This is one of a series of reports published to support the release of results from the 2011 Census. This series of methods and quality reports provides information on the different methods used to collect, process, clean, adjust and protect the census results. The series also reports on the quality assurance of the results and provides quality indicators.

Terms used in the series are explained in the 2011 Census glossary.

Contents

1 Summary ..................................................................................................................................................2
2 Introduction .............................................................................................................................................3
3 Provisional dataset ...................................................................................................................................4
4 Summary of results ..................................................................................................................................5
5 Accounting for attrition in the LS sample .............................................................................................17
Annex A: Composition of the LS sample after each census link ..............................................................22
1 Summary

The ONS Longitudinal Study (LS) links census and life event information for 1 per cent of the population for England and Wales. Information from the 2011 Census linkage has been used to benchmark the level of adjustment required for overcount in the 2011 Census and used in making a national adjustment to the 2011 Census estimates.

This report provides an assessment of the linkage of 2011 Census records to the existing LS study records, using a provisional dataset created for this exercise. The report compares linkage rates at each census from 1981 to 2011, and quantifies and examines the demographic characteristics of those previously in the LS, without a valid exit event, not found at 2011 and seeks to explain the loss to follow up. It also contains data on individuals enumerated more than once at the 2011 Census.

Tracing of LS records in 2011 was extremely successful with 99.1 per cent of records identified from the 2011 Census traced to a NHS patient record. The proportion of LS members enumerated at the 2001 Census and 'not found' in 2011 was 17.9 per cent. A number of factors affecting non-linkage in the LS in 2011 have been identified; including being a young male, male or female immigrant, new born or an immigrant child under 10 years of age.

Types of multiple enumeration found included students enumerated at both their term-time and home address, individuals enumerated more than once on the same form, enumerated on two forms with the same postcode, enumerated at a second address and/or an address of usual residence one year ago, or children enumerated with both parents at different addresses.

The LS estimated rate of multiple enumeration was used as a benchmark in estimating and adjusting for overcount in the 2011 Census estimates. LS linkage rates were also used as a measure of coverage for the 2011 Census. A set of expected non-response rates by age and sex group were calculated from the LS data and used as evidence as to whether further adjustment of the 2011 Census estimates was required.
2 Introduction

The ONS Longitudinal Study (LS) links census and life event information for one per cent of the population of England and Wales.

The original sample was selected from the 1971 Census, and incorporated data on individuals born on one of four selected dates of birth. The sample has been updated at each successive census by taking people with the same four dates of birth in each year and linking them to the existing data.

Life event information has been added to the LS since Census Day in 1971, including birth and immigration (entry events) and death and emigration (exit events) of people with the four dates of birth.

Linkage of data over time is achieved by tracing against NHS patient records. For the 2001 and 2011 Censuses tracing was a combination of an automated process with manual follow up. For censuses prior to 2001, tracing was a manual process. This work is carried out for ONS by the Medical Research Information Service team, part of the Health & Social Care Information Centre (HSCIC) based in Southport. This team were previously part of ONS, and were transferred to HSCIC on 1 April 2008 as a consequence of the 2007 Statistics and Registration Service Act.

Data on the extent to which individuals are enumerated, traced and linked correctly is vital for accurate use of the LS. This report provides an initial assessment of the 2011 linkage process. It compares tracing rates at each census from 1981 to 2011, quantifies and examines the demographic characteristics of those previously in the LS without a valid exit event not found at 2011, and seeks to explain the loss to follow up. It also contains data on individuals enumerated more than once at the 2011 Census.

Tracing of records from the 2011 Census is now complete. This was achieved using an initial tracing extract of census data created prior to the removal of multiple responses. Datasets were provided to HSCIC in batches of delivery groups. The returned traced datasets contained indicators of whether the individual was traced and whether they had been previously identified as an LS member. Where this was the case, a unique identifier was returned allowing these individuals to be linked to their previous LS history. People who had been enumerated more than once in the 2011 Census were identified by cases where more than one record was returned with the same unique identifier.

1 Data from 1971 to 2010 were traced against the NHS Central Register. Data from 2011 were traced against the new MIDAS system which takes patient data from the Personal Demographics Service. MIDAS will be used for all future tracing activity.
3 Provisional dataset

A provisional dataset was created for this exercise comprising existing LS data, the data returned from the tracing exercise and additional data to include 2010 and early 2011 life events:

- An early extract of LS births for 2010 was supplied to HSCIC prior to tracing of the 2011 Census data. These individuals were flagged on the HSCIC system and their identifiers included in the provisional dataset.

- LS births for the first quarter of 2011 could not be processed before tracing began. Babies born during this period who were enumerated at census will have been traced at HSCIC, but will have been incorrectly flagged as having entered the LS at the 2011 Census. Once processing of 2011 LS births is complete, these will be reclassified as having entered the LS at birth. There are estimated to be 2,097 such cases. Examination of the initial 2011 Census extracts shows 1,784 traced cases born in 2011.

- Similarly, LS immigrants since September 2010 could not be processed before tracing began. Immigrants registering with the NHS during this period who were enumerated at census will also have been traced, and also incorrectly flagged as having entered the LS at the 2011 Census. These will also be reclassified once processing of LS immigrants is complete. It is impossible to identify these individuals in the census extract directly but examination of the ‘year and month of entry to UK’ question from the 2011 Census shows approximately 3,000 individuals arriving in the year prior to Census Day and traced as new entrants.

- An extract showing the status of all flagged LS members on Census Day was provided by HSCIC. This allowed for the completion of the exit events in the provisional dataset.
4 Summary of results

4.1. Tracing at HSCIC

Table 1: LS tracing rates

<table>
<thead>
<tr>
<th>Census</th>
<th>Total enumerated</th>
<th>Per cent traced</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>585,034</td>
<td>99.1</td>
</tr>
<tr>
<td>2001</td>
<td>539,996</td>
<td>99.4</td>
</tr>
<tr>
<td>1991</td>
<td>543,875</td>
<td>98.4</td>
</tr>
<tr>
<td>1981</td>
<td>536,137</td>
<td>98.8</td>
</tr>
<tr>
<td>1971</td>
<td>529,901</td>
<td>96.8</td>
</tr>
</tbody>
</table>

Table 1 shows the tracing rates achieved at each census. Currently the rate at 2011 (99.1 per cent) is slightly lower than that achieved at 2001 (99.4 per cent), mainly due to an increase in the number of recent migrants enumerated at census who have not yet registered with the NHS.

The increase in tracing rates since 1991 has been achieved by the combination of use of automated tracing methodologies combined with clerical matching and the availability of more detailed information, including access to census images to aid in query resolution.

4.2. Accounting for the LS sample at each census

On completion of the 2011 Census link exercise the LS contained a total of 1,099,507 members enumerated at one or more censuses, or entering the sample as a new birth or an immigrant. Of these 585,034 were found at the 2011 Census.

Annex A includes a set of tables detailing the history and status of LS members at each census from 2011 back to 1981. These tables are summarised in Table 2.

The major change in the composition of the sample in the decade prior to 2011 was the increase in the number of immigrants. Table 2 shows that almost 72,000 immigrants joined the study between 2001 and 2011 and survived to the 2011 Census (i.e. there was no subsequent death or emigration record). This number is expected to rise to nearer 80,000 once all the event data up to Census Day 2011 has been processed and records reclassified. The comparable numbers of surviving immigrants for the previous decades are 44,000 between 1991 and 2001, and 22,000 between 1981 and 1991.
### Table 2: Number and percentage of traced LS members not accounted for at each Census

<table>
<thead>
<tr>
<th>Found at previous census</th>
<th>Expected at census</th>
<th>Not found at census</th>
<th>Per cent not found</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>482,237</td>
<td>60,521</td>
<td>12.6</td>
</tr>
<tr>
<td>2001</td>
<td>477,210</td>
<td>57,993</td>
<td>12.2</td>
</tr>
<tr>
<td>1991</td>
<td>470,041</td>
<td>47,834</td>
<td>10.2</td>
</tr>
<tr>
<td>1981</td>
<td>456,857</td>
<td>39,817</td>
<td>8.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>69,811</td>
<td>10,411</td>
<td>14.9</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>69,780</td>
<td>9,944</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>71,990</td>
<td>6,283</td>
<td>8.7</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>69,995</td>
<td>5,070</td>
<td>7.2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>71,841</td>
<td>40,589</td>
<td>56.5</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>43,834</td>
<td>28,886</td>
<td>65.9</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>21,541</td>
<td>13,306</td>
<td>61.8</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>28,308</td>
<td>13,451</td>
<td>47.5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>1,010</td>
<td>386</td>
<td>38.2</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>1,106</td>
<td>387</td>
<td>35.0</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>1,471</td>
<td>421</td>
<td>28.6</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>624,899</td>
<td>111,907</td>
<td>17.9</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>591,930</td>
<td>97,210</td>
<td>16.4</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>565,043</td>
<td>67,844</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>555,160</td>
<td>58,338</td>
<td>10.5</td>
<td></td>
</tr>
</tbody>
</table>
4.3. Linkage at each census, 1981 to 2011

Table 2 shows an overall increase in the proportion not found from 16.4 per cent in 2001 to 17.9 per cent in 2011.

There were slight increases in the ‘not found’ proportions for members enumerated at the previous census (12.2 per cent in 2001 to 12.6 per cent in 2011) and for new births (14.3 to 14.9 per cent). The proportion of immigrants not found decreased from 65.9 per cent in 2001 and 61.8 per cent in 1991 to 56.5 per cent in 2011. These 2011 rates may change slightly once all 2011 life events have been processed as some cases will be reclassified.

Table 3 cross-tabulates the figures from the ‘Not found at census’ column of Table 2, breaking down the total not found at each census by how those LS members were first identified in the previous decade. This shows that the increase in immigration in the last decade had a significant impact on the overall linkage failure rate, with this group accounting for 36.3 per cent of those not found in 2011 compared to 29.7 per cent in 2001. The proportion of re-entrants not found has increased from 35.0 to 38.2 per cent (Table 2), but this group has little effect on the overall not found proportion (0.3 per cent in 2011 shown in Table 3).

Table 3: Number and percentage of traced LS members not found at census by where first identified in the previous decade

<table>
<thead>
<tr>
<th></th>
<th>Not found 2011</th>
<th>Not found 2001</th>
<th>Not found 1991</th>
<th>Not found 1981</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Found at previous census</td>
<td>60,521</td>
<td>54.1</td>
<td>57,993</td>
<td>59.7</td>
</tr>
<tr>
<td>Births in previous decade</td>
<td>10,411</td>
<td>9.3</td>
<td>9,944</td>
<td>10.2</td>
</tr>
<tr>
<td>Immigrants in previous decade</td>
<td>40,589</td>
<td>36.3</td>
<td>28,886</td>
<td>29.7</td>
</tr>
<tr>
<td>Re-entrants in previous decade</td>
<td>386</td>
<td>0.3</td>
<td>387</td>
<td>0.4</td>
</tr>
<tr>
<td>Total</td>
<td>111,907</td>
<td>100</td>
<td>97,210</td>
<td>100</td>
</tr>
</tbody>
</table>
4.4. Demographic profiles of LS members found/not found at census

Figures 1 to 4 show the age and sex profiles of LS members found and not found at each census from 1981. As previously noted, the number not found has increased at each census since 1981. Note that the scales for the ‘Not Found’ and ‘Found’ figures are different.

At each census males were less likely to be found than females. For both males and females, those aged 20 to 44 were less likely to be found, with those aged 25 to 29 the largest group.

While immigrants in the previous decade accounted for 36.3 per cent of those not found across all age groups, they had a much greater impact on young adults, with 55.2 per cent of male and 65.7 per cent of female immigrants aged 25 to 29 not found at 2011. These proportions increased from 42.3 per cent for males and 57.2 per cent for females in 2001, and 23.1 per cent and 38.6 per cent respectively in 1991.

In both 2001 and 2011 there was an increase in those aged under ten not found at census. In 2011, 16.6 per cent of all under tens were not found (14.9 per cent of new births and 43.6 per cent of immigrants), while in 2001 16.1 per cent were not found (14.3 per cent of new births and 52.5 per cent of immigrants).

Figure 1: Traced LS members found and not found at the 2011 Census by age and sex
Figure 2: Traced LS members found and not found at the 2001 Census by age and sex

Figure 3: Traced LS members found and not found at the 1991 Census by age and sex
Figure 4: Traced LS members found and not found at the 1981 Census by age and sex

4.5. Multiple enumerations identified at census

The identification of multiple responses, where the same individual is enumerated more than once at census, is crucial for accurate use of LS data. Quantification of the level of overcount identified by the LS has also provided a benchmark for the overcount adjustment made to 2011 Census estimates. This is documented in a 2011 Census: Methods and Quality Report published in July 2012.

In 2011, Census data used in the link process were extracted prior to the removal of multiple responses by Census processing, whereas in 2001 multiple responses were removed prior to the creation of the LS extract. This resulted in higher numbers of multiple responses being identified in the LS in 2011, but allowed comparison between these cases and those identified and removed by Census processing.

Multiple enumerations in the LS were identified where more than one census record was matched to a single NHS record. In total, 12,250 LS members were identified as multiple enumerations matching to 25,299 census records. Multiple enumeration types were derived using: postcode of enumeration; postcode of usual residence one year ago; second address postcode and type; and responses to the term-time address question for students. Table 4 shows the distribution of multiple enumeration types.

---

2 2011 Census: Methods and Quality Report, Overcount Estimation and Adjustment, July 2012
### Table 4: Distribution of multiple enumeration types

<table>
<thead>
<tr>
<th>Type</th>
<th>LS members</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Students enumerated at term-time and home address</strong></td>
<td>4,902</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td>7,348</td>
</tr>
<tr>
<td>Enumerated on same form</td>
<td>928</td>
</tr>
<tr>
<td>Internet and paper return for the same address</td>
<td>687</td>
</tr>
<tr>
<td>Enumerated on different forms for the same address</td>
<td>1,881</td>
</tr>
<tr>
<td>Enumerated with parents at two addresses</td>
<td>840</td>
</tr>
<tr>
<td>Enumerated on two forms where postcode of enumeration on one was postcode of usual residence one year ago on other</td>
<td>373</td>
</tr>
<tr>
<td>Enumerated on two forms where second address postcode and postcode one year ago on one was the same as postcode of enumeration on other</td>
<td>586</td>
</tr>
<tr>
<td>Enumerated on two forms where second address postcode was same as postcode of enumeration on other</td>
<td>538</td>
</tr>
<tr>
<td>Enumerated on communal establishment and household forms</td>
<td>651</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>864</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12,250</strong></td>
</tr>
</tbody>
</table>

Students who lived away from home during term-time were required to be enumerated fully at their term-time address, with limited information entered on their family home form. Thus, while they are included here as multiple LS enumerations, they do not contribute to census over-enumeration.

There were 928 individuals enumerated more than once on the same form (either paper or internet returns). Here individuals either:

- made a mistake in personal details and used the next person’s slot for a correction;
- completed the details roughly at first in one person’s slot and used the next person’s slot for their ‘full’ response; or
- returned the same details in more than one person’s slot.

In 687 cases the LS member was enumerated on both an internet and paper return for the same household using the same questionnaire ID number (QID).
In 1,881 cases the LS member was enumerated on two forms with the same postcode but with different QIDs. These can be further summarised as follows:

- In 1,136 cases the original form posted out to the address had been completed along with a second return.
- In 172 cases both responses were on secondary forms.
- In 327 cases both responses were on original forms where either two forms were delivered to the same address and both completed or, in a small number of cases, the respondent was associated with two properties in the same postcode and returned forms for both.
- The remaining 246 cases were a combination of: household and continuation or individual forms; individual forms completed twice for a person in a communal establishment; or a household and communal return for the same person.

There were 840 cases of children being enumerated with both parents at different addresses.

In 1,497 cases the second enumeration came from a second address and/or an address of usual residence one year ago. This occurred where: a form was returned for the first address and the household then moved home and completed a second form at the second address; young people had left the family home and were enumerated at both the family home and their new address; or individuals had more than one property and completed returns for both. In 118 of these cases the LS member was a student but both addresses were identified as their term-time address.

LS members were enumerated in both a communal establishment and a private household in 651 cases. This can occur where an older person has moved into residential care and is enumerated in both the care and family homes.

The remaining 864 cases could not be allocated a type from the limited census data available at the time, and will be investigated further when the full extract of census data for the LS is available.

4.6. Comparison of multiple enumerations identified by the LS and by Census

Census processing included operations to identify multiple responses from the same individual, to merge information from multiple responses where appropriate, and to remove redundant records. This was largely achieved through the Resolve Multiple Responses (RMR) process, with a number of other specific operations managed through data file amendments.

Table 5 shows the distribution of LS multiple enumerations removed/not removed where both enumerations related to the same address. These are the cases shown as either ‘Enumerated on same form’ or ‘Internet and paper return for the same address’ in Table 4. Of the 1,615 cases identified in the LS 1,258 (78 per cent) were removed by census processes. In all cases that were not removed there were differences in the name, date of birth or gender between the enumerations.
### Table 5: LS members multiply enumerated with same QID removed by RMR

<table>
<thead>
<tr>
<th>Enumeration channel</th>
<th>Person characteristics</th>
<th>Not removed</th>
<th>Removed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enumerated on same form and channel</td>
<td>Name, sex or dob differ</td>
<td>272</td>
<td>448</td>
<td>720</td>
</tr>
<tr>
<td></td>
<td>Name, sex and dob the same</td>
<td>0</td>
<td>208</td>
<td>208</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>272</strong></td>
<td><strong>656</strong></td>
<td><strong>928</strong></td>
</tr>
<tr>
<td>Enumerated on paper and internet forms</td>
<td>Name, sex or dob differ</td>
<td>85</td>
<td>213</td>
<td>298</td>
</tr>
<tr>
<td></td>
<td>Name, sex and dob the same</td>
<td>0</td>
<td>389</td>
<td>389</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>85</strong></td>
<td><strong>602</strong></td>
<td><strong>687</strong></td>
</tr>
</tbody>
</table>

Table 6 shows the distribution of LS multiple enumerations removed/not removed where the enumerations came from different QIDs with the same postcode. The lowest removal rate, 13 per cent, was from cases where two forms were issued to the same household or the individual was associated with two addresses in the same postcode, while 64 per cent of cases where both an original and issued form were completed were removed. Again individuals were more likely to be identified and removed where name, date of birth and gender matched exactly.
Table 6: LS members multiply enumerated with different QID but same postcode removed by RMR

<table>
<thead>
<tr>
<th>Enumeration channel</th>
<th>Person characteristics</th>
<th>Not removed</th>
<th>Removed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Both initial household forms</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Name, sex or dob differ</td>
<td>146</td>
<td>12</td>
<td>158</td>
</tr>
<tr>
<td></td>
<td>Name, sex and dob the same</td>
<td>137</td>
<td>32</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>283</strong></td>
<td><strong>44</strong></td>
<td><strong>327</strong></td>
</tr>
<tr>
<td></td>
<td>One initial and one issued household form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Name, sex or dob differ</td>
<td>279</td>
<td>180</td>
<td>459</td>
</tr>
<tr>
<td></td>
<td>Name, sex and dob the same</td>
<td>128</td>
<td>549</td>
<td>677</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>407</strong></td>
<td><strong>729</strong></td>
<td><strong>1,136</strong></td>
</tr>
<tr>
<td></td>
<td>Both issued household forms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Name, sex or dob differ</td>
<td>56</td>
<td>33</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Name, sex and dob the same</td>
<td>18</td>
<td>65</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>74</strong></td>
<td><strong>98</strong></td>
<td><strong>172</strong></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Name, sex or dob differ</td>
<td>55</td>
<td>50</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Name, sex and dob the same</td>
<td>54</td>
<td>87</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>109</strong></td>
<td><strong>137</strong></td>
<td><strong>246</strong></td>
</tr>
</tbody>
</table>
Table 7 shows the age and gender distribution for all LS members identified as multiple enumerations before and after Census RMR, and Table 8 the same distribution excluding students enumerated at term-time and home addresses.

**Table 7: Multiple enumerations by age and gender**

<table>
<thead>
<tr>
<th>Age group</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>171</td>
<td>165</td>
<td>336</td>
<td>124</td>
<td>113</td>
<td>237</td>
</tr>
<tr>
<td>5-9</td>
<td>260</td>
<td>222</td>
<td>482</td>
<td>230</td>
<td>182</td>
<td>412</td>
</tr>
<tr>
<td>10-14</td>
<td>410</td>
<td>331</td>
<td>741</td>
<td>362</td>
<td>284</td>
<td>646</td>
</tr>
<tr>
<td>15-19</td>
<td>1,234</td>
<td>1,340</td>
<td>2,574</td>
<td>1,165</td>
<td>1,263</td>
<td>2,428</td>
</tr>
<tr>
<td>20-24</td>
<td>1,834</td>
<td>1,863</td>
<td>3,697</td>
<td>1,742</td>
<td>1,771</td>
<td>3,513</td>
</tr>
<tr>
<td>25-29</td>
<td>337</td>
<td>265</td>
<td>602</td>
<td>279</td>
<td>174</td>
<td>453</td>
</tr>
<tr>
<td>30-34</td>
<td>229</td>
<td>179</td>
<td>408</td>
<td>161</td>
<td>97</td>
<td>258</td>
</tr>
<tr>
<td>35-39</td>
<td>204</td>
<td>168</td>
<td>372</td>
<td>129</td>
<td>90</td>
<td>219</td>
</tr>
<tr>
<td>40-44</td>
<td>198</td>
<td>155</td>
<td>353</td>
<td>121</td>
<td>92</td>
<td>213</td>
</tr>
<tr>
<td>45-49</td>
<td>210</td>
<td>171</td>
<td>381</td>
<td>137</td>
<td>89</td>
<td>226</td>
</tr>
<tr>
<td>50-54</td>
<td>191</td>
<td>150</td>
<td>341</td>
<td>129</td>
<td>79</td>
<td>208</td>
</tr>
<tr>
<td>55-59</td>
<td>167</td>
<td>122</td>
<td>289</td>
<td>115</td>
<td>79</td>
<td>194</td>
</tr>
<tr>
<td>60-64</td>
<td>169</td>
<td>150</td>
<td>319</td>
<td>115</td>
<td>93</td>
<td>208</td>
</tr>
<tr>
<td>65-69</td>
<td>118</td>
<td>121</td>
<td>239</td>
<td>83</td>
<td>79</td>
<td>162</td>
</tr>
<tr>
<td>70-74</td>
<td>113</td>
<td>118</td>
<td>231</td>
<td>67</td>
<td>69</td>
<td>136</td>
</tr>
<tr>
<td>75-79</td>
<td>106</td>
<td>124</td>
<td>230</td>
<td>57</td>
<td>59</td>
<td>116</td>
</tr>
<tr>
<td>80-84</td>
<td>120</td>
<td>158</td>
<td>278</td>
<td>50</td>
<td>63</td>
<td>113</td>
</tr>
<tr>
<td>85+</td>
<td>65</td>
<td>200</td>
<td>265</td>
<td>23</td>
<td>85</td>
<td>108</td>
</tr>
<tr>
<td>All ages</td>
<td>6,136</td>
<td>6,002</td>
<td>12,138</td>
<td>5,089</td>
<td>4,761</td>
<td>9,850</td>
</tr>
</tbody>
</table>

* 112 cases with missing values on age or sex.
Table 8: Multiple enumerations by age and gender excluding students at term-time and home address

<table>
<thead>
<tr>
<th>Age group</th>
<th>Identified by LS Male</th>
<th>Identified by LS Female</th>
<th>Identified by LS Total</th>
<th>After RMR Male</th>
<th>After RMR Female</th>
<th>After RMR Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>157</td>
<td>155</td>
<td>312</td>
<td>110</td>
<td>103</td>
<td>213</td>
</tr>
<tr>
<td>5-9</td>
<td>205</td>
<td>178</td>
<td>383</td>
<td>175</td>
<td>138</td>
<td>313</td>
</tr>
<tr>
<td>10-14</td>
<td>272</td>
<td>228</td>
<td>500</td>
<td>224</td>
<td>181</td>
<td>405</td>
</tr>
<tr>
<td>15-19</td>
<td>414</td>
<td>411</td>
<td>825</td>
<td>353</td>
<td>338</td>
<td>691</td>
</tr>
<tr>
<td>20-24</td>
<td>565</td>
<td>479</td>
<td>1,044</td>
<td>481</td>
<td>390</td>
<td>871</td>
</tr>
<tr>
<td>25-29</td>
<td>292</td>
<td>236</td>
<td>528</td>
<td>234</td>
<td>145</td>
<td>379</td>
</tr>
<tr>
<td>30-34</td>
<td>220</td>
<td>175</td>
<td>395</td>
<td>152</td>
<td>93</td>
<td>245</td>
</tr>
<tr>
<td>35-39</td>
<td>203</td>
<td>164</td>
<td>367</td>
<td>128</td>
<td>86</td>
<td>214</td>
</tr>
<tr>
<td>40-44</td>
<td>196</td>
<td>155</td>
<td>351</td>
<td>119</td>
<td>92</td>
<td>211</td>
</tr>
<tr>
<td>45-49</td>
<td>209</td>
<td>169</td>
<td>378</td>
<td>136</td>
<td>87</td>
<td>223</td>
</tr>
<tr>
<td>50-54</td>
<td>189</td>
<td>149</td>
<td>338</td>
<td>127</td>
<td>78</td>
<td>205</td>
</tr>
<tr>
<td>55-59</td>
<td>166</td>
<td>121</td>
<td>287</td>
<td>114</td>
<td>78</td>
<td>192</td>
</tr>
<tr>
<td>60-64</td>
<td>167</td>
<td>150</td>
<td>317</td>
<td>113</td>
<td>93</td>
<td>206</td>
</tr>
<tr>
<td>65-69</td>
<td>116</td>
<td>121</td>
<td>237</td>
<td>81</td>
<td>79</td>
<td>160</td>
</tr>
<tr>
<td>70-74</td>
<td>112</td>
<td>117</td>
<td>229</td>
<td>66</td>
<td>68</td>
<td>134</td>
</tr>
<tr>
<td>75-79</td>
<td>105</td>
<td>124</td>
<td>229</td>
<td>56</td>
<td>59</td>
<td>115</td>
</tr>
<tr>
<td>80-84</td>
<td>119</td>
<td>158</td>
<td>277</td>
<td>49</td>
<td>63</td>
<td>112</td>
</tr>
<tr>
<td>85+</td>
<td>63</td>
<td>199</td>
<td>262</td>
<td>21</td>
<td>84</td>
<td>105</td>
</tr>
<tr>
<td>All ages</td>
<td>3,770</td>
<td>3,489</td>
<td>*7,259</td>
<td>2,739</td>
<td>2,255</td>
<td>4,994</td>
</tr>
</tbody>
</table>

* 79 cases with missing values on age or sex.
5 Accounting for attrition in the LS sample

In addition to census under-enumeration, failure to link LS members can arise from a number of factors.

- Unrecorded migration where the LS member has left England and Wales but their exit has not been captured by NHS systems. Recording of embarkations is dependent on the individual informing their GP prior to leaving, or registering with the NHS in Scotland or Northern Ireland.

- LS members who are enumerated at census but their date of birth is entered or captured in error resulting in their data not being included in the LS extract.

- Individuals entering the LS sample at a previous census or as an immigrant due to erroneously entering an LS date of birth.

To account for cumulative losses due to attrition, an adjustment was made to the data after the 2001 Census link. This exercise has been repeated using the same methodology to make adjustments to the data after the 2011 Census link.

The adjustment model applies rates of loss at each census based on observed loss rates between the 1971 and 1981 Censuses. An underlying assumption of this approach is that enumeration at the 1981 Census was exceptionally high, so any existing LS members failing to appear at the 1981 Census were the subject of one of the three reasons for attrition detailed earlier.

All loss rates referred to in this section allow for both sample replenishment (births, immigrants and new members found at census) and observed exits (deaths and embarkations). At each point in time, separate loss rates broken down by sex and five year age group are estimated and applied for: those present at the previous census; those born since the previous census; and immigrants since the previous census.

The adjustment for 1981 to 1991 applied these loss rates to each of the three types of LS member to give an estimated number of LS members expected to be found at the 1991 Census.

The loss rates were similarly applied for 1991 to 2001. In this case, the rates for ‘those present at the previous census’ were applied to the estimated number present in 1991 rather than the observed number present.

The 1991 to 2001 adjustment was subsequently repeated for 2001 to 2011, again applying loss rates to the estimated number present in 2001 rather than the observed number.

This approach, illustrated in Figure 5, gave estimates of the numbers of LS members expected to appear at each of the 1991, 2001 and 2011 Censuses. The difference between these estimates and the observed numbers of LS members actually appearing provides an estimate of the level of under-enumeration at each census.
Figures 6 and 7 show the proportion of LS members not found at each census from 1981 to 2011 and the proportion after adjusting for attrition as outlined above. There are high loss rates for both males and females aged 65 to 69 in 1981 and for the same cohort aged 75 to 79 in 1991. This has been attributed to the pre-computerised processing of embarkations at and around retirement age and results in an over adjustment for 65 to 69 year olds at each subsequent census. The combination of this artefact and relatively small numbers at older ages makes it difficult to draw conclusions for those aged 65 and over.

Figures 8 and 9 show the hypothesised under-enumeration rate at the 2001 and 2011 Censuses based on the LS where those estimated to have been lost to attrition in the previous decade have been removed from both the sample expected at census and the total lost to follow up. This shows under-enumeration rates highest for under-tens and those in their twenties for males and females at both censuses. There was a slight improvement in 2011 compared to 2001 for males aged up to 29, and a more marked improvement for females up to age 40. Males aged 30 and over and females aged 40 and over show slightly higher rates of under-enumeration in 2011.

The results of this work were subsequently used in making a national adjustment to the 2011 Census estimates. Details of how they were used are included in a 2011 Census: Methods and Quality Report published in July 2012\(^3\).

---

\(^3\) 2011 Census: Methods and Quality Report, Making a national adjustment to the 2011 Census, July 2012
Figure 5: Modelling LS attrition at 1971/81 attrition rates

Observed Attrition 1971 - 1981

Census 71 (N_0)

Ex (n)

Census 81 (N_1)

C 81 Found

C 81 Attr (n_1)

70's Births (B_1)

Ex (n)

C 81 Found

C 81 Attr (n_1)

70's Immigrants (I_1)

Ex (n)

C 81 Found

C 81 Attr (n_1)

Exits:
Ex Observed Deaths and Embarks

Attrition rates:
Census AtrN r_1/(N_0-s_1)
Births AtrB n_1/(B_1+l_1)
Immigrants Atr I_1/(l_1-k_1)

Projected present at 81:
N_1 = N_0 + B_1 + I_1 - s_1 - l_1 - k_1 - r_1 + l_1 - h_1

Estimated Attrition 1961 - 1991

Census 81P (N_0)

Ex (n)

C 81 Attr (n_0)

50's Births (B_2)

Ex (n)

C 81 Found

C 81 Attr (n_1)

50's Immigrants (I_2)

Ex (n)

C 81 Found

C 81 Attr (n_1)

Exits:
Ex Observed Deaths and Embarks

Attrition rates:
Census AtrN r_0/(N_0-s_0)
Births AtrB n_0/(B_0+l_0)
Immigrants Atr I_0/(l_0-k_0)

Projected present at 81:
N_0 = N_0 + B_0 + I_0 - s_0 - l_0 - k_0 - r_0 + l_0 - h_0

Estimated Attrition 1991 - 2001

Census 91P (N_1)

Ex (n)

C 91 Attr (n_0)

2000's Births (B_3)

Ex (n)

C 91 Found

C 91 Attr (n_1)

2000's Immigrants (I_3)

Ex (n)

C 91 Found

C 91 Attr (n_1)

Exits:
Ex Observed Deaths and Embarks

Attrition rates:
Census AtrN r_0/(N_0-s_0)
Births AtrB n_0/(B_0+l_0)
Immigrants Atr I_0/(l_0-k_0)

Projected present at 91:
N_0 = N_0 + B_0 + I_0 - s_0 - l_0 - k_0 - r_0 + l_0 - h_0

Estimated Attrition 2001 - 2011

Census 01P (N_1)

Ex (n)

C 01 Attr (n_0)

2010's Births (B_4)

Ex (n)

C 01 Found

C 01 Attr (n_1)

2010's Immigrants (I_4)

Ex (n)

C 01 Found

C 01 Attr (n_1)

Exits:
Ex Observed Deaths and Embarks

Attrition rates:
Census AtrN r_0/(N_0-s_0)
Births AtrB n_0/(B_0+l_0)
Immigrants Atr I_0/(l_0-k_0)

Projected present at 01:
N_0 = N_0 + B_0 + I_0 - s_0 - l_0 - k_0 - r_0 + l_0 - h_0

Ex (l, k, d)

N * Exit rate from 2001 enumerations

Ex (l, k, d)

N * Exit rate from actual 2001 enumerations

Ex (l, k, d)

N * Exit rate from actual 2001 enumerations
Figure 6: Percentage of male LS members not found, 1981-2011, adjusted for cumulative loss from the LS in 1991, 2001 and 2011

Figure 7: Percentage of female LS members not found, 1981-2011, adjusted for cumulative loss from the LS in 1991, 2001 and 2011
Figure 8: Hypothesised male under-enumeration, 2001 and 2011

Figure 9: Hypothesised female under-enumeration, 2001 and 2011
Annex A: Composition of the LS sample after each census link

Tables A1 to A4 summarise the composition of the LS sample after each census link from 1981 to 2011.

Rows describe the individuals LS history:

**Traced sample:** LS members whose Census record was matched to their NHS record

**Untraced sample:** LS members whose Census record was not matched to their NHS record

**Found at [previous] Census:** LS members enumerated in the previous census

**Births:** LS members born on one of the four LS dates of birth in the period since the previous census (for example, 2000s births implies births between 30 April 2001 and 27 March 2011)

**Immigrants:** LS members registering with the NHS with one of the four LS dates of birth in the period since the previous census

**Re-entrants:** LS members who had embarked (notified the NHS that they were going abroad) prior to the previous census and had re-entered prior to the current census

**Found at [current] Census with no previous LS history:** LS members with a record from the most recent census and no previous census, birth or immigrant record

**Expected at [previous] Census but not found:** LS members who entered the sample prior to the previous census, who had not died or embarked at that point, but who were not found at that census

**Deaths prior to [previous] census:** LS members who had died prior to the previous census

**Embarkations prior to [previous] census:** LS members who embarked prior to the previous census and who have not returned prior to the current census

Columns describe the individual’s status at census:

**Found:** LS members found at the current census

**Not found:** LS members who had not died or embarked prior to Census Day and who were not found at the current census

**Died before [current] Census:** LS members who died on or before the date of the current census

**Embarked before [current] Census:** LS members who had embarked prior to the date of the current census and who had not re-entered or died. Where an individual had embarked without a re-entry but was enumerated at census their status was reassigned to ‘Found’.
Table A1: The LS sample in terms of tracing and linkage in 2011

<table>
<thead>
<tr>
<th>Expected at 2011 Census</th>
<th>Not expected at 2011 Census</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Died before 2011 Census</td>
<td>Embarked before 2011 Census</td>
</tr>
<tr>
<td>Traced sample</td>
<td>Found</td>
<td>Not found</td>
</tr>
<tr>
<td>Found at 2001 Census or (re-) entered since</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 2001 Census</td>
<td>421,716</td>
<td>60,521</td>
</tr>
<tr>
<td>Births</td>
<td>59,400</td>
<td>10,411</td>
</tr>
<tr>
<td>*61,184</td>
<td>*10,724</td>
<td>*380</td>
</tr>
<tr>
<td>Immigrants</td>
<td>31,252</td>
<td>40,589</td>
</tr>
<tr>
<td>*34,252</td>
<td>*43,548</td>
<td>*231</td>
</tr>
<tr>
<td>Re-entrants</td>
<td>624</td>
<td>386</td>
</tr>
<tr>
<td>New LS members</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 2011 Census</td>
<td>21,949</td>
<td></td>
</tr>
<tr>
<td>with no previous LS history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not found at 2001 Census</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected at 2001 Census but not found</td>
<td>44,865</td>
<td>108,071</td>
</tr>
<tr>
<td>Deaths prior to 2001 Census</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embarkations prior to 2001 Census</td>
<td>38</td>
<td>15,802</td>
</tr>
<tr>
<td>Untraced sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 2011 Census with no previous LS history</td>
<td>5,151</td>
<td></td>
</tr>
<tr>
<td>Found at 2001 Census</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected at 2001 Census but not found</td>
<td>29</td>
<td>15,840</td>
</tr>
<tr>
<td>Total</td>
<td>585,034</td>
<td>238,184</td>
</tr>
</tbody>
</table>

* Figure in italics show the estimated revised number of births, immigrants and new entrants on completion of event capture for the decade prior to the 2011 Census.
Table A2: The LS sample in terms of tracing and linkage in 2001

<table>
<thead>
<tr>
<th></th>
<th>Expected at 2001 Census</th>
<th>Not expected at 2001 Census</th>
<th>Died before 2001 Census</th>
<th>Embarked before 2001 Census</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Found</td>
<td>Not found</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traced sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 1991 Census or (re-) entered since</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 1991 Census</td>
<td>419,127</td>
<td>57,993</td>
<td>58,014</td>
<td>2,061</td>
<td><strong>537,195</strong></td>
</tr>
<tr>
<td>Births</td>
<td>59,836</td>
<td>9,944</td>
<td>467</td>
<td>301</td>
<td><strong>70,548</strong></td>
</tr>
<tr>
<td>Immigrants</td>
<td>14,948</td>
<td>28,886</td>
<td>212</td>
<td>1,835</td>
<td><strong>45,881</strong></td>
</tr>
<tr>
<td>Re-entrants</td>
<td>719</td>
<td>387</td>
<td>40</td>
<td>42</td>
<td><strong>1,188</strong></td>
</tr>
<tr>
<td>New LS members</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 2001 Census with no previous LS history</td>
<td>14,852</td>
<td></td>
<td></td>
<td></td>
<td><strong>14,852</strong></td>
</tr>
<tr>
<td>Not found at 1991 Census</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected at 1991 Census but not found</td>
<td>27,018</td>
<td>66,155</td>
<td>6,121</td>
<td>1,491</td>
<td><strong>100,785</strong></td>
</tr>
<tr>
<td>Deaths prior to 1991 Census</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>125,989</strong></td>
</tr>
<tr>
<td>Embarkations prior to 1991 Census</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>11,252</strong></td>
</tr>
<tr>
<td>Untraced sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 2001 Census with no previous LS history</td>
<td>3,417</td>
<td></td>
<td></td>
<td></td>
<td><strong>3,417</strong></td>
</tr>
<tr>
<td>Found at 1991 Census</td>
<td>67</td>
<td>6,613</td>
<td></td>
<td></td>
<td><strong>6,680</strong></td>
</tr>
<tr>
<td>Expected at 1991 Census but not found</td>
<td>17</td>
<td>9,274</td>
<td></td>
<td></td>
<td><strong>9,291</strong></td>
</tr>
<tr>
<td>Sub-total</td>
<td>539,996</td>
<td>179,252</td>
<td>190,883</td>
<td>16,947</td>
<td><strong>927,078</strong></td>
</tr>
<tr>
<td>Not entered by 2001 Census</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>172,429</strong></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>1,099,507</strong></td>
</tr>
</tbody>
</table>
Table A3: The LS sample in terms of tracing and linkage in 1991

<table>
<thead>
<tr>
<th>Expected at 1991 Census</th>
<th>Not expected at 1991 Census</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Died before 1991 Census</td>
<td>Embarked before 1991 Census</td>
<td></td>
</tr>
<tr>
<td>Found</td>
<td>Not found</td>
<td></td>
</tr>
</tbody>
</table>

**Traced sample**

- Found at 1981 Census or (re-) entered since
  - Found at 1981 Census: 422,207
  - Births: 65,707
  - Immigrants: 8,235
  - Re-entrants: 1,050

**New LS members**

- Found at 1991 Census with no previous LS history: 16,740

**Not found at 1981 Census**

- Expected at 1981 Census but not found: 21,075
- Deaths prior to 1981 Census: 61,083
- Embarkations prior to 1981 Census: 39

**Untraced sample**

- Found at 1991 Census with no previous LS history: 8,606
- Expected at 1981 Census but not found: 38

**Sub-total**

- 543,875
- 110,076
- 125,990
- 12,439
- 792,379

**Not entered by 1991 Census**

- 307,128

**Total**

- 1,099,507
Table A4: The LS sample in terms of tracing and linkage in 1981

<table>
<thead>
<tr>
<th>Expected at 1981 Census</th>
<th>Not expected at 1981 Census</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Died before 1981 Census</td>
<td>Embarked before 1981 Census</td>
<td></td>
</tr>
<tr>
<td>Found</td>
<td>Not found</td>
<td></td>
</tr>
<tr>
<td>Traced sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 1971 Census or entered since</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 1971 Census</td>
<td>417,040</td>
<td>39,817</td>
</tr>
<tr>
<td>Births</td>
<td>64,925</td>
<td>5,070</td>
</tr>
<tr>
<td>Immigrants</td>
<td>14,857</td>
<td>13,451</td>
</tr>
<tr>
<td>New LS members</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 1981 Census with no previous LS history</td>
<td>33,113</td>
<td></td>
</tr>
<tr>
<td>Untraced sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Found at 1981 Census with no previous LS history</td>
<td>4,955</td>
<td></td>
</tr>
<tr>
<td>Found at 1971 Census</td>
<td>1,247</td>
<td>6,809</td>
</tr>
<tr>
<td>Sub-total</td>
<td>536,137</td>
<td>65,147</td>
</tr>
<tr>
<td>Not entered by 1991 Census</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>