarised in Figure 9, show that even where an incorrect postcode has been imputed, the distance between the imputed postcode and the actual, correct postcode is less than 2km in over half of the imputations. Further details showing distances by imputation type can be found in Appendix D.

Figure 9 Distance between imputed and correct postcodes - for all imputations where origin and destination both in Scotland, and where unit postcode incorrectly imputed



This imputation method was used for the first time in the production of 2001 migration data. Table 2 illustrates the extent of missing or invalid postcodes for each type of imputation in the 2001 data, and gives an indication of the expected number of correctly imputed area codes, according to the results from the sample.

Table 2		Expected number of records with correct geographic indicator							
Imputation type	Number of records	Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Missing origin postcode - within Scotland	15373	15373	15373	14774	7790	15133	10645	8062	5818
Missing destination postcode - within Scotland	22846	22846	22347	22134	12066	22485	16813	13000	9134
Both missing - within Scotland (origin PC)	56511	56511	54361	54382	14935	56087	36067	19776	1165
Both missing - within Scotland (destination PC)	56511	56511	54371	54371	14745	56119	36576	20338	1229
Missing origin postcode - Transfers In	1742	1720	1632	1621	529	1731	1103	750	221
Missing destination postcode – Trans Out	840	840	772	772	329	829	431	272	57

In future years, the imputation process will be simpler. Stages 1, 3 and 5 will be unnecessary, because the postcode information on the September 2001 extract is now complete, having been through the imputation process for producing 2001 data. This imputed 2001 extract will be used as 'extract 1' for next year's migration data processing.

Given the increased accuracy of the imputation when both origin and destination information is available, we propose to use the postcodes imputed for migration purposes to feed back into the SAPE currently registered population, thus improving the quality of the SAPEs.

PAMS is asked to approve the imputation methodology outlined here, and to agree that this method should be used to feed back postcode information into the Small Area Population Estimates.

STEP 4 - Control CHI migration to HB level migration by age and sex

Step 4, Part 1 - Create control file

It is recognised that the NHSCR is the best available source of health board level migration data, and that migration data obtained from the CHI for smaller areas should be controlled to the NHSCR data. In order to do this, we need to create a dataset of summarised NHSCR-derived migration, by origin HB, destination HB, age and sex. This is required for all moves, not just those wholly within Scotland, and was created for 2001 as follows:

- Within-Scotland migration (80113) and migration to and from the rest of the UK, by age, sex and HB was obtained as usual, from the output of the migration phase of EMP[2]. Flows from Scotland to the rest of the UK and the rest of the UK to Scotland were controlled to figures agreed with ONS (51480 and 54885 respectively). Moves to and from the Armed Forces were obtained from the NHSCR.
- Scotland level international migration *totals* were calculated as 18700 in (excluding asylum seekers), and 19100 out, and were obtained from three sources, the International Passenger Survey (IPS), the Irish CSA, and the Home Office. These totals were apportioned to health boards using the health board proportions in the NHSCR international flows.
- An age/sex distribution was applied to the health board international migration totals. The distribution used was derived from the NHSCR data, but excluding the international migration flows themselves. This produces HB migration by origin, destination, age and sex. It should be noted that this is not a change in methodology, just a change in the way the existing methodology is applied.

[2] EMP is GROS' Estimates, Migration, and Projections data production system

- Finally, the within-Scotland, cross border and international data was combined, and where a move had an origin or destination outside Scotland, it was allocated a health board of '16'.

Step 4, Part 2 – Control CHI migration data

The following method was used to create a 2001 CHI migration dataset that is fully consistent with NHSCR-derived flows at health board level by age and sex.

For each combination of origin health board, destination health board, sex and syoa:

Stage 1 Where the number of records on the CHI is equal to the number of records on the NHSCR, copy CHI records from input data to controlled dataset.

Stage 2 Where the number of records on the CHI is greater than the number of records on the NHSCR, delete CHI records randomly one at a time from the appropriate sub-set of CHI records until the desired figure is achieved.

Stage 3 Where the number of records on the CHI is less than the number of records on the NHSCR, do one of the following:

1. If the ratio of NHSCR to CHI records in a particular origin/destination/age/sex group is three or less, then duplicate randomly selected CHI records from the appropriate subset until desired number is achieved. Note – we impose a limit on the number of times each CHI record is duplicated to avoid distorting the geographic/age/sex distributions of the data.

2. If there are no CHI records in a particular origin/destination/age/sex group available to duplicate, then use the method illustrated in the following example:

   Eg. To control the flow health board1 to health board2 for males aged 3 to a desired figure of 2, we randomly select two CHI records which have an origin of health board1, a destination of health board2, a sex of male and an age in the range 0-10. Set their age equal to 3, and then add these two records to our controlled dataset.

3. If there are insufficient CHI records in a particular origin/destination/age/sex group available to duplicate, ie. the NHSCR:CHI ratio is greater than three, then use a combination of the above methods as illustrated in the following example:

   Eg. To control the flow health board1 to health board2 for males aged 3 to a desired figure of 24, where there are only 7 records in the appropriate CHI origin/destination age/sex group, we first duplicate each of the CHI records three times (as in stage 3, part 1). We now need to duplicate a further 3 CHI records to achieve the desired total of 24. To do this, we randomly select 3 CHI records that have an origin of health board1, a destination of health board2, a sex of male and an age in the range 0-10. Set their age equal to 3, and then add these 3 records to our controlled dataset.

It is important that when controlling the migration flows at health board level, we do not distort the distribution of CHI flows to and from the smaller geographical areas within each health board. For example, if, in the raw CHI data, moves into Edinburgh City comprised 70% of all moves into Lothian health board, then this proportion should remain the same in the controlled CHI data**.** The method was tested using migration data from mid00 – mid01, and this was found to be the case. For details, see Appendix E.

PAMS is asked to agree the method proposed for controlling CHI between-HB moves to NHSCR between-HB moves.

Further work

Although the methods proposed for the 2002 migration data are an improvement on previous years, two aspects would benefit from further work:

Firstly, further work on the reasons for the discrepancies between the CHI and NHSCR flows for moves outside Scotland might enable us to eventually produce council level migration tables by age and sex for the following: within-Scotland moves; moves between Scottish councils and the rest of the UK; and moves between Scottish councils and overseas. Ideally, we would control CHI overseas migration by HB, age and sex to an NHSCR-derived overseas migration dataset. Similarly, CHI migration to and from the rest of the UK could be controlled to the NHSCR-derived 'rest of UK' flows by HB, age and sex. This would enable us to produce migration output as follows:

Council	Within Scotland			To/from rest of UK			To/from overseas			Total Migration		
	In	Out	Net	In	Out	Net	In	Out	Net	In	Out	Net
Aberdeen City Aberdeenshire ↓ West Lothian												

PAMS is asked to approve further work to improve the quality of the CHI transfers data with a view to splitting the migration of people to and from Scotland into moves to & from 'rest of UK' and 'overseas'.

Secondly, the method used to control CHI migration to NHSCR-derived migration only controls moves between health boards. Moves within a health board are not addressed. In the absence of any alternative migration dataset to control to, the information gained by controlling the CHI data to the NHSCR data for within-Scotland moves between health boards could be used. This would be based on the assumption that the scaling required to control the CHI data to the NHSCR data for moves between HBs will be the same as for moves within HBs. The proposed method is as follows:

1. Summarise the CHI and NHSCR datasets by origin and destination HB to give (for example):

Origin HB	Destination HB	CHI count	NHSCR count
01	02	284	203
01	03	133	93
01	04	57	45

2. Create grossing factors (NHSCR count / CHI count) for each origin/destination combination:

Origin HB	Destination HB	CHI count	NHSCR count	Gross factor
01	02	284	203	0.715
01	03	133	93	0.699
01	04	57	45	0.789

3. Create average grossing factors for each origin, and average grossing factors for each destination HB. These two values can be used to create an average grossing factor for moves within that HB (avgross):

HB	grossin	grossout	avgross
01	0.71	0.65	0.68
02	0.67	0.69	0.68
etc			

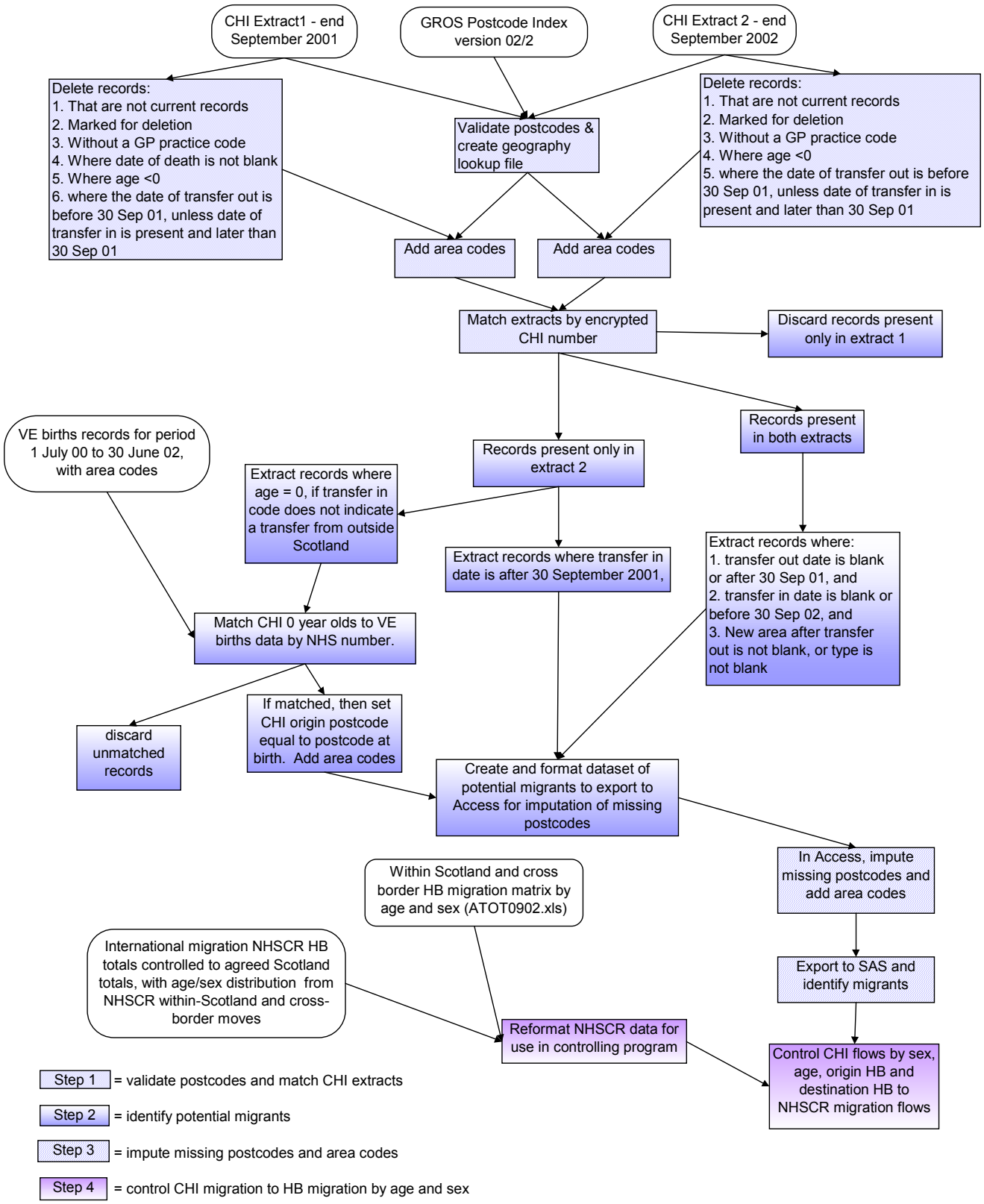
4. Apply the average grossing factor for a health board to moves within that health board to derive the desired number of CHI within-HB moves, then randomly delete or duplicate records to achieve that desired total:

HB	avgross	CHI orig	CHI adj
01	0.68	16070	10927
02	0.68	47264	32139
etc			

This method was approved by a Peer Review for use in ONS's internal migration data, but was not used to produce GROS 2001 migration data. The 2001 Census migration data will be useful for evaluating this method with a view to implementing it in the production of 2002 migration data, or 2003 if Census migration data is not made available by end March 2003.

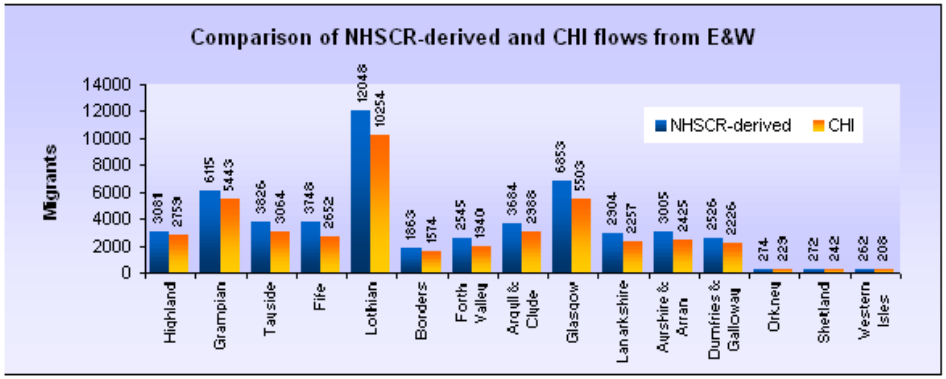
PAMS is asked to agree the method in principle, subject to validation against Census migration data.

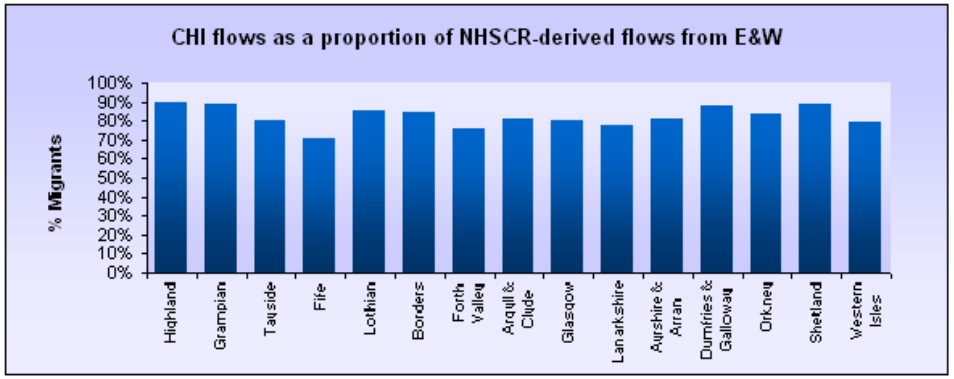
Appendix A – Summarised Methodology

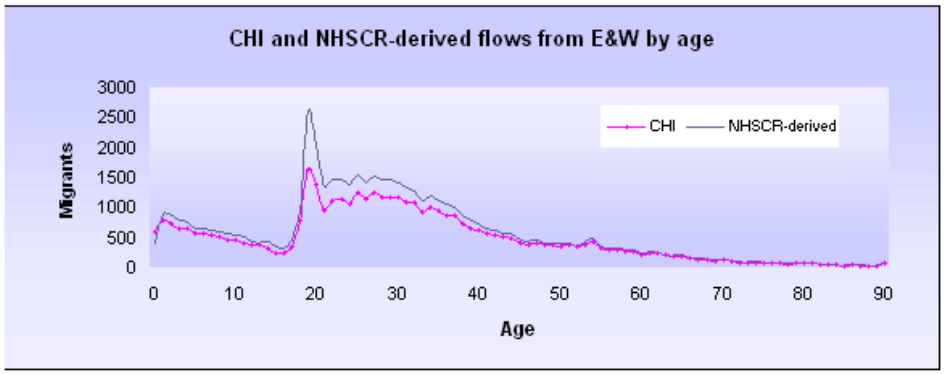


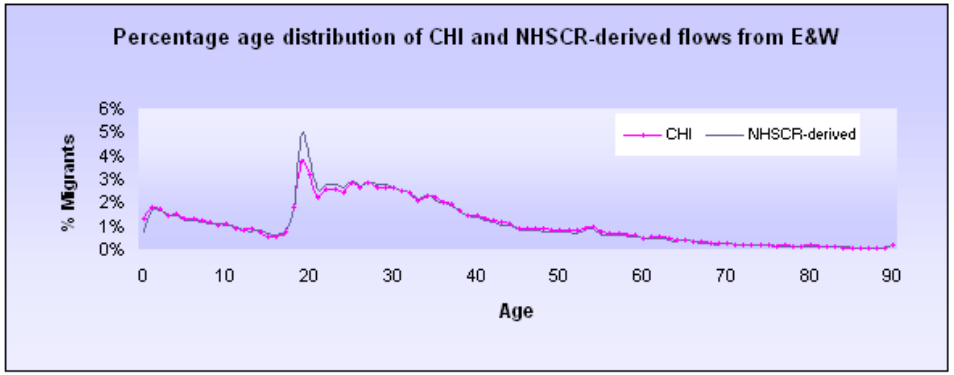
Appendix B

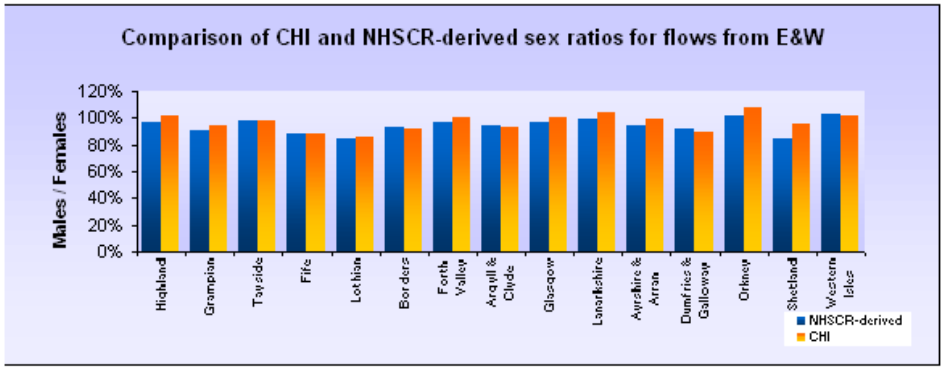
Comparison of raw CHI and NHSCR-derived INFLOWS into SCOTLAND from ENGLAND & WALES, 2000-2001



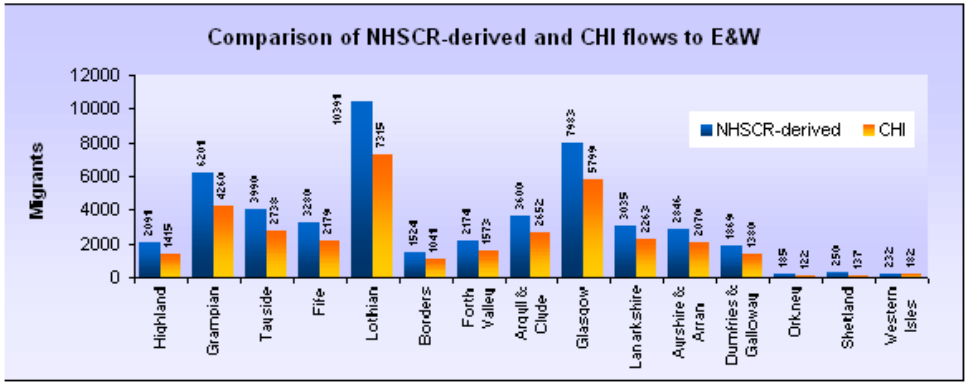


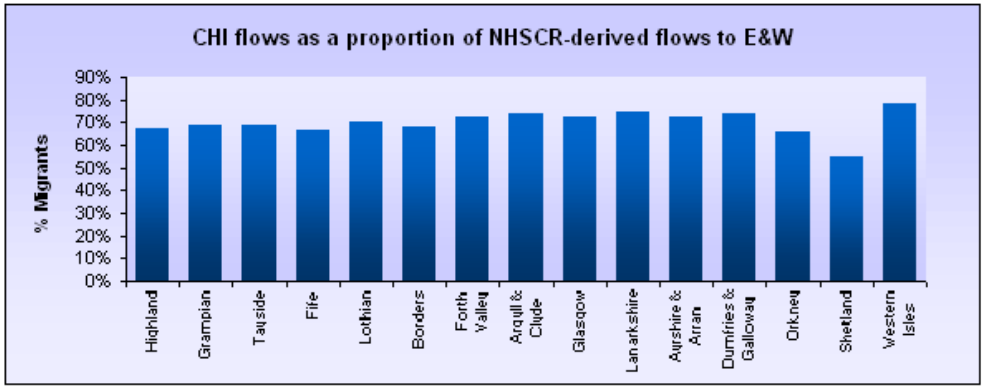


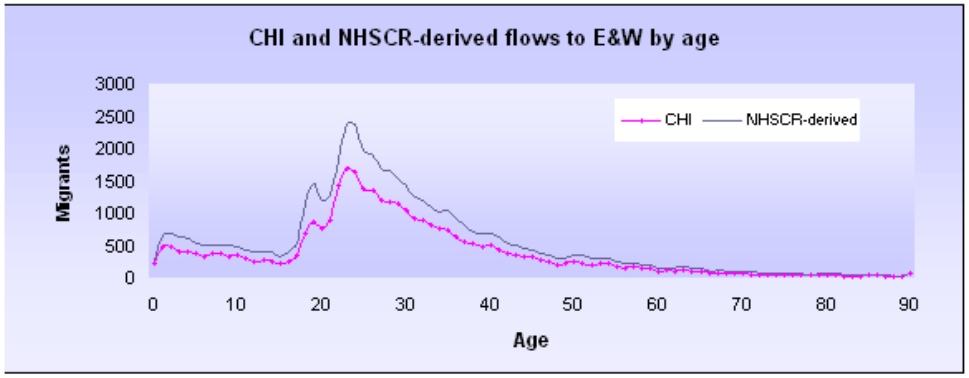


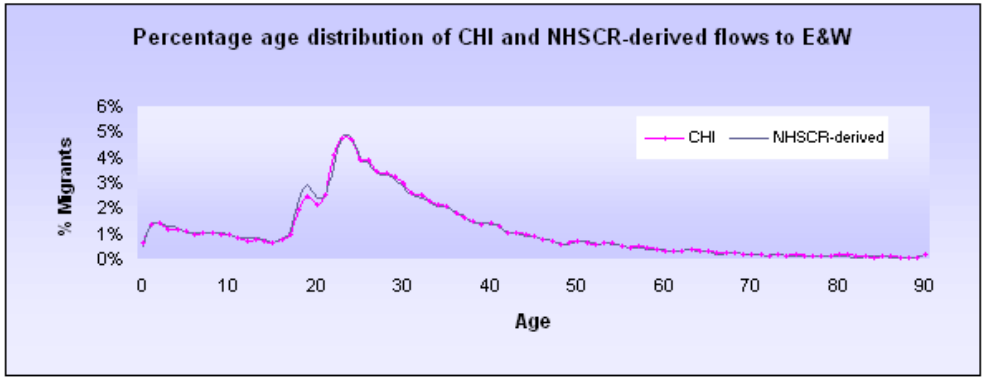


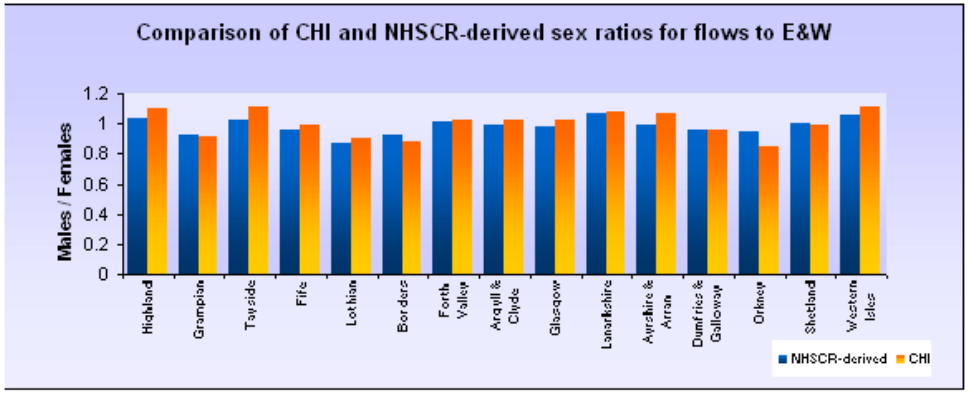
Comparison of raw CHI and NHSCR-derived OUTFLOWS from SCOTLAND into ENGLAND & WALES, 2000-2001



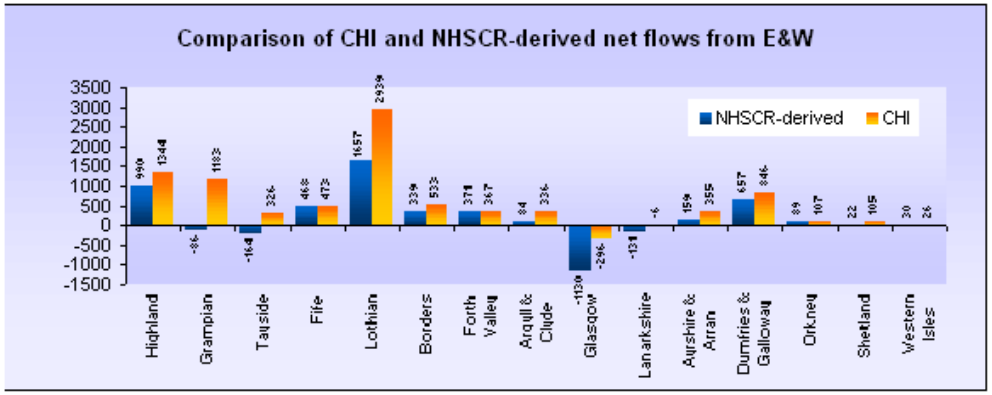


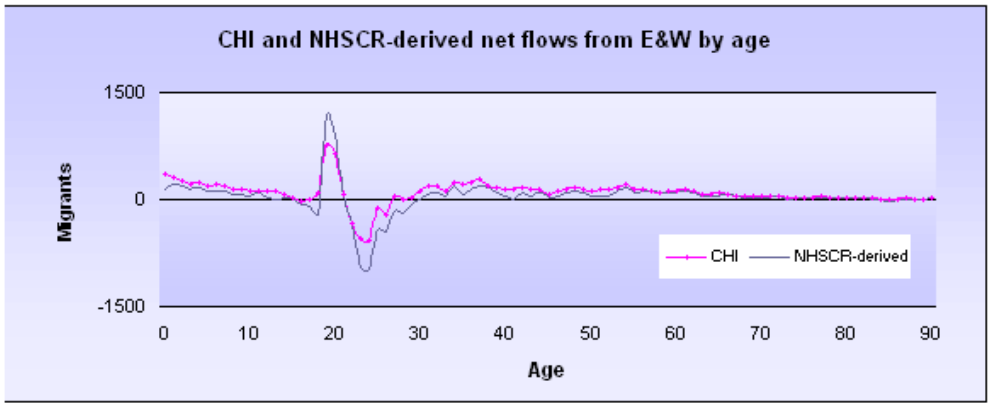




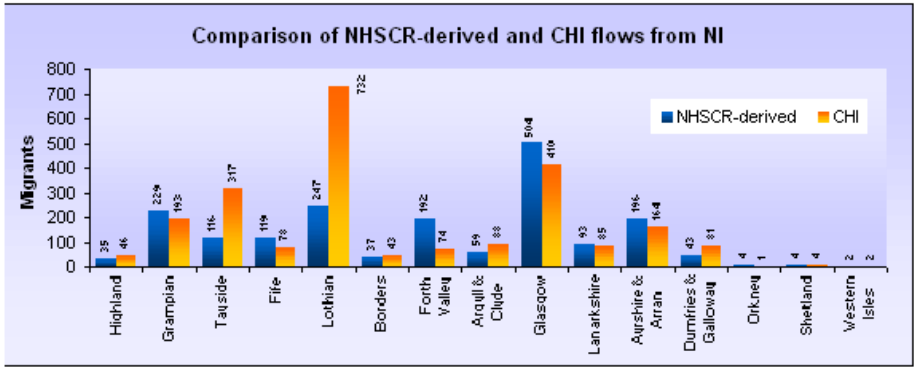


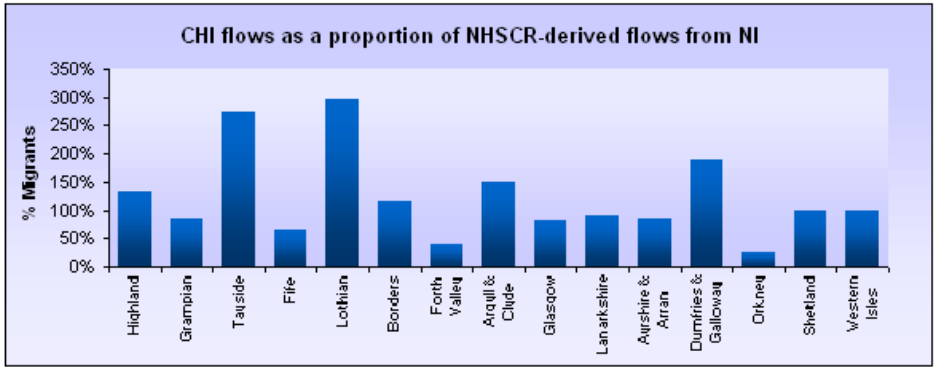
Comparison of raw CHI and NHSCR-derived NETFLOWS between SCOTLAND and ENGLAND & WALES, 2000-2001

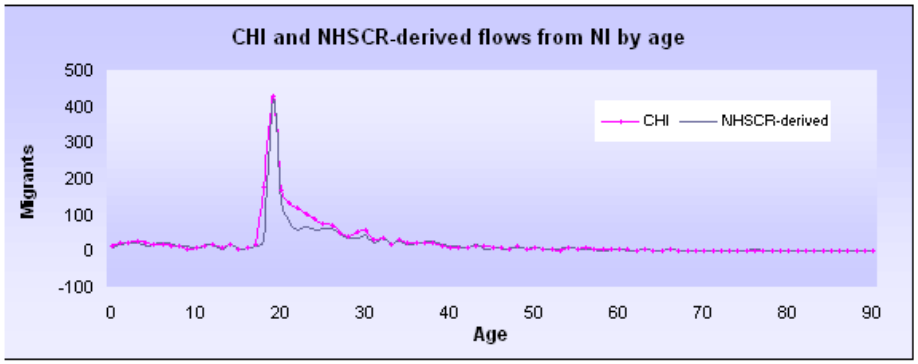


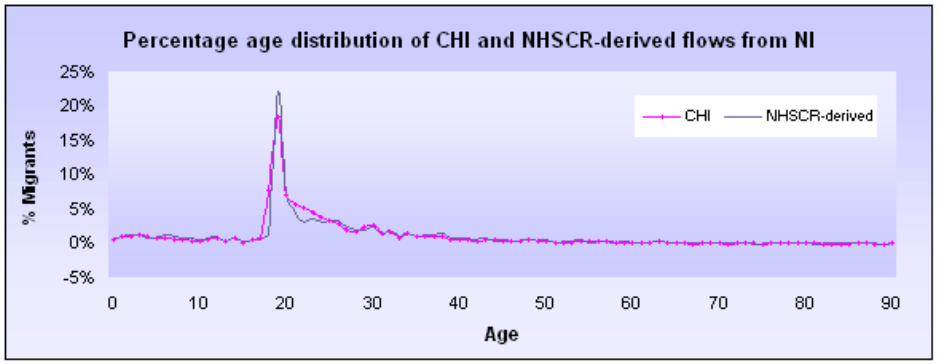


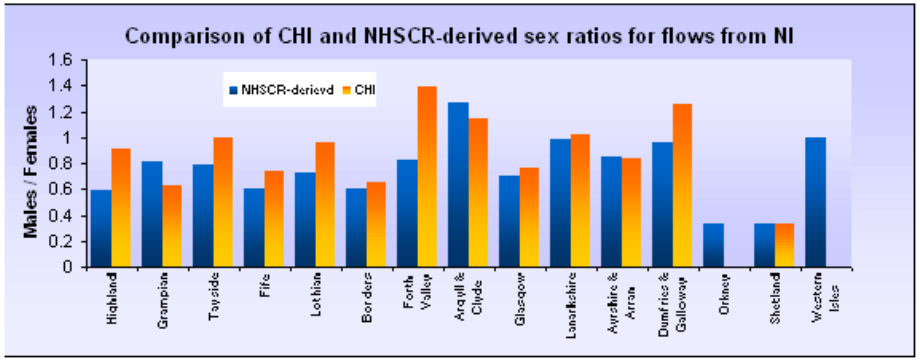
Comparison of raw CHI and NHSCR-derived INFLOWS into SCOTLAND from NORTHERN IRELAND, 2000-2001



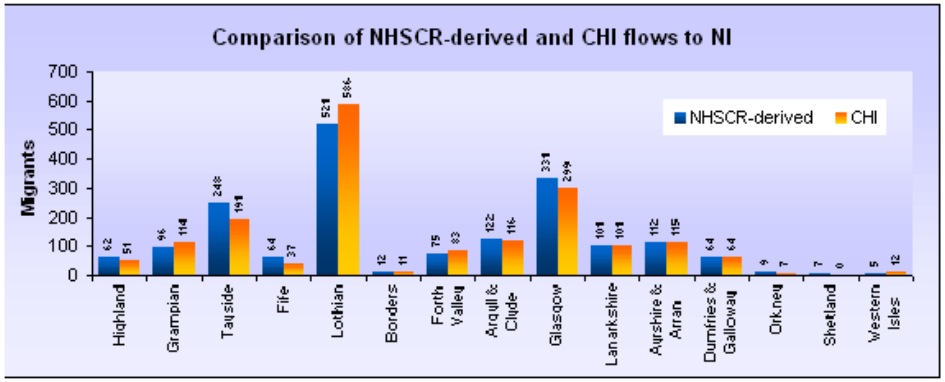


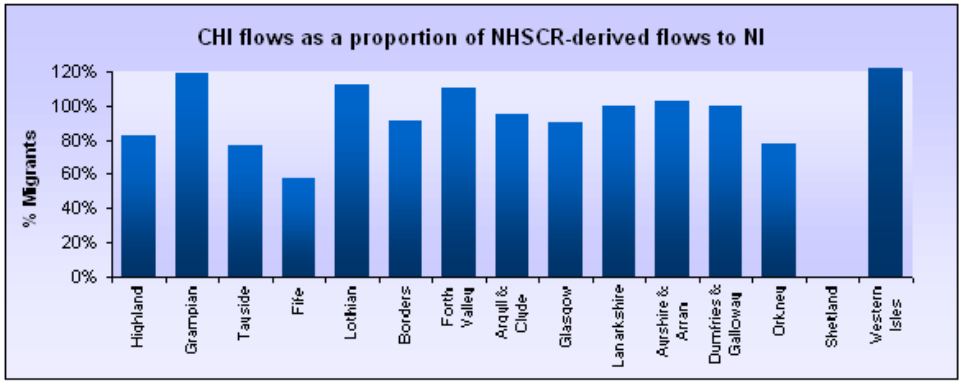


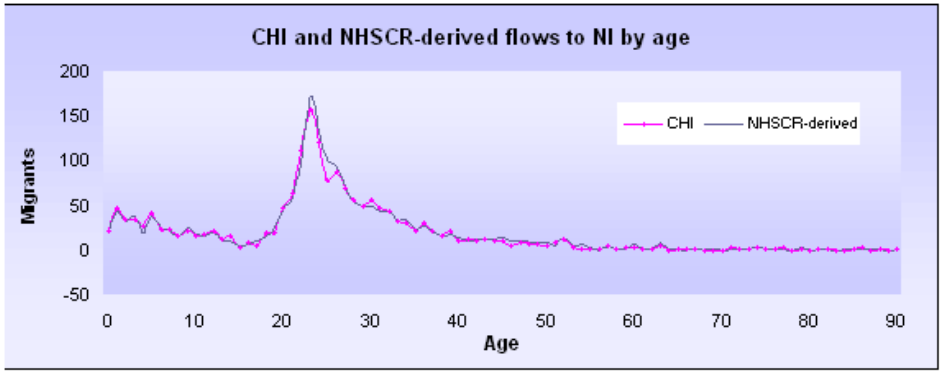


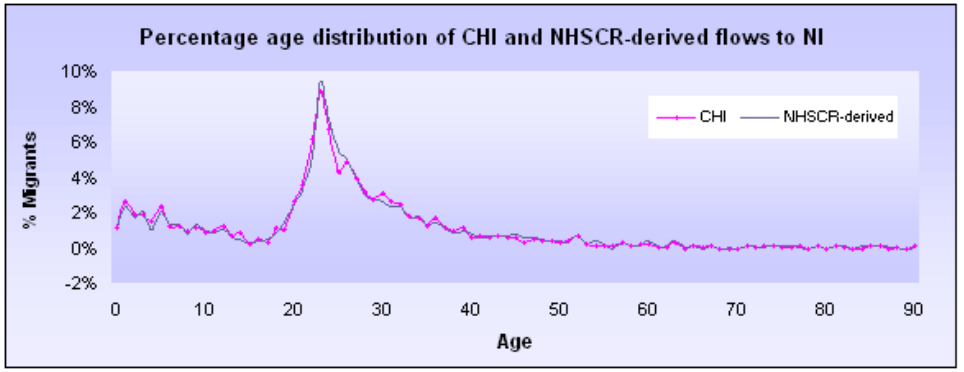


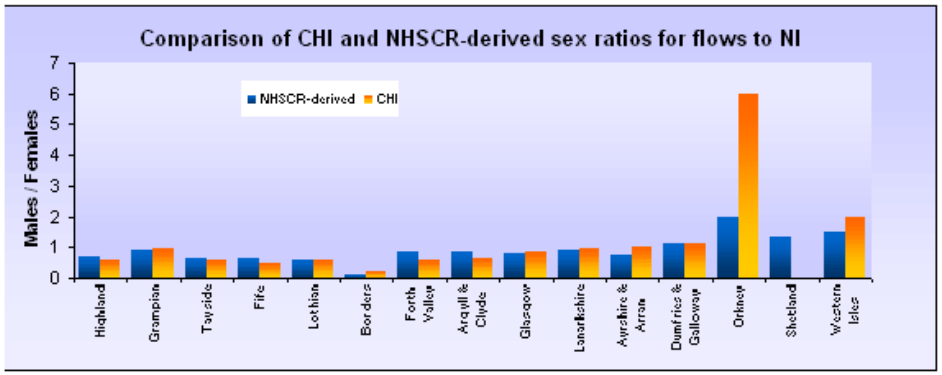
Comparison of raw CHI and NHSCR-derived OUTFLOWS from SCOTLAND into NORTHERN IRELAND, 2000-2001



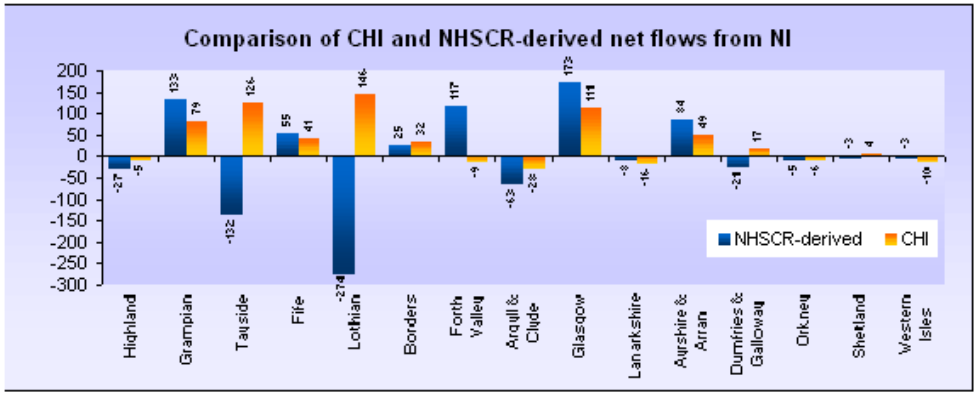


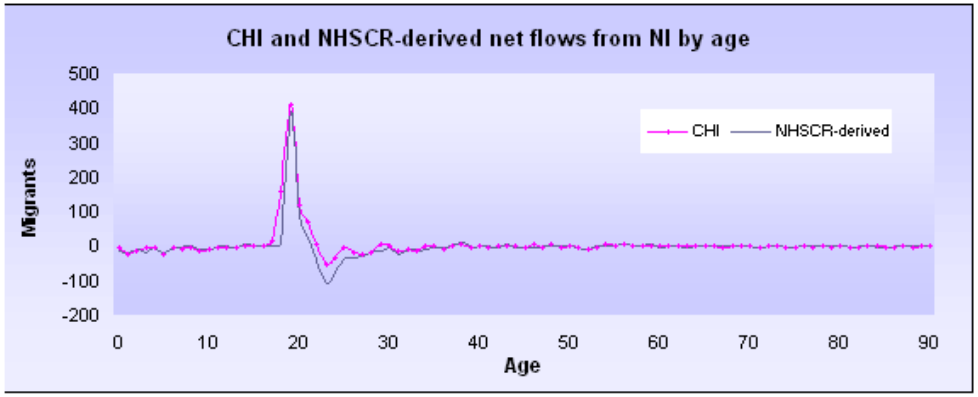




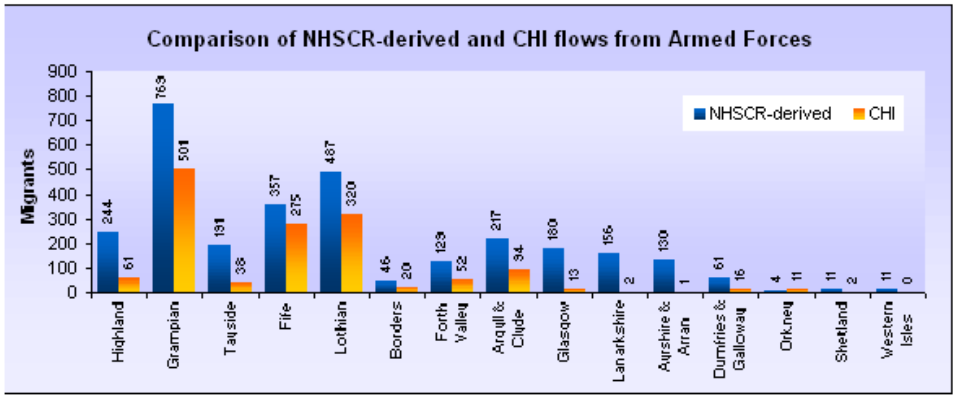


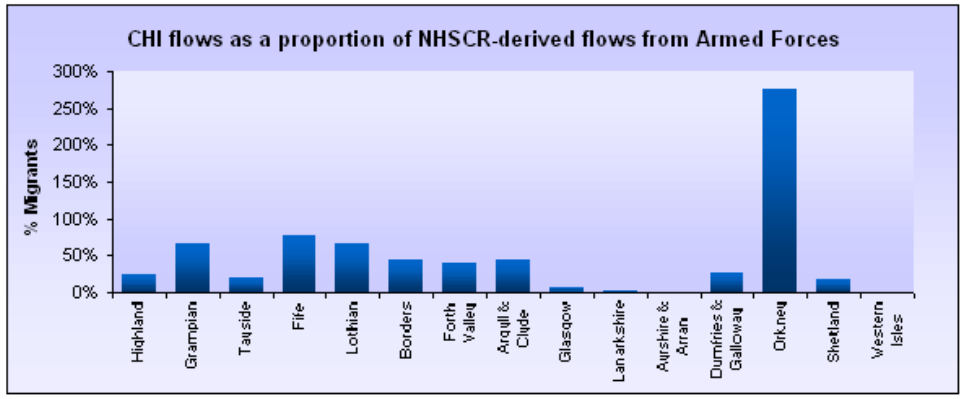
Comparison of raw CHI and NHSCR-derived NETFLOWS between SCOTLAND and NORTHERN IRELAND, 2000-2001

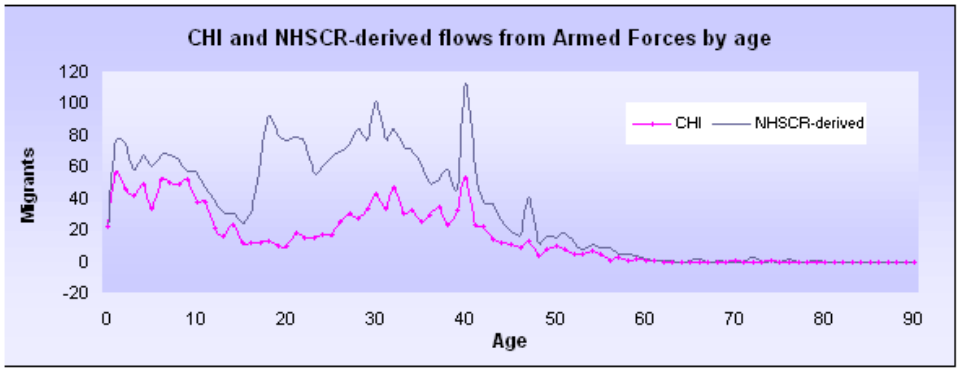


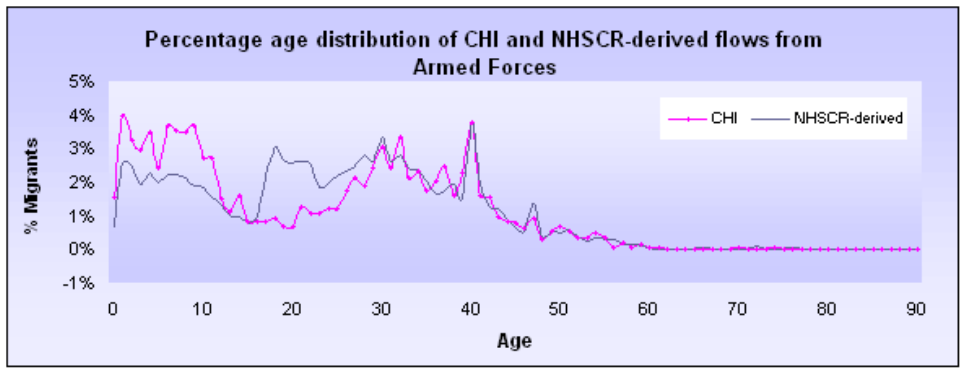


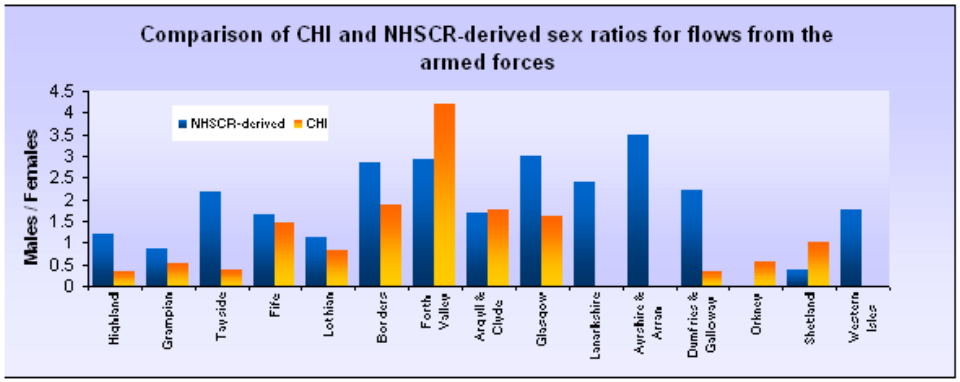
Comparison of raw CHI and NHSCR-derived INFLOWS into SCOTLAND from ARMED FORCES, 2000-2001



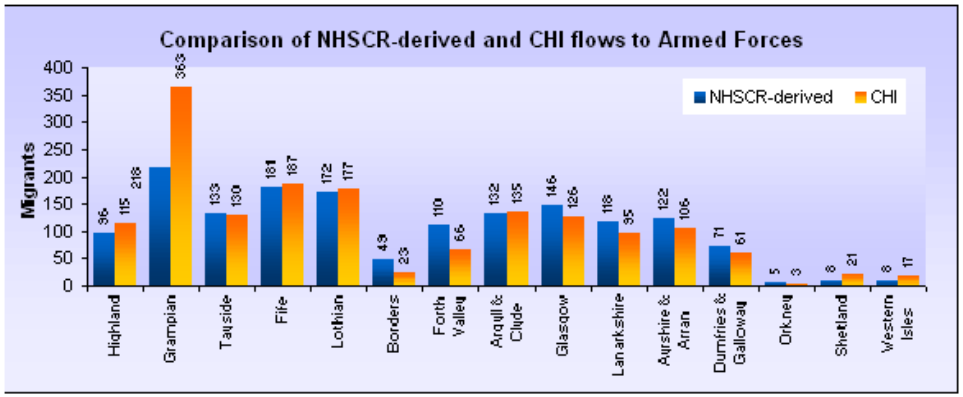


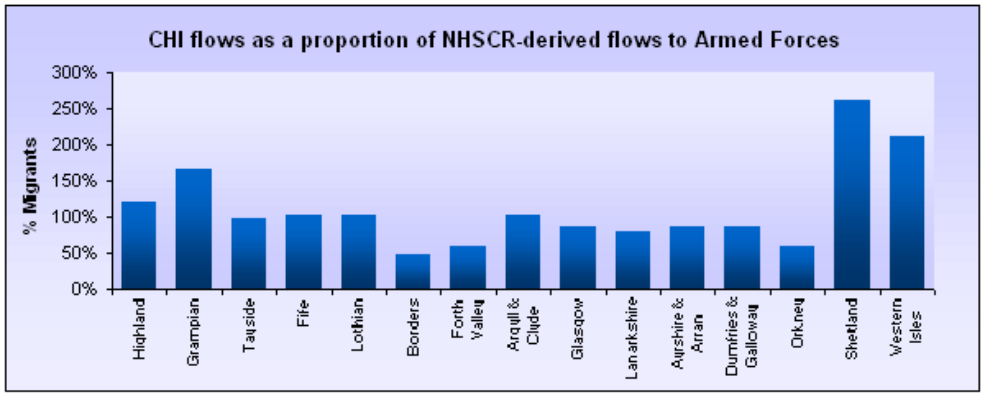


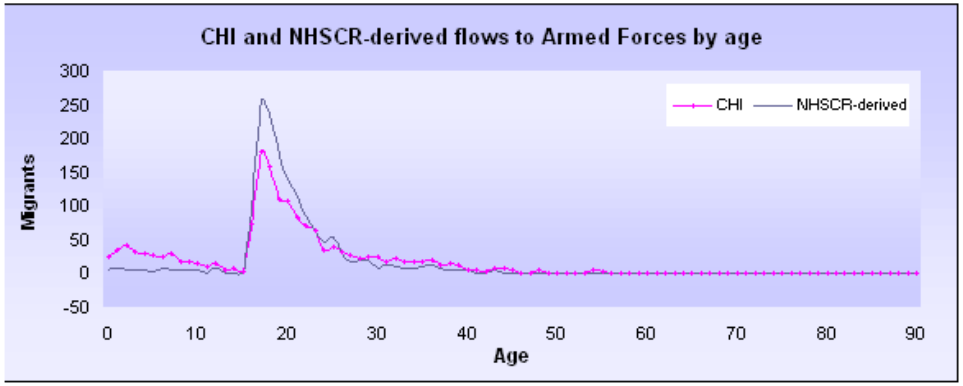


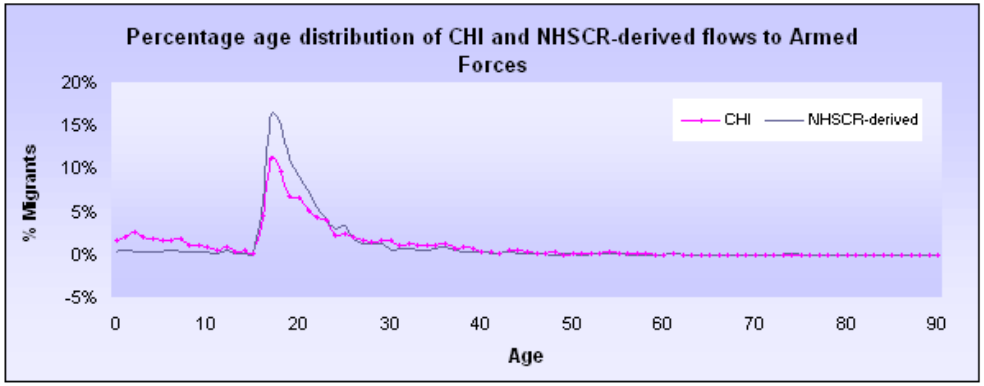


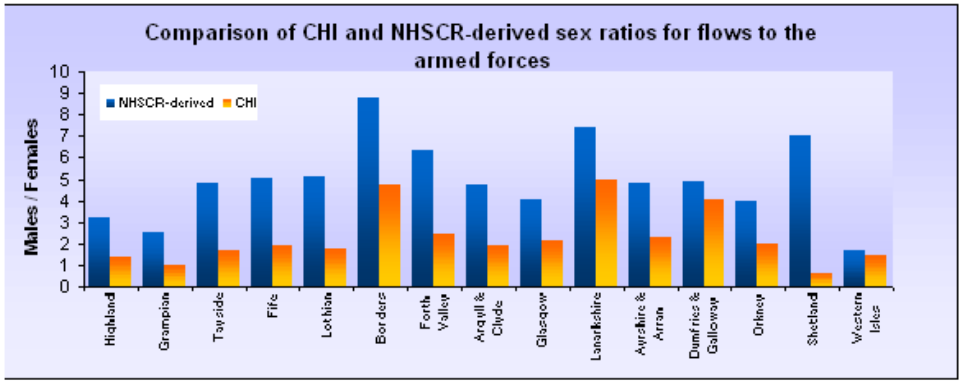
Comparison of raw CHI and NHSCR-derived OUTFLOWS from SCOTLAND to ARMED FORCES, 2000-2001



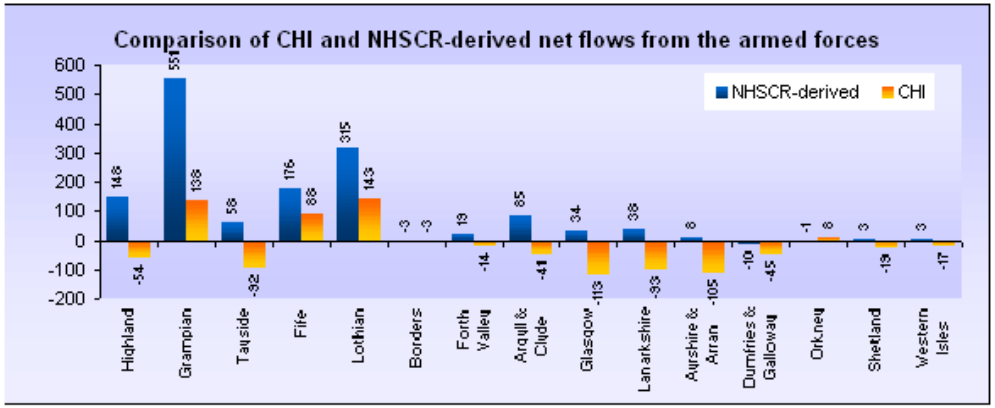


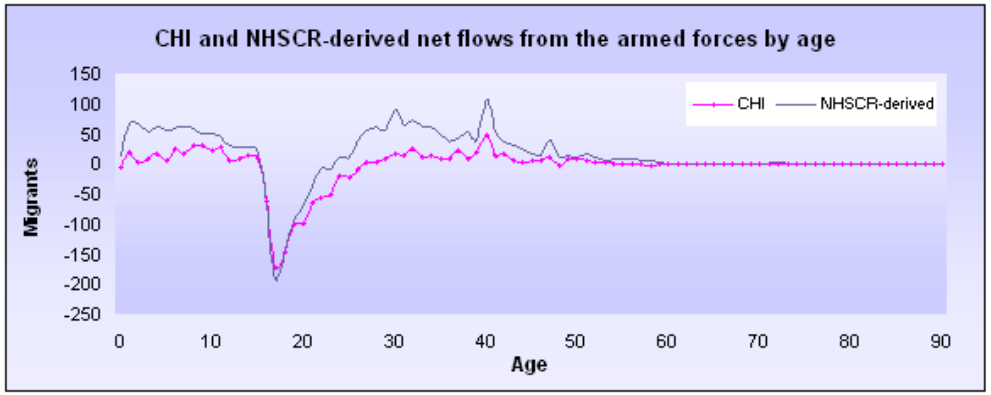




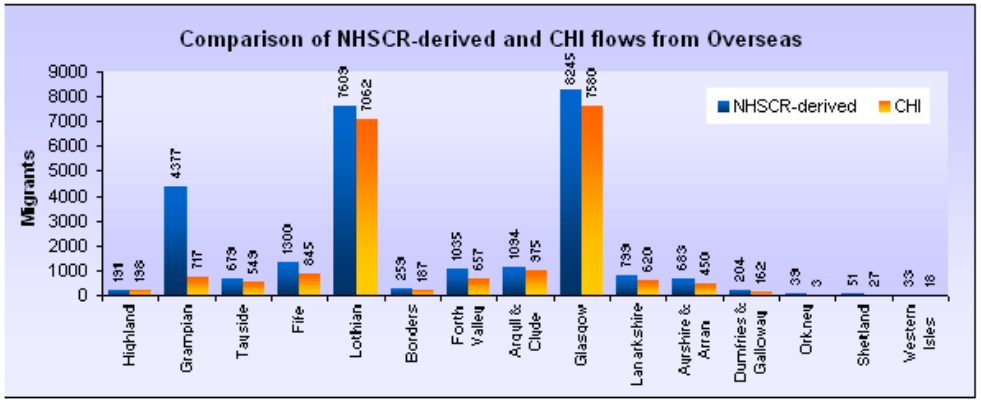


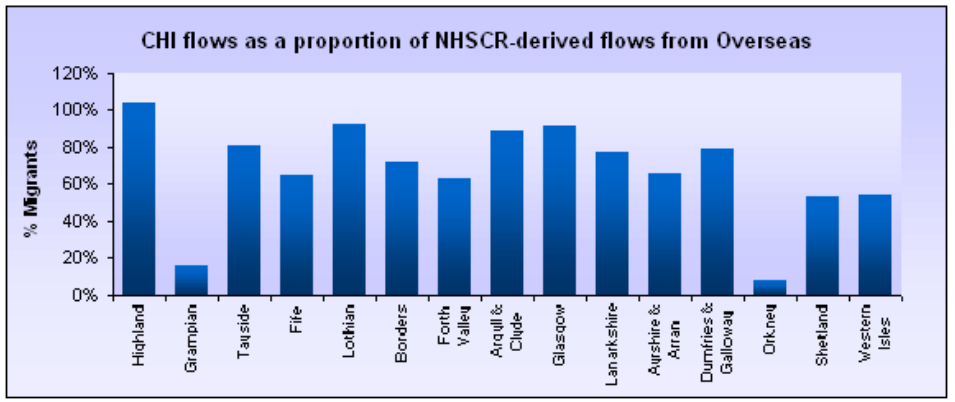
Comparison of raw CHI and NHSCR-derived NETFLOWS between SCOTLAND and ARMED FORCES, 2000-2001

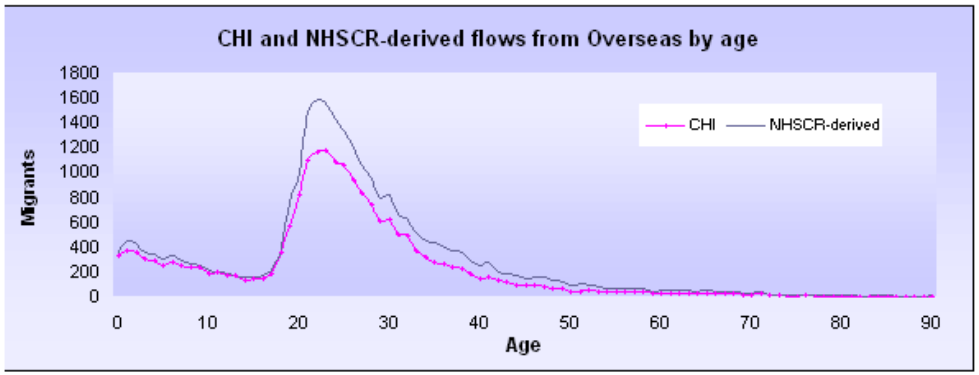


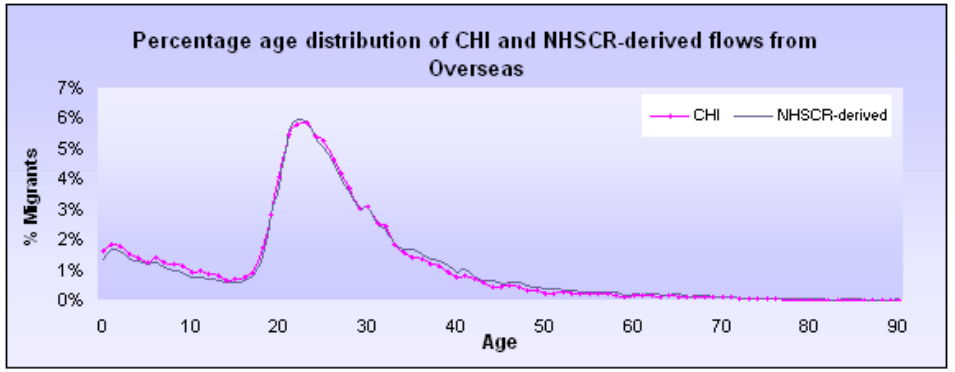


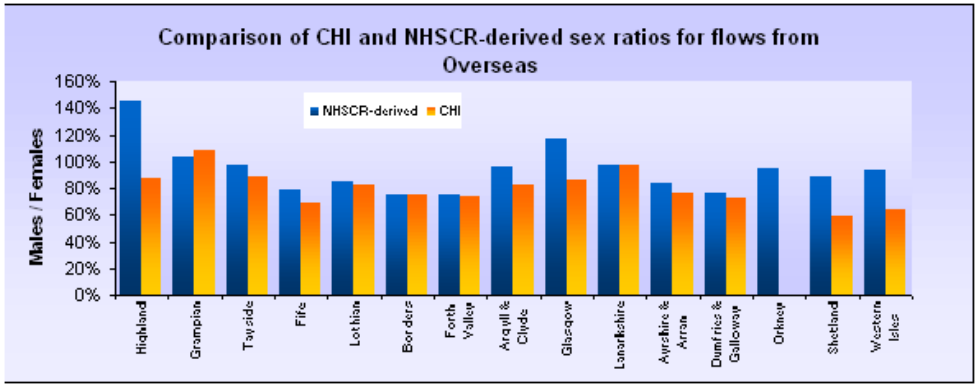
Comparison of raw CHI and NHSCR-derived INFLOWS into SCOTLAND from OVERSEAS, 2000-2001



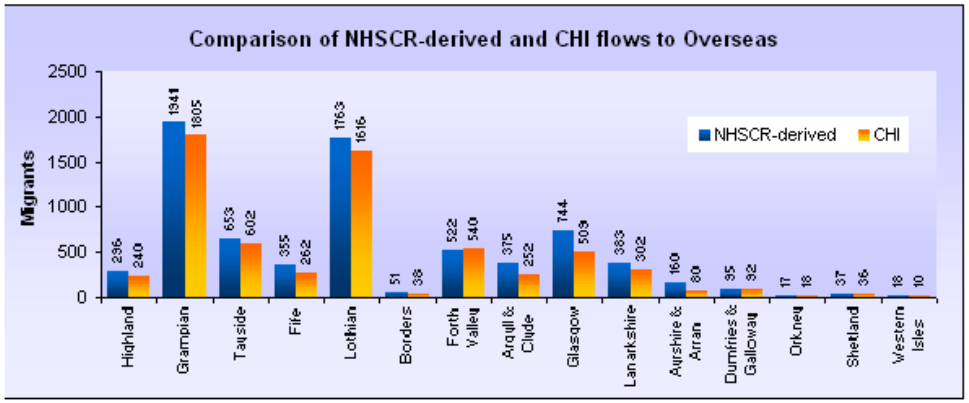


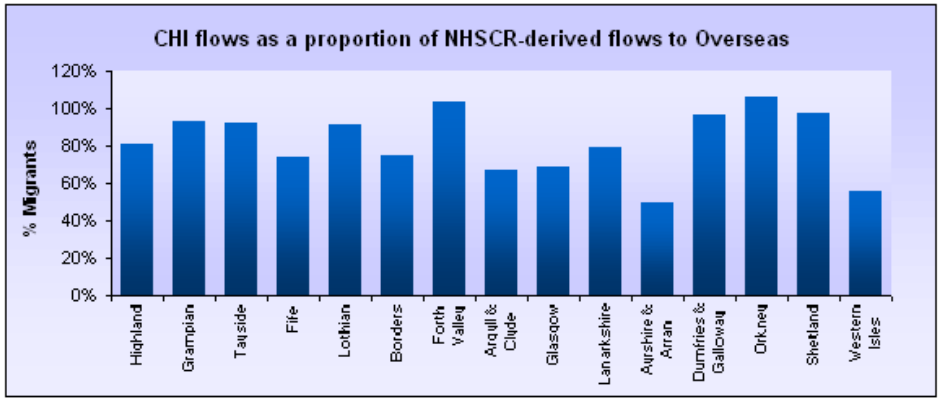


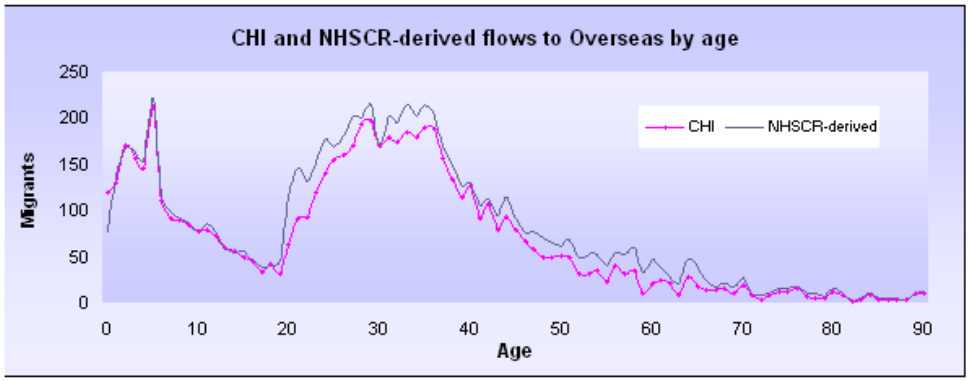


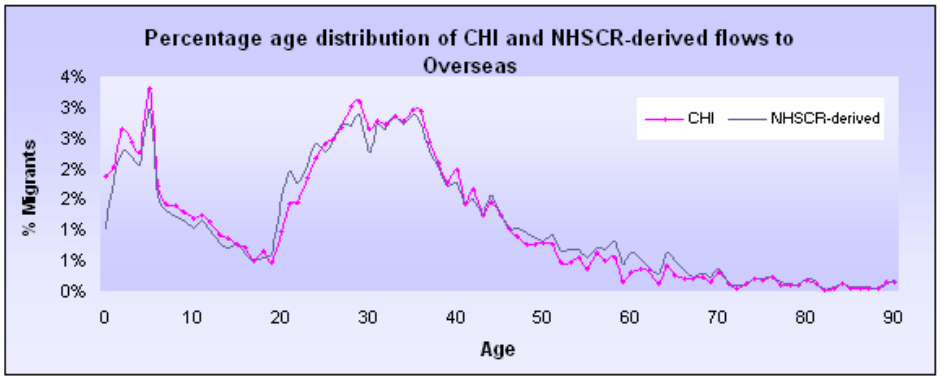


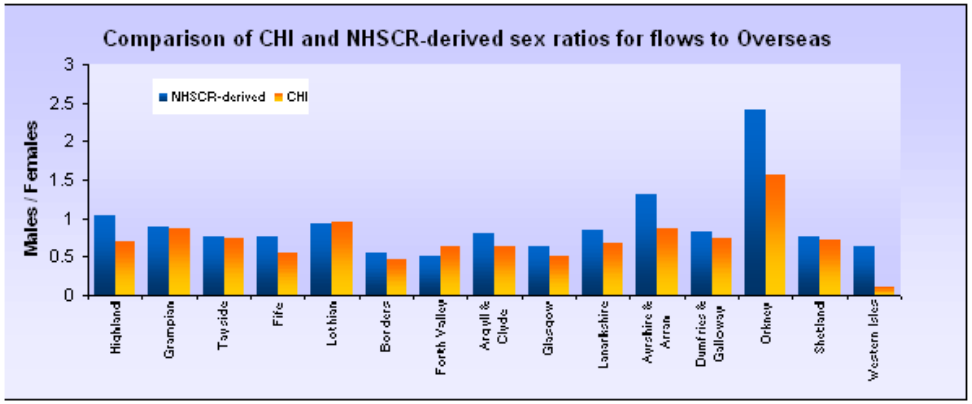
Comparison of raw CHI and NHSCR-derived OUTFLOWS from SCOTLAND to OVERSEAS, 2000-2001



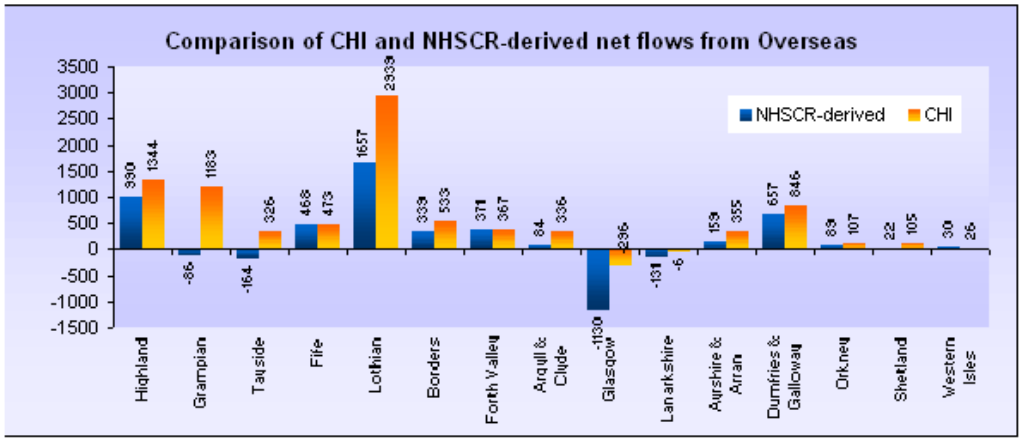


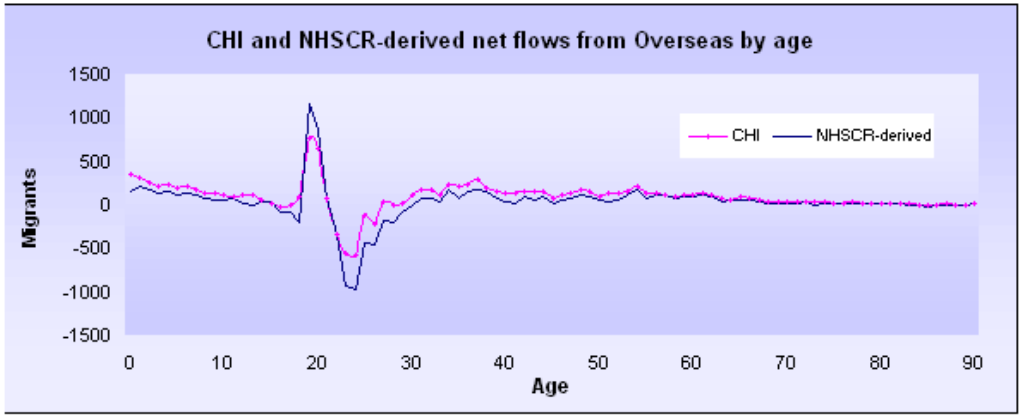






Comparison of raw CHI and NHSCR-derived NETFLOWS between SCOTLAND and OVERSEAS, 2000-2001





Appendix C – Imputation methodology

Stage 1 – within Scotland migrants, where origin postcode missing

The matching hierarchy is shown in Table 1a. If there are no potential donor records with all 8 variables matching (ie. type 9), then destination postcode is discarded as a matching variable, and a match is sought from the set of records with matching origin and destination health boards and GP practices, age group, sex and destination council (type 8) – and so on. Once a match is found, its postcode is assigned to the recipient record, but if more than one match is found, a postcode is chosen randomly from all available matched records.

Table 1a. Origin postcode missing

	Imputation Type								
Variables available for matching	Type 9	Type 8	Type 7	Type 6	Type 5	Type 4	Type 3	Type 2	Type 1
Origin CHI HB	✓	✓	✓	✓	✓	✓	✓	✓	✓
Destination HB	✓	✓	✓	✓	✓	✓	✓	✓	
Origin GP Practice	✓	✓	✓	✓	✓	✓			
Destination GP Practice	✓	✓	✓	✓	✓		✓		
Age Group	✓	✓	✓	✓					
Sex	✓	✓	✓						
Destination Council	✓	✓							
Destination Postcode	✓								

Stage 2 – within Scotland migrants, where destination postcode missing

The process for each stage is the same, the only difference is the variables available for matching, and the order in which they are used. See table 1b.

Table 1b. Destination postcode missing

	Imputation Type								
Variables available for matching	Type 9	Type 8	Type 7	Type 6	Type 5	Type 4	Type 3	Type 2	Type 1
Destination CHI HB	✓	✓	✓	✓	✓	✓	✓	✓	✓
Origin HB	✓	✓	✓	✓	✓	✓	✓	✓	
Destination GP Practice	✓	✓	✓	✓	✓	✓			
Origin GP Practice	✓	✓	✓	✓	✓		✓		
Age Group	✓	✓	✓	✓					
Sex	✓	✓	✓						
Origin Council	✓	✓							
Origin Postcode	✓								

Stage 3 – within Scotland migrants, where both postcodes missing

Table 1c. Both postcodes missing

	Imputation Type						
Variables available for matching	Type 7	Type 6	Type 5	Type 4	Type 3	Type 2	Type 1
Origin CHI HB	✓	✓	✓	✓	✓	✓	✓
Destination CHI HB	✓	✓	✓	✓	✓	✓	
Origin GP Practice	✓	✓	✓	✓			
Destination GP Practice	✓	✓	✓		✓		
Age Group	✓	✓					
Sex	✓						

Stage 4 – Transfers out of Scotland, where origin postcode missing

Table 1d. Transfers Out - Origin postcode missing

	Imputation Type				
Variables available for matching	Type 5	Type 4	Type 3	Type 2	Type 1
Destination HB (16)	✓	✓	✓	✓	✓
Origin CHI HB	✓	✓	✓	✓	
Origin GP Practice	✓	✓	✓		
Age Group	✓	✓			
Sex	✓				

Stage 5 – Transfers into Scotland, where destination postcode missing

Table 1e. Transfers In - Destination postcode missing

	Imputation Type				
Variables available for matching	Type 5	Type 4	Type 3	Type 2	Type 1
Origin HB (16)	✓	✓	✓	✓	✓
Destination CHI HB	✓	✓	✓	✓	
Destination GP Practice	✓	✓	✓		
Age Group	✓	✓			
Sex	✓				

Appendix D

Table 1a Number of records in origin sample with correctly imputed geographic information

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	110	110	110	105	71	107	83	73	57
Grampian	33	33	33	31	18	32	22	20	17
Tayside	182	182	182	180	90	179	132	92	72
Fife	64	64	64	59	34	64	52	36	27
Lothian	347	347	347	335	177	345	236	189	128
Borders	50	50	50	49	32	49	43	36	20
Forth Valley	64	64	64	57	24	61	42	30	17
Argyll & Clyde	50	50	50	49	28	49	31	26	10
Greater Glasgow	325	325	325	316	143	325	205	131	107
Lanarkshire	78	78	78	74	40	71	59	44	31
Ayrshire & Arran	70	70	70	64	38	70	45	40	29
Dumfries & Galloway	19	19	19	18	11	18	16	12	9
Orkney	5	5	5	5	1	5	2	2	0
Shetland	5	5	5	5	5	5	5	5	5
Western Isles	9	9	9	9	3	9	4	4	5
Total	1,411	1,411	1,411	1,356	715	1,389	977	740	534

Table 1b Percentage of records in origin sample with correctly imputed geographic information

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	110	100	100	95	65	97	75	66	52
Grampian	33	100	100	94	55	97	67	61	52
Tayside	182	100	100	99	49	98	73	51	40
Fife	64	100	100	92	53	100	81	56	42
Lothian	347	100	100	97	51	99	68	54	37
Borders	50	100	100	98	64	98	86	72	40
Forth Valley	64	100	100	89	38	95	66	47	27
Argyll & Clyde	50	100	100	98	56	98	62	52	20
Greater Glasgow	325	100	100	97	44	100	63	40	33
Lanarkshire	78	100	100	95	51	91	76	56	40
Ayrshire & Arran	70	100	100	91	54	100	64	57	41
Dumfries & Galloway	19	100	100	95	58	95	84	63	47
Orkney	5	100	100	100	20	100	40	40	0
Shetland	5	100	100	100	100	100	100	100	100
Western Isles	9	100	100	100	33	100	44	44	56
Total	1,411	100	100	96	51	98	69	52	38

Table 1c Number of records in origin sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	139	139	139	136	74	135	95	79	47
5-9	90	90	90	90	48	89	66	47	31
10-14	70	70	70	65	36	69	49	36	34
15-19	83	83	83	77	48	81	56	46	34
20-24	164	164	164	159	78	163	105	79	54
25-29	175	175	175	166	78	174	113	88	64
30-34	155	155	155	148	67	152	101	69	54
35-39	109	109	109	104	56	105	77	56	36
40-44	79	79	79	75	43	78	59	46	32
45-49	68	68	68	64	34	68	51	39	33
50-54	45	45	45	44	27	44	36	25	19
55-59	45	45	45	45	27	45	34	28	20
60-64	40	40	40	39	19	39	27	21	10
65-69	28	28	28	27	15	27	17	16	14
70-74	24	24	24	24	15	24	21	16	9
75+	97	97	97	93	50	96	70	49	43
Total	1,411	1,411	1,411	1,356	715	1,389	977	740	534

Table 1d Percentage of records in origin sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	139	100	100	98	53	97	68	57	34
5-9	90	100	100	100	53	99	73	52	34
10-14	70	100	100	93	51	99	70	51	49
15-19	83	100	100	93	58	98	67	55	41
20-24	164	100	100	97	48	99	64	48	33
25-29	175	100	100	95	45	99	65	50	37
30-34	155	100	100	95	43	98	65	45	35
35-39	109	100	100	95	51	96	71	51	33
40-44	79	100	100	95	54	99	75	58	41
45-49	68	100	100	94	50	100	75	57	49
50-54	45	100	100	98	60	98	80	56	42
55-59	45	100	100	100	60	100	76	62	44
60-64	40	100	100	98	48	98	68	53	25
65-69	28	100	100	96	54	96	61	57	50
70-74	24	100	100	100	63	100	88	67	38
75+	97	100	100	96	52	99	72	51	44
Total	1,411	100	100	96	51	98	69	52	38

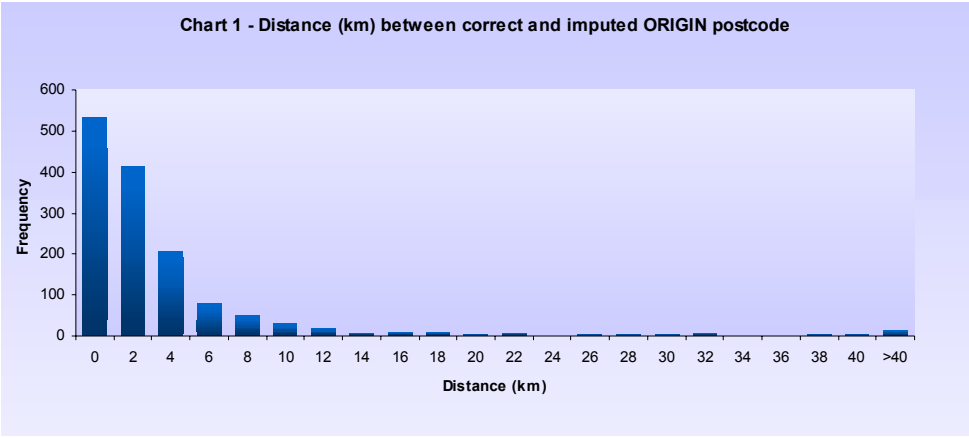


Table 2a **Number of records in destination sample with correctly imputed geographic information**

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	106	106	106	97	66	100	78	72	50
Grampian	106	106	106	103	55	105	69	54	41
Tayside	210	210	203	206	120	207	167	126	93
Fife	128	128	128	126	77	127	108	91	59
Lothian	391	391	380	380	186	389	248	197	135
Borders	49	49	49	46	36	49	44	41	23
Forth Valley	131	131	127	127	62	128	92	74	45
Argyll & Clyde	84	84	75	79	39	77	58	42	28
Greater Glasgow	232	232	225	226	108	232	141	105	75
Lanarkshire	472	472	468	460	244	462	393	272	195
Ayrshire & Arran	195	195	190	190	110	194	146	114	90
Dumfries & Galloway	16	16	16	13	10	16	13	10	7
Orkney	0	0	0	0	0	0	0	0	0
Shetland	9	9	9	9	6	9	8	8	5
Western Isles	22	22	22	22	17	22	18	18	14
Total	2,151	2,151	2,104	2,084	1,136	2,117	1,583	1,224	860

Table 2b **Percentage of records in destination sample with correctly imputed geographic information**

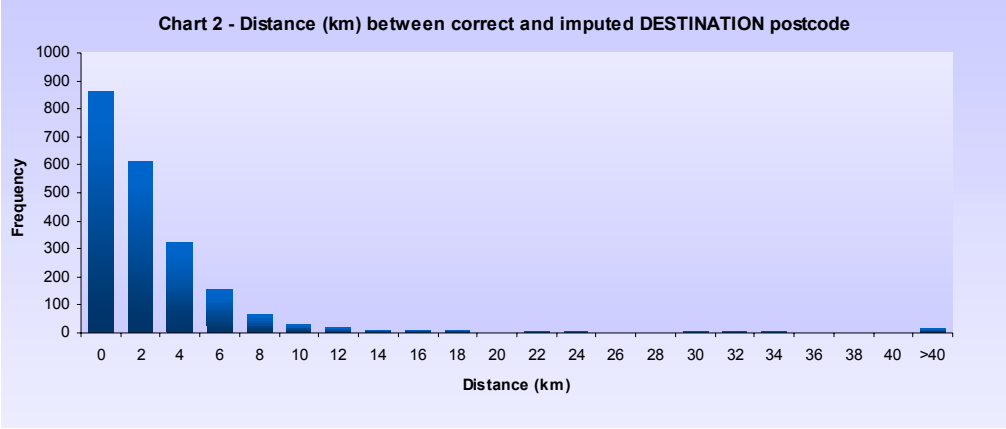
Health Board	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	106	100	100	92	62	94	74	68	47
Grampian	106	100	100	97	52	99	65	51	39
Tayside	210	100	97	98	57	99	80	60	44
Fife	128	100	100	98	60	99	84	71	46
Lothian	391	100	97	97	48	99	63	50	35
Borders	49	100	100	94	73	100	90	84	47
Forth Valley	131	100	97	97	47	98	70	56	34
Argyll & Clyde	84	100	89	94	46	92	69	50	33
Greater Glasgow	232	100	97	97	47	100	61	45	32
Lanarkshire	472	100	99	97	52	98	83	58	41
Ayrshire & Arran	195	100	97	97	56	99	75	58	46
Dumfries & Galloway	16	100	100	81	63	100	81	63	44
Orkney	0	0	0	0	0	0	0	0	0
Shetland	9	100	100	100	67	100	89	89	56
Western Isles	22	100	100	100	77	100	82	82	64
Total	2,151	100	98	97	53	98	74	57	40

Table 2c Number of records in destination sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	255	255	249	247	123	252	201	145	71
5-9	126	126	123	122	68	123	94	78	54
10-14	129	129	128	128	78	127	104	80	58
15-19	178	178	170	168	87	175	115	91	72
20-24	224	224	220	219	112	221	152	117	90
25-29	228	228	223	222	102	225	152	106	80
30-34	222	222	219	213	130	217	172	141	100
35-39	177	177	171	170	96	173	133	103	80
40-44	135	135	133	132	77	133	101	81	56
45-49	91	91	88	87	50	88	68	54	36
50-54	67	67	66	64	43	66	52	44	33
55-59	64	64	64	63	36	64	50	38	28
60-64	54	54	53	52	25	54	36	28	19
65-69	27	27	26	27	19	27	23	19	17
70-74	41	41	41	41	19	41	31	22	15
75+	133	133	130	129	71	131	99	77	51
Total	2,151	2,151	2,104	2,084	1,136	2,117	1,583	1,224	860

Table 2d Percentage of records in destination sample with correctly imputed geographic information

Age Group	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	255	100	98	97	48	99	79	57	28
5-9	126	100	98	97	54	98	75	62	43
10-14	129	100	99	99	60	98	81	62	45
15-19	178	100	96	94	49	98	65	51	40
20-24	224	100	98	98	50	99	68	52	40
25-29	228	100	98	97	45	99	67	46	35
30-34	222	100	99	96	59	98	77	64	45
35-39	177	100	97	96	54	98	75	58	45
40-44	135	100	99	98	57	99	75	60	41
45-49	91	100	97	96	55	97	75	59	40
50-54	67	100	99	96	64	99	78	66	49
55-59	64	100	100	98	56	100	78	59	44
60-64	54	100	98	96	46	100	67	52	35
65-69	27	100	96	100	70	100	85	70	63
70-74	41	100	100	100	46	100	76	54	37
75+	133	100	98	97	53	98	74	58	38
Total	2,151	100	98	97	53	98	74	57	40



1. Imputation accuracy for missing ORIGIN postcode, where both postcodes missing

Table 3a1 **Number of records in sample with correctly imputed geographic information**

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	276	276	276	264	115	272	167	131	12
Grampian	79	79	74	73	28	79	50	31	0
Tayside	311	311	291	305	69	308	219	107	4
Fife	295	295	295	289	95	293	227	127	5
Lothian	1,366	1,366	1,352	1,352	374	1,360	839	482	32
Borders	155	155	155	153	68	154	115	68	8
Forth Valley	363	363	361	361	115	360	257	159	7
Argyll & Clyde	121	121	116	119	44	120	82	54	6
Greater Glasgow	1,040	1,040	933	964	181	1,040	502	212	16
Lanarkshire	721	721	696	679	147	701	518	226	14
Ayrshire & Arran	457	457	432	427	123	457	316	175	2
Dumfries & Galloway	34	34	34	31	10	34	29	12	1
Orkney	16	16	16	16	6	16	11	11	0
Shetland	2	2	2	2	0	2	1	1	0
Western Isles	99	99	99	99	35	99	72	71	3
Total	5,335	5,335	5,132	5,134	1,410	5,295	3,405	1,867	110

Table 3b1 **Percentage of records in sample with correctly imputed geographic information**

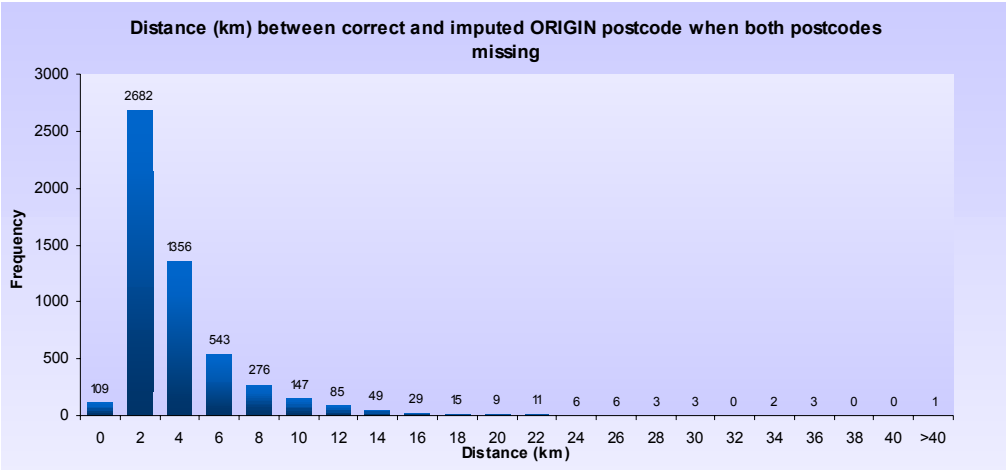
Health Board	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	276	100	100	96	42	99	61	47	4
Grampian	79	100	94	92	35	100	63	39	0
Tayside	311	100	94	98	22	99	70	34	1
Fife	295	100	100	98	32	99	77	43	2
Lothian	1,366	100	99	99	27	100	61	35	2
Borders	155	100	100	99	44	99	74	44	5
Forth Valley	363	100	99	99	32	99	71	44	2
Argyll & Clyde	121	100	96	98	36	99	68	45	5
Greater Glasgow	1,040	100	90	93	17	100	48	20	2
Lanarkshire	721	100	97	94	20	97	72	31	2
Ayrshire & Arran	457	100	95	93	27	100	69	38	0
Dumfries & Galloway	34	100	100	91	29	100	85	35	3
Orkney	16	100	100	100	38	100	69	69	0
Shetland	2	100	100	100	0	100	50	50	0
Western Isles	99	100	100	100	35	100	73	72	3
Total	5,335	100	96	96	26	99	64	35	2

Table 3c1 Number of records in sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	266	266	262	255	72	266	181	97	7
5-9	359	359	346	341	98	358	243	141	5
10-14	345	345	330	330	99	342	226	126	5
15-19	269	269	260	261	63	268	190	85	4
20-24	323	323	309	316	80	320	176	96	9
25-29	427	427	409	412	104	422	262	140	14
30-34	457	457	440	440	114	450	284	144	9
35-39	493	493	472	473	109	489	291	153	6
40-44	411	411	397	400	104	408	240	127	8
45-49	326	326	312	315	105	322	200	130	6
50-54	307	307	296	295	70	305	203	111	4
55-59	256	256	243	243	79	255	158	90	6
60-64	272	272	262	260	61	268	191	92	3
65-69	207	207	201	202	59	207	130	84	3
70-74	204	204	198	198	64	204	136	81	3
75+	413	413	395	393	129	411	294	170	18
Total	5,335	5,335	5,132	5,134	1,410	5,295	3,405	1,867	110

Table 3d1 Percentage of records in sample with correctly imputed geographic information

Age Group	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	266	100	98	96	27	100	68	36	3
5-9	359	100	96	95	27	100	68	39	1
10-14	345	100	96	96	29	99	66	37	1
15-19	269	100	97	97	23	100	71	32	1
20-24	323	100	96	98	25	99	54	30	3
25-29	427	100	96	96	24	99	61	33	3
30-34	457	100	96	96	25	98	62	32	2
35-39	493	100	96	96	22	99	59	31	1
40-44	411	100	97	97	25	99	58	31	2
45-49	326	100	96	97	32	99	61	40	2
50-54	307	100	96	96	23	99	66	36	1
55-59	256	100	95	95	31	100	62	35	2
60-64	272	100	96	96	22	99	70	34	1
65-69	207	100	97	98	29	100	63	41	1
70-74	204	100	97	97	31	100	67	40	1
75+	413	100	96	95	31	100	71	41	4
Total	5,335	100	96	96	26	99	64	35	2



2. Imputation accuracy for missing DESTINATION postcode, where both postcodes missing

Table 3a2 **Number of records in sample with correctly imputed geographic information**

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	277	276	276	265	118	273	170	136	13
Grampian	79	79	74	72	25	79	51	32	0
Tayside	310	311	291	305	65	308	223	111	4
Fife	295	295	295	289	95	294	232	129	5
Lothian	1,365	1,366	1,353	1,353	362	1,360	848	497	34
Borders	155	155	155	153	68	154	115	68	8
Forth Valley	364	363	362	362	114	360	261	163	9
Argyll & Clyde	120	121	117	119	46	120	84	56	6
Greater Glasgow	1,042	1,040	933	964	180	1,040	510	221	16
Lanarkshire	720	721	696	678	146	702	522	229	14
Ayrshire & Arran	457	457	430	425	121	457	321	180	3
Dumfries & Galloway	34	34	34	31	11	34	29	13	1
Orkney	16	16	16	16	6	16	11	11	0
Shetland	2	2	2	2	0	2	1	1	0
Western Isles	99	99	99	99	35	99	75	73	3
Total	5,335	5,335	5,133	5,133	1,392	5,298	3,453	1,920	116

Table 3b2 **Percentage of records in sample with correctly imputed geographic information**

Health Board	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	277	100	100	96	43	99	62	49	5
Grampian	79	100	94	91	32	100	65	41	0
Tayside	310	100	94	98	21	99	72	36	1
Fife	295	100	100	98	32	100	79	44	2
Lothian	1,365	100	99	99	27	100	62	36	2
Borders	155	100	100	99	44	99	74	44	5
Forth Valley	364	100	100	100	31	99	72	45	2
Argyll & Clyde	120	100	97	98	38	99	69	46	5
Greater Glasgow	1,042	100	90	93	17	100	49	21	2
Lanarkshire	720	100	97	94	20	97	72	32	2
Ayrshire & Arran	457	100	94	93	26	100	70	39	1
Dumfries & Galloway	34	100	100	91	32	100	85	38	3
Orkney	16	100	100	100	38	100	69	69	0
Shetland	2	100	100	100	0	100	50	50	0
Western Isles	99	100	100	100	35	100	76	74	3
Total	5,335	100	96	96	26	99	65	36	2

Table 3c2 Number of records in sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	266	266	262	254	69	266	183	98	7
5-9	359	359	345	340	95	358	248	144	5
10-14	345	345	330	330	102	342	231	134	5
15-19	269	269	262	262	62	269	192	90	5
20-24	323	323	311	317	78	321	184	106	12
25-29	427	427	408	410	99	422	268	147	14
30-34	457	457	440	443	114	450	293	152	9
35-39	493	493	471	472	105	490	295	157	6
40-44	411	411	397	399	98	408	240	127	8
45-49	326	326	312	315	105	322	200	131	6
50-54	307	307	297	296	72	305	206	114	4
55-59	256	256	243	243	81	255	158	90	6
60-64	272	272	262	260	62	268	191	93	3
65-69	207	207	201	202	59	207	131	84	3
70-74	204	204	197	197	63	204	137	81	3
75+	413	413	395	393	128	411	296	172	20
Total	5,335	5,335	5,133	5,133	1,392	5,298	3,453	1,920	116

Table 3d2 Percentage of records in sample with correctly imputed geographic information

Age Group	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	266	100	98	95	26	100	69	37	3
5-9	359	100	96	95	26	100	69	40	1
10-14	345	100	96	96	30	99	67	39	1
15-19	269	100	97	97	23	100	71	33	2
20-24	323	100	96	98	24	99	57	33	4
25-29	427	100	96	96	23	99	63	34	3
30-34	457	100	96	97	25	98	64	33	2
35-39	493	100	96	96	21	99	60	32	1
40-44	411	100	97	97	24	99	58	31	2
45-49	326	100	96	97	32	99	61	40	2
50-54	307	100	97	96	23	99	67	37	1
55-59	256	100	95	95	32	100	62	35	2
60-64	272	100	96	96	23	99	70	34	1
65-69	207	100	97	98	29	100	63	41	1
70-74	204	100	97	97	31	100	67	40	1
75+	413	100	96	95	31	100	72	42	5
Total	5,335	100	96	96	26	99	65	36	2

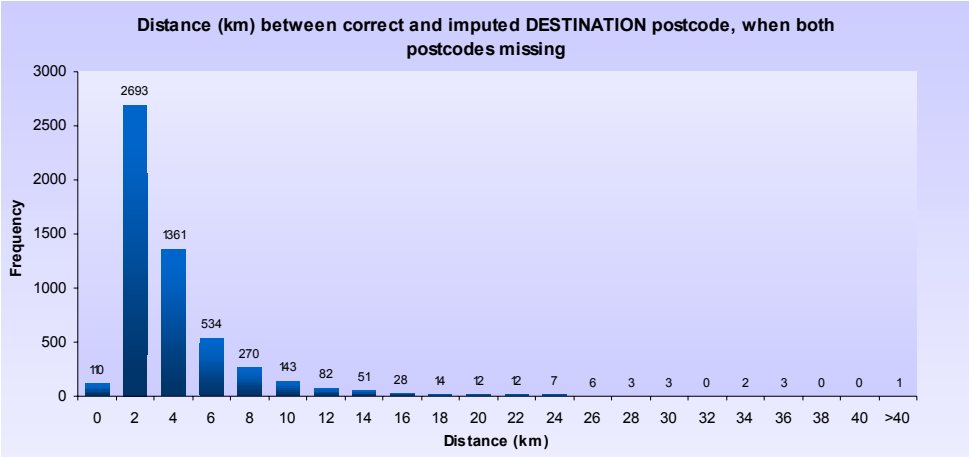


Table 4a Number of records in 'Transfer Out' sample with correctly imputed geographic information

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	1	1	1	1	1	1	1	0	0
Grampian	2	2	2	2	0	2	1	0	0
Tayside	5	5	5	5	1	5	2	2	1
Fife	4	4	4	4	2	4	4	3	0
Lothian	32	32	30	30	11	32	14	8	2
Borders	3	3	3	3	3	3	3	3	0
Forth Valley	9	9	8	8	5	9	5	3	0
Argyll & Clyde	0	0	0	0	0	0	0	0	0
Greater Glasgow	8	8	7	7	3	8	3	3	1
Lanarkshire	6	6	4	4	1	5	3	1	0
Ayrshire & Arran	3	3	3	3	2	3	2	1	1
Dumfries & Galloway	0	0	0	0	0	0	0	0	0
Orkney	0	0	0	0	0	0	0	0	0
Shetland	0	0	0	0	0	0	0	0	0
Western Isles	1	1	1	1	0	1	0	0	0
Total	74	74	68	68	29	73	38	24	5

Table 4b Percentage of records in 'Transfer Out' sample with correctly imputed geographic information

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	1	100	100	100	100	100	100	0	0
Grampian	2	100	100	100	0	100	50	0	0
Tayside	5	100	100	100	20	100	40	40	20
Fife	4	100	100	100	50	100	100	75	0
Lothian	32	100	94	94	34	100	44	25	6
Borders	3	100	100	100	100	100	100	100	0
Forth Valley	9	100	89	89	56	100	56	33	0
Argyll & Clyde	0	0	0	0	0	0	0	0	0
Greater Glasgow	8	100	88	88	38	100	38	38	13
Lanarkshire	6	100	67	67	17	83	50	17	0
Ayrshire & Arran	3	100	100	100	67	100	67	33	33
Dumfries & Galloway	0	0	0	0	0	0	0	0	0
Orkney	0	0	0	0	0	0	0	0	0
Shetland	0	0	0	0	0	0	0	0	0
Western Isles	1	100	100	100	0	100	0	0	0
Total	74	100	92	92	39	99	51	32	7

Table 4c Number of records in 'Transfer Out' sample with correctly imputed geographic information

		Number of records with correct geographic indicator							
Age Group	Number of records in sample	Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	5	0	0	0	0	0	0	0	0
5-9	4	4	4	4	3	4	4	4	1
10-14	0	0	0	0	0	0	0	0	0
15-19	5	5	4	4	1	5	3	1	0
20-24	20	20	20	20	9	20	10	7	0
25-29	11	11	11	11	4	11	5	4	0
30-34	7	7	7	7	3	7	4	1	1
35-39	11	11	9	9	5	11	6	4	1
40-44	2	2	2	2	0	2	0	0	0
45-49	1	1	1	1	1	1	1	0	0
50-54	2	2	2	2	0	2	0	0	0
55-59	1	1	1	1	0	1	1	0	0
60-64	1	1	1	1	0	1	1	0	0
65-69	1	1	1	1	1	1	1	1	1
70-74	0	0	0	0	0	0	0	0	0
75+	3	3	3	3	1	3	1	1	0
Total	74	69	66	66	28	69	37	23	4

Table 4d Percentage of records in 'Transfer Out' sample with correctly imputed geographic information

		Number of records with correct geographic indicator							
Age Group	Number of records in sample	Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	5	0	0	0	0	0	0	0	0
5-9	4	100	100	100	75	100	100	100	25
10-14	0	0	0	0	0	0	0	0	0
15-19	5	100	80	80	20	100	60	20	0
20-24	20	100	100	100	45	100	50	35	0
25-29	11	100	100	100	36	100	45	36	0
30-34	7	100	100	100	43	100	57	14	14
35-39	11	100	82	82	45	100	55	36	9
40-44	2	100	100	100	0	100	0	0	0
45-49	1	100	100	100	100	100	100	0	0
50-54	2	100	100	100	0	100	0	0	0
55-59	1	100	100	100	0	100	100	0	0
60-64	1	100	100	100	0	100	100	0	0
65-69	1	100	100	100	100	100	100	100	100
70-74	0	0	0	0	0	0	0	0	0
75+	3	100	100	100	33	100	33	33	0
Total	74	93	89	89	38	93	50	31	5

Table 5a Number of records in 'Transfer In' sample with correctly imputed geographic information

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	7	7	7	7	3	7	6	5	0
Grampian	14	14	14	14	1	14	7	3	0
Tayside	9	9	8	9	6	9	7	6	2
Fife	3	3	3	3	0	3	3	1	0
Lothian	63	63	63	63	17	63	35	22	13
Borders	2	2	2	1	0	2	1	1	0
Forth Valley	5	5	5	5	2	5	5	5	0
Argyll & Clyde	10	10	8	8	3	10	9	5	2
Greater Glasgow	28	28	25	25	14	28	18	14	2
Lanarkshire	11	9	7	6	1	10	6	4	0
Ayrshire & Arran	6	6	6	6	1	6	3	2	1
Dumfries & Galloway	0	0	0	0	0	0	0	0	0
Orkney	0	0	0	0	0	0	0	0	0
Shetland	0	0	0	0	0	0	0	0	0
Western Isles	0	0	0	0	0	0	0	0	0
Total	158	156	148	147	48	157	100	68	20

Table 5b Percentage of records in 'Transfer In' sample with correctly imputed geographic information

Health Board	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	7	100	100	100	43	100	86	71	0
Grampian	14	100	100	100	7	100	50	21	0
Tayside	9	100	89	100	67	100	78	67	22
Fife	3	100	100	100	0	100	100	33	0
Lothian	63	100	100	100	27	100	56	35	21
Borders	2	100	100	50	0	100	50	50	0
Forth Valley	5	100	100	100	40	100	100	100	0
Argyll & Clyde	10	100	80	80	30	100	90	50	20
Greater Glasgow	28	100	89	89	50	100	64	50	7
Lanarkshire	11	82	64	55	9	91	55	36	0
Ayrshire & Arran	6	100	100	100	17	100	50	33	17
Dumfries & Galloway	0	0	0	0	0	0	0	0	0
Orkney	0	0	0	0	0	0	0	0	0
Shetland	0	0	0	0	0	0	0	0	0
Western Isles	0	0	0	0	0	0	0	0	0
Total	158	99	94	93	30	99	63	43	13

Table 5c Number of records in 'Transfer In' sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	12	12	11	10	4	12	8	6	2
5-9	8	8	6	6	4	8	5	5	0
10-14	2	2	2	2	1	2	1	1	1
15-19	25	25	25	25	9	25	16	12	7
20-24	41	41	40	39	9	41	19	15	3
25-29	18	16	16	16	7	18	13	11	2
30-34	17	17	16	16	2	16	6	3	0
35-39	11	11	10	10	2	11	9	4	2
40-44	4	4	4	4	2	4	4	1	1
45-49	5	5	5	5	2	5	5	3	1
50-54	4	4	4	4	2	4	3	3	1
55-59	3	3	2	2	0	3	3	0	0
60-64	3	3	3	3	2	3	3	2	0
65-69	3	3	2	3	1	3	3	1	0
70-74	2	2	2	2	1	2	2	1	0
75+	0	0	0	0	0	0	0	0	0
Total	158	156	148	147	48	157	100	68	20

Table 5d Percentage of records in 'Transfer In' sample with correctly imputed geographic information

Age Group	Number of records in sample	Percentage of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	12	100	92	83	33	100	67	50	17
5-9	8	100	75	75	50	100	63	63	0
10-14	2	100	100	100	50	100	50	50	50
15-19	25	100	100	100	36	100	64	48	28
20-24	41	100	98	95	22	100	46	37	7
25-29	18	89	89	89	39	100	72	61	11
30-34	17	100	94	94	12	94	35	18	0
35-39	11	100	91	91	18	100	82	36	18
40-44	4	100	100	100	50	100	100	25	25
45-49	5	100	100	100	40	100	100	60	20
50-54	4	100	100	100	50	100	75	75	25
55-59	3	100	67	67	0	100	100	0	0
60-64	3	100	100	100	67	100	100	67	0
65-69	3	100	67	100	33	100	100	33	0
70-74	2	100	100	100	50	100	100	50	0
75+	0	0	0	0	0	0	0	0	0
Total	158	99	94	93	30	99	63	43	13

Table 6a **Number of records in currently registered sample with correctly imputed geographic information**

Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	489	489	489	475	210	483	284	226	17
Grampian	228	228	218	214	61	227	150	83	4
Tayside	719	719	667	699	141	708	457	195	7
Fife	489	489	489	476	135	483	362	189	8
Lothian	2,088	2,088	2,061	2,061	544	2,078	1,225	695	52
Borders	253	253	253	248	129	251	197	138	10
Forth Valley	548	548	538	538	147	543	363	220	9
Argyll & Clyde	266	266	254	262	72	262	174	94	4
Greater Glasgow	1,599	1,599	1,424	1,467	279	1,599	745	316	20
Lanarkshire	1,273	1,273	1,229	1,196	234	1,212	910	388	15
Ayrshire & Arran	712	712	675	664	201	711	465	251	10
Dumfries & Galloway	71	71	71	66	20	71	55	36	1
Orkney	27	27	27	27	13	27	20	20	2
Shetland	16	16	16	16	5	16	10	10	1
Western Isles	128	128	128	128	40	128	85	81	3
Total	8,906	8,906	8,539	8,537	2,231	8,799	5,502	2,942	163

Table 6b **Percentage of records in currently registered sample with correctly imputed geographic information**

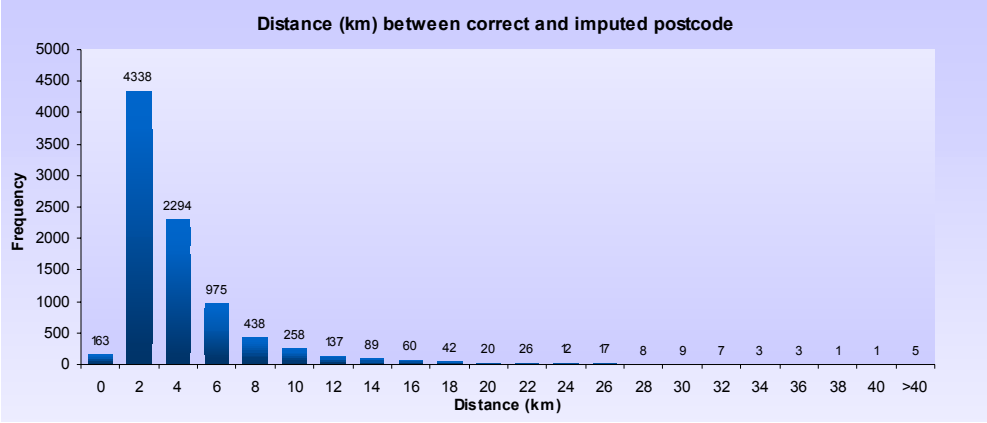
Health Board	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
Highland	489	100	100	97	43	99	58	46	3
Grampian	228	100	96	94	27	100	66	36	2
Tayside	719	100	93	97	20	98	64	27	1
Fife	489	100	100	97	28	99	74	39	2
Lothian	2,088	100	99	99	26	100	59	33	2
Borders	253	100	100	98	51	99	78	55	4
Forth Valley	548	100	98	98	27	99	66	40	2
Argyll & Clyde	266	100	95	98	27	98	65	35	2
Greater Glasgow	1,599	100	89	92	17	100	47	20	1
Lanarkshire	1,273	100	97	94	18	95	71	30	1
Ayrshire & Arran	712	100	95	93	28	100	65	35	1
Dumfries & Galloway	71	100	100	93	28	100	77	51	1
Orkney	27	100	100	100	48	100	74	74	7
Shetland	16	100	100	100	31	100	63	63	6
Western Isles	128	100	100	100	31	100	66	63	2
Total	8,906	100	96	96	25	99	62	33	2

Table 6c Number of records in currently registered sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	661	661	633	629	171	652	418	215	6
5-9	575	575	553	549	149	568	359	194	8
10-14	544	544	518	517	136	534	348	184	11
15-19	531	531	512	513	123	525	334	169	20
20-24	712	712	689	695	157	705	394	195	23
25-29	831	831	804	801	187	820	482	259	17
30-34	836	836	809	809	218	828	532	277	10
35-39	780	780	747	746	179	773	482	248	8
40-44	625	625	592	595	158	618	370	203	6
45-49	485	485	458	460	120	481	284	156	4
50-54	420	420	401	402	111	413	261	158	7
55-59	366	366	345	346	85	358	233	117	5
60-64	366	366	351	349	98	359	234	121	6
65-69	262	262	253	251	78	260	151	97	3
70-74	269	269	258	258	80	268	183	105	6
75+	643	643	616	617	181	637	437	244	23
Total	8,906	8,906	8,539	8,537	2,231	8,799	5,502	2,942	163

Table 6d Percentage of records in currently registered sample with correctly imputed geographic information

Age Group	Number of records in sample	Number of records with correct geographic indicator							
		Health Board	Council	District	Ward	Postcode Area	Postal District	Postcode Sector	Postcode Unit
0-4	661	100	96	95	26	99	63	33	1
5-9	575	100	96	95	26	99	62	34	1
10-14	544	100	95	95	25	98	64	34	2
15-19	531	100	96	97	23	99	63	32	4
20-24	712	100	97	98	22	99	55	27	3
25-29	831	100	97	96	23	99	58	31	2
30-34	836	100	97	97	26	99	64	33	1
35-39	780	100	96	96	23	99	62	32	1
40-44	625	100	95	95	25	99	59	32	1
45-49	485	100	94	95	25	99	59	32	1
50-54	420	100	95	96	26	98	62	38	2
55-59	366	100	94	95	23	98	64	32	1
60-64	366	100	96	95	27	98	64	33	2
65-69	262	100	97	96	30	99	58	37	1
70-74	269	100	96	96	30	100	68	39	2
75+	643	100	96	96	28	99	68	38	4
Total	8,906	100	96	96	25	99	62	33	2



Appendix E

Table 1a – The absolute difference between the controlled and uncontrolled health board flows, for moves between health boards in Scotland, and for moves to and from Scotland.

Table 1a
Impact of controlling CHI migration data to NHSCR Health Board flows (controlled - original)

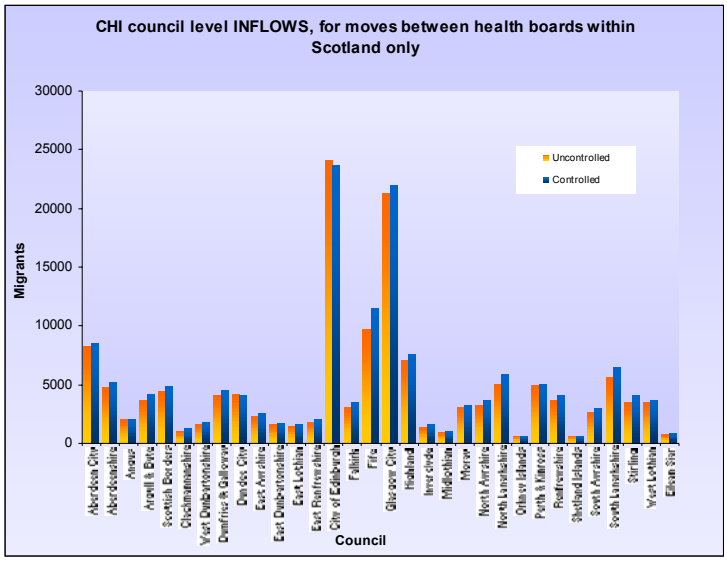
Area of Destination \ Area of Origin	Highland	Grampian	Tayside	Fife	Lothian	Borders	Forth Valley	Argyll & Clyde	Greater Glasgow	Lanarkshire	Ayrshire & Arran	Dumfries & Galloway	Orkney	Shetland	Western Isles	Scotland Total	Overseas
Highland	0	99	67	12	65	10	10	20	75	18	23	3	2	2	7	**413**	115
Grampian	57	0	126	28	71	18	42	14	92	9	19	6	12	24	-1	**517**	404
Tayside	22	113	0	10	56	15	-29	42	43	25	7	24	5	-1	1	**333**	-352
Fife	17	66	81	0	135	14	60	15	49	23	14	9	-2	4	3	**488**	1229
Lothian	-40	65	48	74	0	68	33	9	112	61	45	-10	0	-16	-2	**447**	-341
Borders	7	15	15	25	70	0	2	9	9	19	1	0	-1	3	-5	**169**	298
Forth Valley	23	32	27	10	119	-5	0	49	94	-12	25	-8	-1	3	1	**357**	838
Argyll & Clyde	38	56	49	26	84	7	63	0	344	46	77	42	3	0	10	**845**	435
Greater Glasgow	45	113	109	62	179	22	98	284	0	263	164	43	-3	7	2	**1388**	-282
Lanarkshire	23	47	38	23	102	44	41	72	316	0	53	16	0	2	7	**784**	718
Ayrshire & Arran	11	7	20	37	43	-3	8	78	141	57	0	11	0	2	6	**418**	683
Dumfries & Galloway	7	21	17	1	23	-4	15	25	52	-14	18	0	-1	1	0	**161**	248
Orkney	3	8	3	-2	6	-5	0	3	3	0	0	2	0	0	0	**21**	44
Shetland	-1	11	2	2	5	-1	3	0	6	-4	-1	0	-2	0	0	**20**	25
Western Isles	2	9	3	1	7	1	-1	0	9	18	6	1	-4	0	0	**52**	66
Scotland Total	**214**	**662**	**605**	**309**	**965**	**181**	**345**	**620**	**1345**	**509**	**451**	**139**	**8**	**31**	**29**	**6413**	**4128**
Overseas	1191	4980	2394	1775	5934	604	1443	1668	3646	1478	1121	652	93	166	70	**27215**	

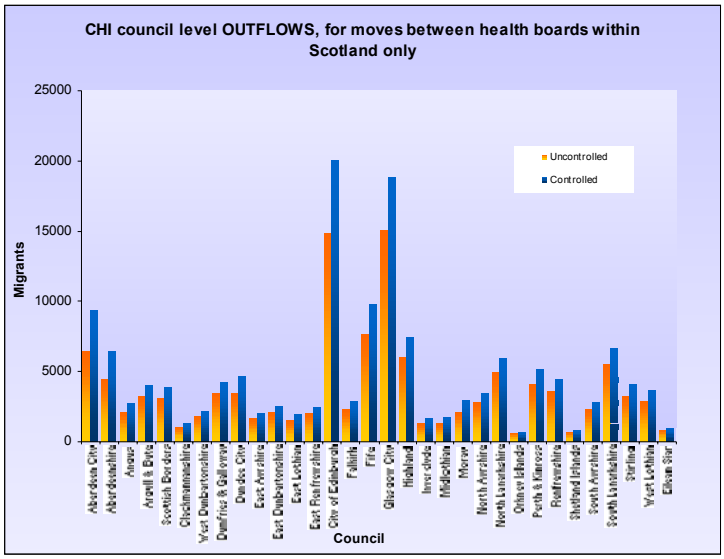
Table 1b – The percentage difference between the controlled and uncontrolled health board flows, for moves between health boards in Scotland, and for moves to and from Scotland.

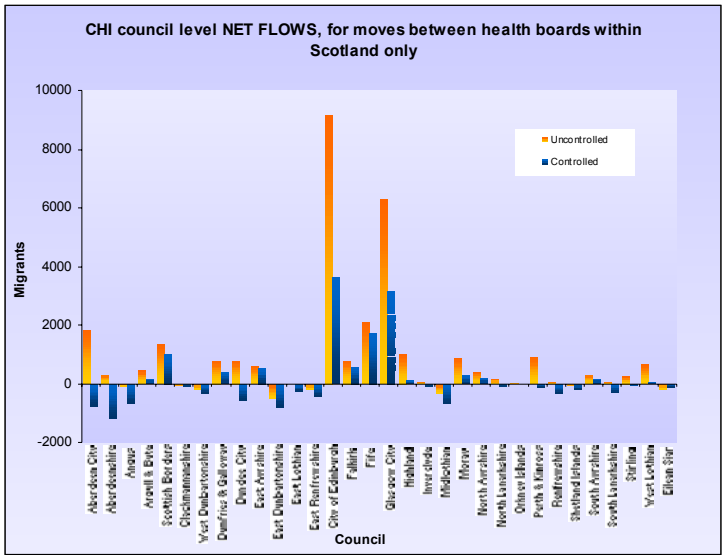
Table 1b
Impact of controlling CHI migration data to NHSCR Health Board flows, (% change)

Area of Destination \ Area of Origin	Highland	Grampian	Tayside	Fife	Lothian	Borders	Forth Valley	Argyll & Clyde	Greater Glasgow	Lanarkshire	Ayrshire & Arran	Dumfries & Galloway	Orkney	Shetland	Western Isles	Scotland Total	Overseas
Highland		11%	21%	7%	12%	15%	6%	6%	16%	10%	12%	5%	2%	4%	4%	**11%**	3%
Grampian	5%		11%	6%	7%	13%	13%	4%	15%	3%	8%	6%	10%	14%	-1%	**8%**	4%
Tayside	5%	12%		1%	5%	13%	-5%	13%	7%	9%	3%	23%	24%	-4%	6%	**5%**	-7%
Fife	9%	14%	6%		6%	15%	14%	6%	10%	9%	9%	12%	-9%	24%	19%	**8%**	31%
Lothian	-5%	4%	4%	5%		7%	3%	2%	7%	7%	9%	-2%	0%	-15%	-4%	**4%**	-2%
Borders	9%	14%	14%	28%	4%		2%	10%	5%	22%	2%	0%	-50%	75%	-38%	**7%**	16%
Forth Valley	13%	9%	6%	3%	11%	-6%		13%	10%	-2%	12%	-8%	-9%	19%	5%	**7%**	30%
Argyll & Clyde	18%	23%	20%	17%	19%	12%	26%		10%	11%	11%	42%	38%	0%	21%	**14%**	10%
Greater Glasgow	7%	12%	15%	15%	13%	13%	12%	11%		10%	12%	12%	-10%	16%	1%	**11%**	-2%
Lanarkshire	16%	21%	23%	13%	14%	39%	12%	14%	9%		15%	15%	0%	67%	47%	**12%**	24%
Ayrshire & Arran	8%	3%	12%	26%	14%	-5%	4%	7%	8%	12%		4%	0%	33%	100%	**8%**	22%
Dumfries & Galloway	9%	26%	21%	2%	8%	-5%	22%	18%	19%	-11%	5%		-11%	6%	0%	**10%**	10%
Orkney	3%	11%	12%	-29%	12%	-56%	0%	50%	14%	0%	0%	33%		0%	0%	**6%**	17%
Shetland	-3%	12%	10%	11%	8%	-10%	38%	0%	43%	-44%	-10%	0%	-13%		-	**6%**	8%
Western Isles	2%	13%	12%	7%	18%	20%	-10%	0%	9%	82%	29%	17%	-29%	0%		**10%**	28%
Scotland Total	**5%**	**10%**	**10%**	**6%**	**9%**	**9%**	**8%**	**9%**	**10%**	**8%**	**10%**	**8%**	**2%**	**6%**	**4%**	**9%**	**6%**
Overseas	65%	76%	65%	67%	61%	54%	64%	53%	54%	53%	47%	41%	62%	86%	32%	**61%**	

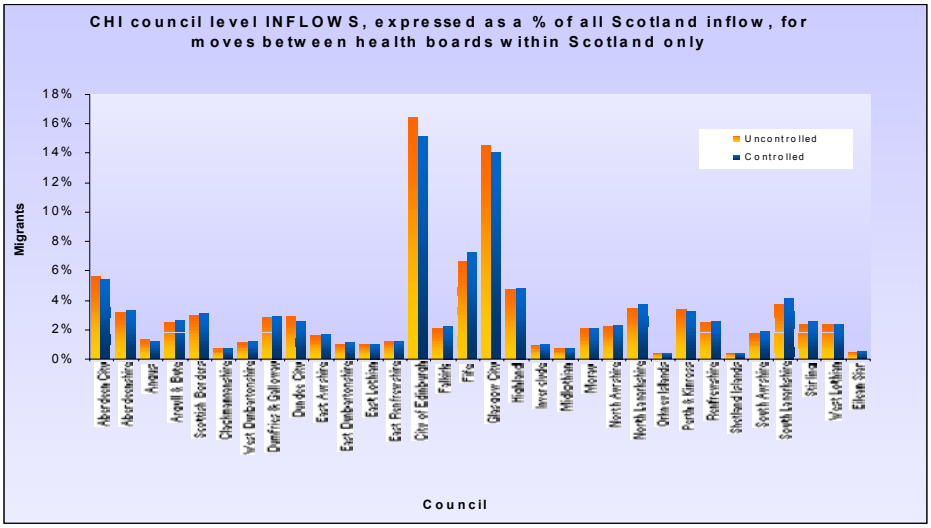
Council Flows – In, out and net flows for moves between health boards within Scotland only, shown by Council.

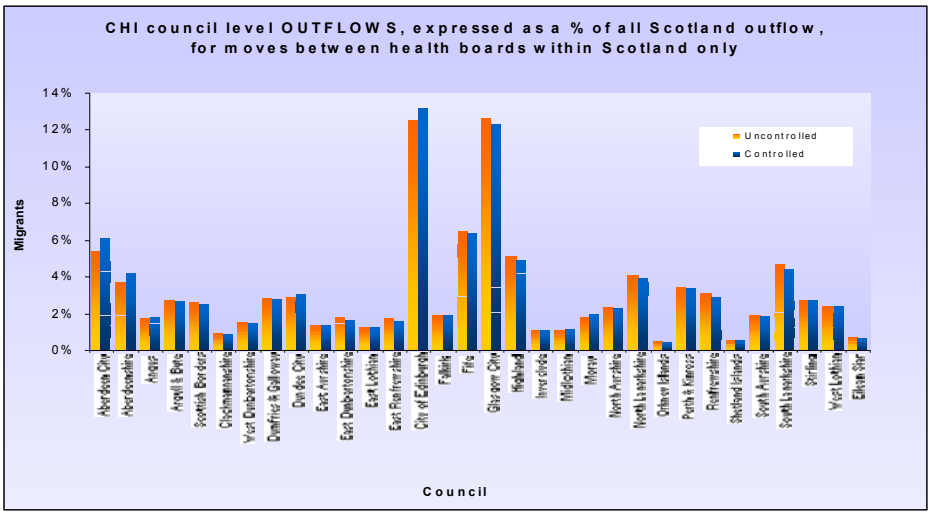




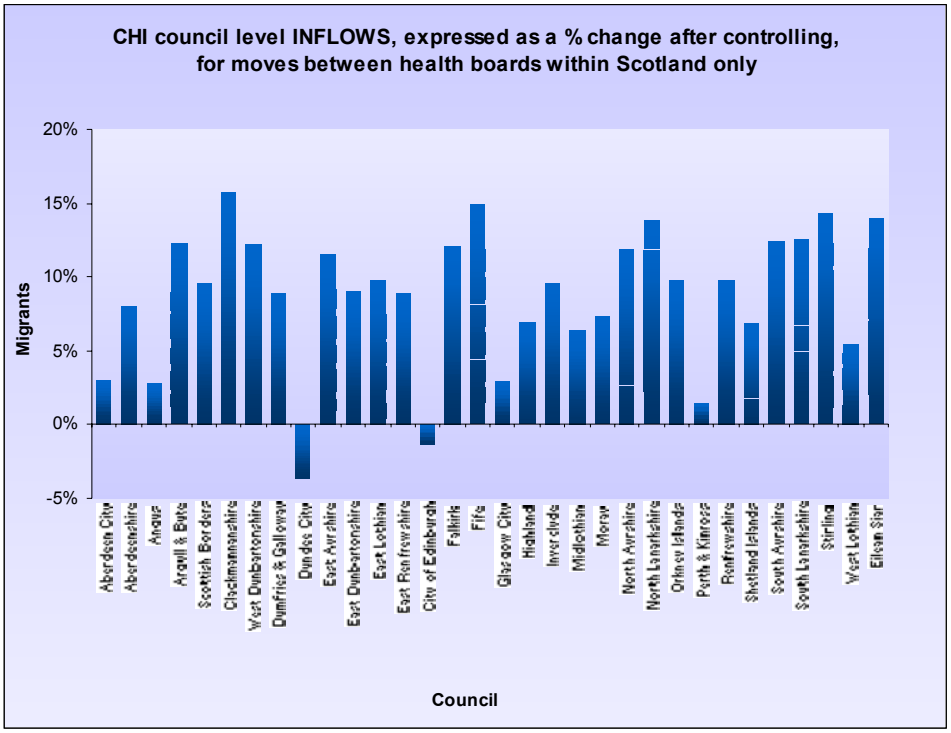


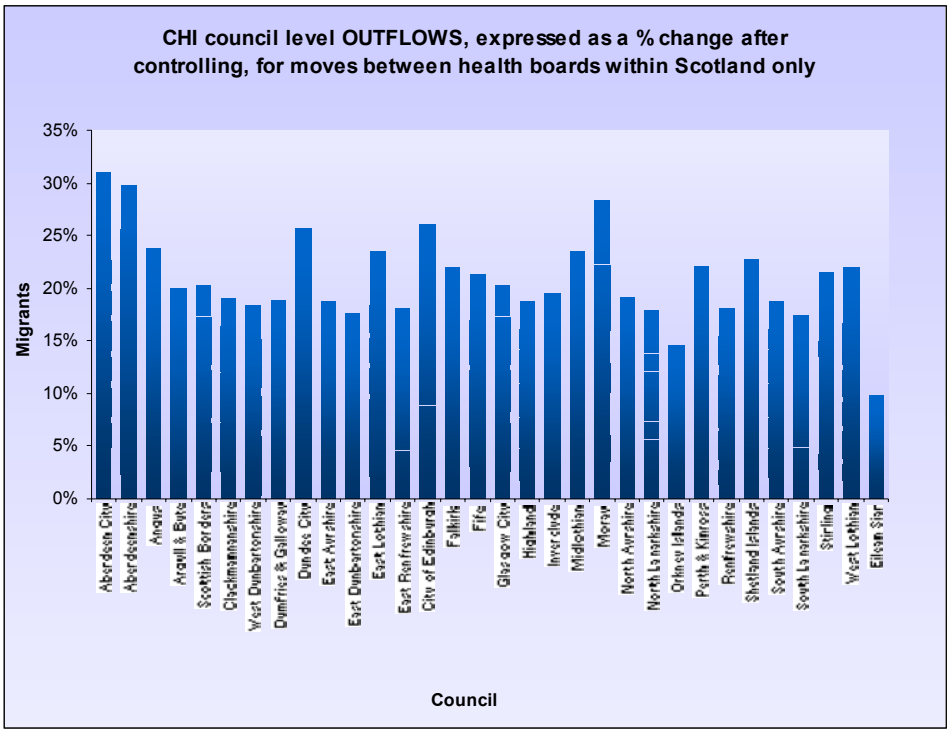
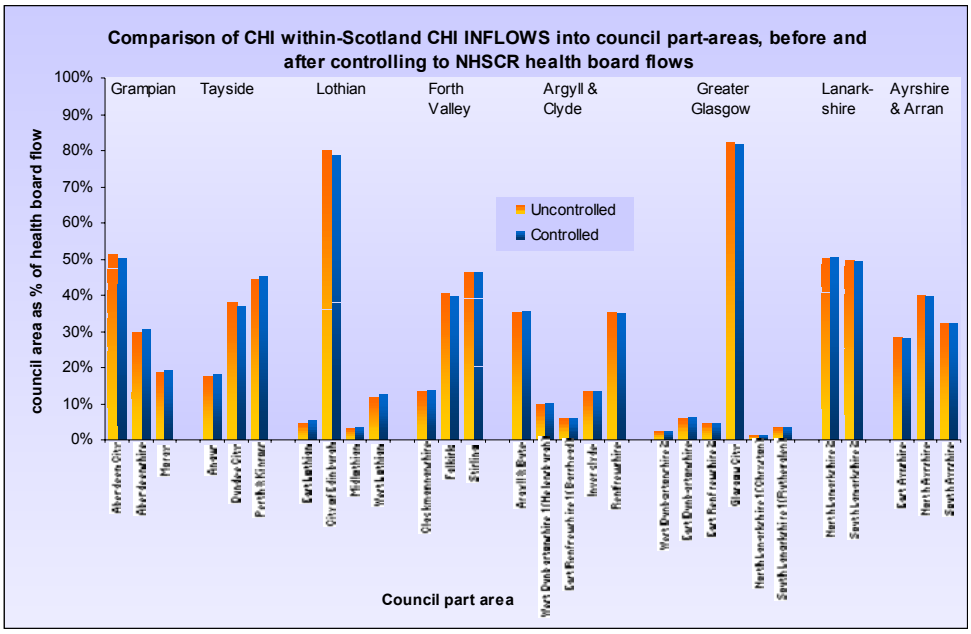
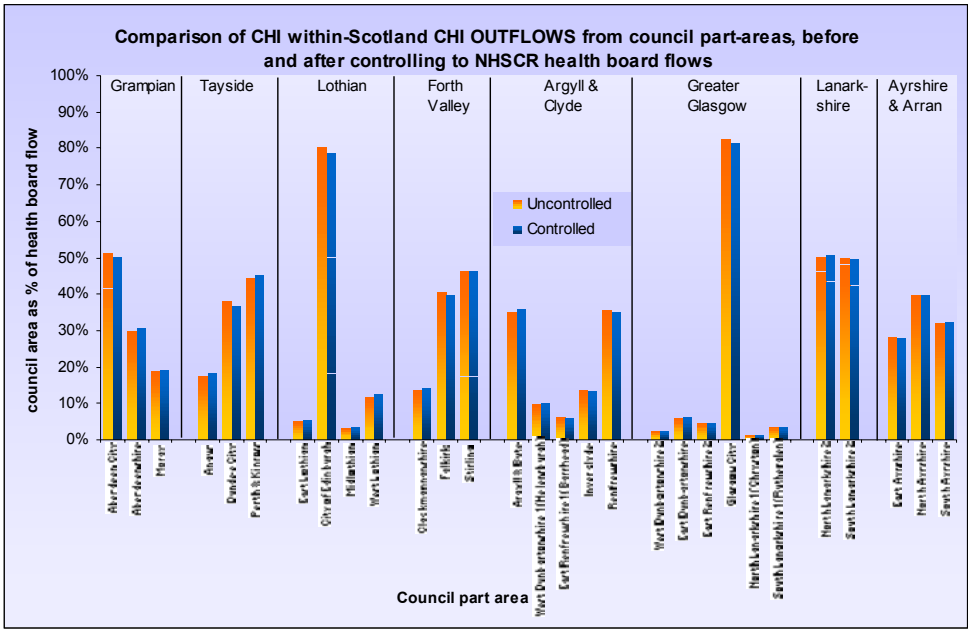
Council Proportions – Council level in and out flows, expressed as a %'age of all Scotland flow, for moves between health boards within Scotland only.





% Controlling – The percentage difference between the controlled and uncontrolled council flows, for moves between health boards in Scotland only.





PA's in HB – Flows to/from council part-areas within health boards, expressed as a percentage of the flow into each health board, before and after controlling.