A population-based case-control study for examining early life influences on geographical variation in adult mortality in England and Wales using stomach cancer and stroke as examples

Ravi Maheswaran,a David P Strachan,b Brian Dodgeonc and Nicola G Bestd

Background  Geographical variation in mortality is influenced by factors operating in early life and in adulthood. The relative contributions of these factors may be examined by comparing the extent to which adult mortality is related to places of residence in early life and at death. We describe a population-based case-control design, in which all deaths are used as cases and the Office for National Statistics (ONS) Longitudinal Study (LS) survivors are used as controls.

Methods  Cases were all deaths from stomach cancer and stroke in England and Wales 1993–1995 amongst people born between January 1930 and September 1939 and for whom place of enumeration in 1939 could be imputed from the first three characters of their National Health Service number. Controls were all LS members born in the same period, enumerated in the 1991 census, resident in England and Wales in mid-1994 and for whom place of enumeration in 1939 could be similarly imputed. Logistic regression was used, adjusting for birth year, sex and social class. A previous mapping exercise by ONS generated comparable geographical units (counties) for 1939 enumeration and area of residence in 1991 or at death. ‘Non-migrant’ (i.e. 1939 ‘county’ the same as county in 1991 or at death) case:control ratios were calculated to indicate background mortality risk in counties, with adjustment for imprecision using Bayesian smoothing methods. These ratios were then used in modelling risk for inter-county migrants.

Results  There were 2590 stomach cancer and 7778 stroke deaths and 28 400 men and 28 180 women as controls. For men, 64%, 61% and 67% of stomach cancer deaths, stroke deaths and controls respectively could be assigned a county of enumeration in 1939. The corresponding percentages for women were 76%, 72% and 75%. For stomach cancer, after adjustment for county of enumeration in 1939, a significant association with the non-migrant case:control ratio for county of residence in 1991 or at death was observed ($P = 0.010$), indicating an association between current area of residence and stomach cancer mortality. There was no evidence of an independent effect of county of enumeration in 1939. For stroke, there was a highly significant trend in relation to 1939 county ($P = 0.0004$) and a less significant association with county of residence in 1991 or at death ($P = 0.016$).

---

a Public Health GIS Unit, School of Health and Related Research, The University of Sheffield, Sheffield S1 4DA, UK.
b Department of Public Health Sciences, St George’s Hospital Medical School, Cranmer Terrace, London SW17 0RE, UK.
c Centre for Longitudinal Studies, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL, UK.
d Department of Epidemiology and Public Health, Imperial College School of Medicine, St Mary’s Campus, Norfolk Place, London W2 1PG, UK.

Correspondence: Dr Ravi Maheswaran, Public Health GIS Unit, School of Health and Related Research, The University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK. E-mail: r.maheswaran@sheffield.ac.uk
There is substantial geographical variation in mortality from several diseases in England and Wales but the reasons for this are not clear. Elements of the recent debate on geographical variation in mortality include the role of socioeconomic deprivation, individual versus ecological factors and the relative importance of factors operating in early life, based on the Barker hypothesis, and in adult life. One method of examining the extent to which geographical variation in mortality is influenced by factors operating in early life and in adulthood is to examine the relative extents to which adult mortality is related to place of residence in early life and to place of residence at death.

Relatively few studies have examined the influence of early versus adult life factors on geographical variation in mortality, with mixed results. An ecological study found high correlations between recent age-standardized mortality rates from ischaemic heart disease, rheumatic heart disease and chronic bronchitis in 212 areas of England and Wales and infant mortality rates in the same areas in 1921–1925. However, the association with ischaemic heart disease mortality was substantially reduced after adjustment for current socioeconomic deprivation. A national study of proportional mortality by place of birth and place of death suggested a significant influence from birthplace for deaths from ischaemic heart disease, stroke and chronic bronchitis. A strong association between birthplace and stomach cancer has also been reported. In the British Regional Heart Study, however, the incidence of ischaemic heart disease was more closely related to place of examination in middle age than to birthplace. A study of inter-regional migrants in England and Wales using data from the Office for National Statistics (ONS) Longitudinal Study (LS) found that ischaemic heart disease and stroke mortality during 1971–1988 were related to both area of residence in 1939 and area of residence in 1971.

The study designs employed in the above studies include ecological studies, studies based on proportional mortality and cohort studies. The first two have inherent methodological limitations. Cohort studies are more rigorous but even in the LS, the largest cohort study in the UK (comprising 1% of the national population), the numbers of deaths from specific causes are not sufficiently high for detailed geographical analysis. In order to overcome this limitation, we have developed a population-based case-control study design, in which all deaths (i.e. a 100% sample) for whom place of residence in 1939 could be determined are used as ‘cases’ and LS survivors are used as ‘controls’. The method focuses on generations born in the 1930s and in this paper we demonstrate it on stomach cancer and stroke. We chose these diseases because there is substantial geographical variation and evidence for environmental influences, including early life influences, on risk in adult life.

We investigated the extent to which geographical variation in mortality from stomach cancer and stroke in England and Wales is related to place of residence in early life and place of residence at death.

Methods

The method we describe relies entirely on routinely collected data and had been approved by St Mary’s Hospital, London Local Research Ethics Committee. Mortality records were provided by ONS and the fields supplied included sex, individual’s own social class using the Registrar General’s classification, date of birth, date of death, county of residence at death, underlying cause of death and the first three characters of the National Health Service (NHS) number. From 1993 onwards ONS used automated search algorithms to link incoming death registrations with an NHS number held at the NHS Central Register. This is a unique alphanumeric identifier for each individual eligible for care under the British health service.

A substantial number of NHS numbers issued at the inception of the NHS in 1948 were derived directly from wartime food rationing numbers, which in turn were allocated at National Registration on 29 September 1939 at the start of the Second World War. The 1939 based NHS alphanumeric numbers were largely of the form ABCDnnn/n, the first three characters indicating the enumeration areas from which local government district in 1939 can be derived. During the coding of LS data, ONS carried out a mapping exercise that grouped local government districts from 1939 to post-1974 county boundaries. This generated comparable geographical units for 1939 enumeration and county of residence at death for cases or at the 1991 census for controls (defined below). Imputation of 1939 areas was not possible for individuals whose NHS number was not of the format characteristic of wartime identity numbers. The limitation of the method in relation to pre-war evacuation of children from major cities to the countryside is addressed in the Discussion.

Cases were defined as all deaths from stomach cancer (International Classification of Diseases Ninth Revision (ICD-9 151) and stroke (ICD-9 430–438) in England and Wales 1993–1995 amongst people born between 1 January 1930 and 29 September 1939 inclusive and for whom place of enumeration in 1939 could be imputed from their NHS number. The method therefore examines mortality in middle age (53–65 years). Social class for cases had been derived by ONS from information provided on death certificates by next of kin.

Controls were drawn from the ONS LS. This is a 1% representative sample of the population of England and Wales, linking data from Censuses from 1971 onwards with birth and
death registrations. We selected as controls, all LS members who were born in the same period as the cases, enumerated in the 1991 census, still alive and resident in England and Wales in mid-1994 and for whom place of enumeration in 1939 could be imputed from the NHS number. The LS controls’ own social class was based on their 1991 census occupation data. Whilst the LS contains information on place of residence in 1971 and 1981, comparable information is not available for deaths and these time points could not therefore be used. We restricted our focus to the generations born in the 1930s because unpublished information from the LS showed that there was a substantial drop in the proportion of NHS numbers of the format described in the generations born before 1930, especially for men, related to enlistment for the armed forces.

Statistical analysis

Each case group in turn was compared with the LS controls by multiple logistic regression modelling using the Generalized Linear Interactive Modelling (GLIM) package. Counts of cases and controls were generated for combinations of birth year (1930–1934, 1935–1939), sex (male, female), adult social class (I, II, III, III M, IV, V, unclassified), county of residence in 1991 or at death (54 categories) and county of enumeration in 1939 (54 categories). All models were adjusted for birth year, sex and social class. In addition, a score reflecting background mortality risk for each county was derived, based on the case-control ratio amongst non-migrants, and used in the logistic regression modelling (described below). The key stages in the analysis were as follows:

1. The birth-year-sex-social class-adjusted case-control ratio was derived for subjects whose 1939 county was the same as 1991 county or county of residence at death. We refer to these subjects as ‘non-migrants’ but it is important to note that there is no information on their place of residence between these two time points. This ratio in effect reflects the mortality rate amongst non-migrants born in 1930–1939 in each county as the numerator is all deaths in the population at risk and the denominator is a 1% sample of the population (person-years) at risk.

2. The natural logarithm of the adjusted non-migrant case-control ratio was used as a quantitative score to rank the counties in the analyses for inter-county migrants. The ratio was first linked to 1939 county in models adjusting for county in 1991 or at death as a categorical variable, and then linked to county in 1991 or at death in models adjusting for 1939 county as a categorical variable. This approach is an extension of methods used in earlier publications. It addresses the question as to whether mortality risk is higher among people who come from a higher risk area, implying a genetic, developmental or early environmental effect, or whether the risk is more closely related to current area of residence, implying an effect of more recent exposure.

3. Finally, in recognition of the relative imprecision of the non-migrant adjusted case-control ratios, Bayesian smoothing methods were applied to the logarithm of the ratios, before assigning these as scores to the 1939 county or to county in 1991 or at death. These smoothed estimates were obtained by ‘shrinking’ the logarithms of the adjusted case:control ratios for each county obtained in part (1) above towards a combination of the local (nearest neighbours based on contiguous counties) and global mean log case:control ratios across counties. The amount by which the estimate for each county is smoothed or shrunk is inversely proportional to the standard error of the (unsmoothed) log-adj usted mortality ratio for that county. Smoothing was carried out using random effects models similar to those used in conventional Bayesian disease mapping.

Results

There were 2590 stomach cancer deaths and 7778 stroke deaths in 1993–1995 in England and Wales amongst people born between 1 January 1930 and 29 September 1939. In the LS, 28 400 men and 28 180 women were born in the same time period, enumerated in the 1991 census and living in England and Wales in mid-1994 (Table 1).

Table 1 Cases (deaths from stomach cancer and stroke in 1993–1995) and controls (Longitudinal Study [LS] subjects enumerated in the 1991 census and alive in mid-1994), amongst people born between 1 January 1930 and 29 September 1939; England and Wales

<table>
<thead>
<tr>
<th>Social class</th>
<th>Deaths (Stomach cancer)</th>
<th>Controls (Longitudinal Study)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>I</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>206</td>
<td>94</td>
</tr>
<tr>
<td>III manual</td>
<td>471</td>
<td>33</td>
</tr>
<tr>
<td>IV</td>
<td>250</td>
<td>54</td>
</tr>
<tr>
<td>V</td>
<td>100</td>
<td>31</td>
</tr>
<tr>
<td>Uncoded</td>
<td>25</td>
<td>258</td>
</tr>
<tr>
<td>Total</td>
<td>1881</td>
<td>709</td>
</tr>
</tbody>
</table>

Valid 1999 county and county in 1991 or at death

(%) Total: 64 70 61 72 67 75

Year of birth

1930–1934: 787 348 1760 1670 9170 11230
1935–29 September 1939: 410 191 959 743 9718 9964

Social class

I: 44 1
II: 206 94
III manual: 471 33
IV: 250 54
V: 100 31
Uncoded: 25 258
Deaths (Stomach cancer) | Controls (Longitudinal Study)
Of the deaths, 61–64% of men and 72–76% of women could be assigned to a county of residence in 1939 using the method described. The percentages were similar or slightly higher (67% and 75%, respectively) for LS controls (Table 1). In order to display the unadjusted data in tabular form, we grouped the 54 counties into fifths of the distribution based on the case:control ratios of non-migrants with approximately equal numbers of controls in each fifth. We calculated the relative risk of mortality in each fifth for migrants and non-migrants. These are case:control ratios in each fifth relative to the overall case:control ratio for migrants and non-migrants combined.

For stomach cancer, the unadjusted relative risk rose from 0.6 in the lowest fifth to 1.59 in the highest fifth for non-migrants. For migrants classified by county in 1939, the unadjusted relative risk rose from 0.79 to 1.12 but the trend was inconsistent. For migrants classified by county in 1991 or at death, the unadjusted relative risk rose from 0.72 to 1.15 with a consistent rise across fifths of the distribution (Table 2).

For stroke, the unadjusted relative risk rose from 0.87 in the lowest fifth to 1.36 in the highest fifth for non-migrants. For migrants classified by county in 1939, the unadjusted relative risk rose from 0.78 to 1.00 but the trend was inconsistent. For migrants classified by county in 1991 or at death, the unadjusted relative risk rose from 0.74 to 1.01 but with an inconsistent trend (Table 3).

The ability of the method to detect early and later life effects will depend on the extent to which migrants moved to counties with higher or lower risk. For both stomach cancer and stroke, 74% of migrant controls moved to a country that was not in the same fifth of the distribution for risk. As an indication of the extremes, 16% of migrant controls in counties in the lowest fifth of the distribution for risk of stomach cancer moved to counties in the highest and second highest fifths of the distribution for risk. Conversely, 44% of people in counties in the highest fifth of the distribution for risk moved to counties in the lowest and second lowest fifths for risk. The percentages with regard to stroke mortality risk were 17% and 36% respectively (data not shown). The information indicates that there were sizeable population movements to counties at higher and lower risk.

Figures 1 and 2 show the unsmoothed and smoothed relative risks (relative case-control ratios) of mortality among non-migrants from stomach cancer and stroke, respectively, for the 54 counties in England and Wales, with Greater London as the reference category. The smoothed maps show the well-recognized North-South gradient for both conditions more clearly than the unsmoothed maps.

Table 4 shows the results of the $\chi^2$ tests of heterogeneity and trend by logistic regression modelling. For stomach cancer, after adjustment for birth year, sex and social class there was no significant heterogeneity across the 54 counties when classified by county in 1939 and county in 1991 or at death. When the counties were scored by the unsmoothed non-migrant case:control ratio there was no significant association between stomach cancer mortality and either 1939 county or county in 1991 or at death. However, when the smoothed case:control ratios were

### Table 2 Unadjusted effect of area of residence in 1939 and 1991 or at death on relative risk of mortality from stomach cancer, England and Wales

<table>
<thead>
<tr>
<th>Counties (grouped by relative risk in non-migrants)</th>
<th>Non-migrants</th>
<th>Migrants by area in 1939</th>
<th>Migrants by area in 1991 or at death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaths</td>
<td>Controls</td>
<td>Relative risk</td>
</tr>
<tr>
<td>Lowest fifth</td>
<td>120</td>
<td>4645</td>
<td>0.60</td>
</tr>
<tr>
<td>2nd fifth</td>
<td>177</td>
<td>4664</td>
<td>0.88</td>
</tr>
<tr>
<td>3rd fifth</td>
<td>203</td>
<td>4516</td>
<td>1.03</td>
</tr>
<tr>
<td>4th fifth</td>
<td>273</td>
<td>4611</td>
<td>1.36</td>
</tr>
<tr>
<td>Highest fifth</td>
<td>325</td>
<td>4727</td>
<td>1.59</td>
</tr>
<tr>
<td>Total</td>
<td>1098</td>
<td>23163</td>
<td>1.05</td>
</tr>
</tbody>
</table>

The relative risks are case:control ratios, relative to the overall case:control ratio of 1736:40 084 for migrants and non-migrants combined. Migrants are defined as people whose counties of residence in 1939 and 1991 or at death are different.

### Table 3 Unadjusted effect of area of residence in 1939 and 1991 or at death on relative risk of mortality from stroke, England and Wales

<table>
<thead>
<tr>
<th>Counties (grouped by relative risk in non-migrants)</th>
<th>Non-migrants</th>
<th>Migrants by area in 1939</th>
<th>Migrants by area in 1991 or at death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaths</td>
<td>Controls</td>
<td>Relative risk</td>
</tr>
<tr>
<td>Lowest fifth</td>
<td>461</td>
<td>4153</td>
<td>0.87</td>
</tr>
<tr>
<td>2nd fifth</td>
<td>556</td>
<td>4725</td>
<td>0.92</td>
</tr>
<tr>
<td>3rd fifth</td>
<td>729</td>
<td>5136</td>
<td>1.11</td>
</tr>
<tr>
<td>4th fifth</td>
<td>630</td>
<td>4089</td>
<td>1.20</td>
</tr>
<tr>
<td>Highest fifth</td>
<td>878</td>
<td>5060</td>
<td>1.36</td>
</tr>
<tr>
<td>Total</td>
<td>3254</td>
<td>23163</td>
<td>1.07</td>
</tr>
</tbody>
</table>

The relative risks are case:control ratios, relative to the overall case:control ratio of 5132:40 084 for migrants and non-migrants combined. Migrants are defined as people whose counties of residence in 1939 and 1991 or at death are different.
used, a significant association \( (P = 0.010) \) emerged with county in 1991 or at death. This provides evidence of an association between current area of residence and stomach cancer mortality, independent of birth year, sex, adult social class and county of enumeration in 1939. There was no evidence of an independent effect of county of enumeration in 1939.

For stroke mortality, there was evidence of statistically significant heterogeneity in relation to 1939 county, which remained significant as a trend in relation to the unsmoothed non-migrant case:control ratio, and highly significant \( (P = 0.0004) \) when the smoothed case:control ratios were used. There was a less significant association \( (P = 0.016) \) with county in 1991 or at death when smoothed non-migrant case:control ratios were modelled. Thus, for stroke there was evidence of independent associations of mortality risk with both early life and later effects of places of residence.

### Table 4  
Chi-square tests of heterogeneity and trend from logistic regression models, adjusted for birth-year, sex, social class and county in 1939 or county in 1991 or at death

<table>
<thead>
<tr>
<th></th>
<th>County in 1939</th>
<th>County in 1991 or at death</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stomach cancer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity across counties (53 d.f.)</td>
<td>52.40</td>
<td>56.49</td>
</tr>
<tr>
<td>Trend by unsmoothed log relative risk in non-migrants (1 d.f.)</td>
<td>0.39</td>
<td>1.72</td>
</tr>
<tr>
<td>Trend by smoothed log relative risk in non-migrants (1 d.f.)</td>
<td>2.65</td>
<td>6.66**</td>
</tr>
<tr>
<td>Adjusted relative risk for migrants versus non-migrants</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td><strong>Stroke</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity across counties (53 d.f.)</td>
<td>72.60*</td>
<td>68.30</td>
</tr>
<tr>
<td>Trend by unsmoothed log relative risk in non-migrants (1 d.f.)</td>
<td>5.97*</td>
<td>2.57</td>
</tr>
<tr>
<td>Trend by smoothed log relative risk in non-migrants (1 d.f.)</td>
<td>12.64***</td>
<td>5.82*</td>
</tr>
<tr>
<td>Adjusted relative risk for migrants versus non-migrants</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

* \( P < 0.05 \), ** \( P < 0.01 \), *** \( P < 0.001 \).
Discussion

We developed the population-based case-control approach to overcome some of the limitations of existing methods in relation to examining geographical variation in mortality risk. A key benefit of this design is the use of 100% of national mortality data that greatly increases its power. If the LS had been used as a straightforward cohort it would have contained only 1% of these deaths, substantially reducing the power to detect significant patterns and associations.

The estimates of risk for individual counties amongst non-migrants were based on relatively small numbers. In order to address this limitation, we employed Bayesian smoothing methods to improve the estimates of non-migrant mortality risk at the county level. The benefit of this approach is reflected in the larger $\chi^2$ values in tests for trends for both stomach cancer and stroke mortality. It is useful to note, however, that there are a number of methodological and statistical issues related to disease mapping and these are discussed in detail in recent publications.21,22

It could be argued that this approach is overly complicated and a simpler method would be to examine mortality in relation to place of birth and place of death where these are recorded on death certificates, as has been done by others using proportional mortality.7 An important advantage of our case-control approach is that controls in the LS are a representative sample of the total population denominator from which the cases arose, and are therefore likely to reliably represent the exposure distribution in the source population. The patterns observed with our case-control approach would not be distorted by differences in death rates from other common causes, a recognized limitation of the proportional mortality approach. A further potential problem is the accuracy of place of birth information from death certificates, which is provided by an informant many years after birth.

There are a number of potential limitations to the method we have described in this paper. The first two relate to the linkage of NHS numbers to death and LS records. Approximately 30% of records could not be linked to informative NHS numbers with respect to determination of county of residence in 1939. This is because some NHS numbers allocated in 1939 did not have a format that could be used to deduce place of enumeration, or that numbers were reallocated for people who entered national service after the war. It is possible that a selection process operated and that the 30% whose NHS numbers could not be linked to place of enumeration were different from the rest, both with respect to place of birth and with respect to risk of disease in later life. Although it seems unlikely that there were systematic differences between cases and controls with respect to linkage, this potential bias cannot be ruled out.

The second and potentially more serious limitation concerns population movements between birth and national registration in 1939. These movements may have been due to slum

![Unsmoothed and smoothed relative risk of mortality from stroke in 1993−1995 among non-migrants born between 1 January 1930 and 29 September 1939 in 54 counties in England and Wales, with Greater London as the reference category](image-url)
clearances, sub-urban development or the evacuation of children before the outbreak of war. There were large-scale evacuations of schoolchildren from major cities to rural areas in the first week of September 1939 in anticipation of bombing raids. The place of enumeration of these people would not be the same as place of birth and therefore may not accurately reflect their early life environment. The magnitude of population movements was assessed by comparing the percentage distribution of LS controls born in 1930–1939 by county of enumeration in 1939 with the percentage distribution of all births in England and Wales in an equivalent time period. The problem was most marked in Inner London which contained 3.4% of the LS controls but 9.7% of births. However, it was not overwhelming —the LS population enumerated in rural districts was inflated by one-quarter while the equivalent excess in municipal and urban districts was one-seventh relative to the distribution of all births.

A further potential limitation is the extent to which county is a good proxy for early and later life factors. These factors could be differently related to county, with county being a good indicator for one set of factors but not the other. The degree of variation within and between counties of putative early and later life factors therefore needs to be considered when interpreting results produced by the method described in this paper.

The extent to which county is a proxy for early and later life factors is also dependent on the age at which migration occurred for this birth cohort. For example, if migration occurred relatively late in life, place of residence in 1939 would represent a mixture of early and later life factors and would therefore be an inadequate discriminator for factors operating in these two periods of the life course. Whilst we do not have information directly from this study regarding the age at migration, an indication may be obtained from a study of migration between counties among LS members born in the 1930s with information on county of residence at three time points (1939, 1971 and 1991). Over half the cohort (22,085 individuals or 56%) were coded to the same county at all three time points. Of the 17,324 inter-county ‘migrants’, 87% had moved between 1939 and 1971 i.e. by the age of 32 to 41 years. This suggests that county does provide some discrimination between early and later life factors for this birth cohort.

We adjusted for social class as this is an important potential confounder. Information on social class was obtained from different sources for cases and controls, potentially leading to the problem referred to as numerator-denominator bias. This may have blunted the association between social class and mortality, thereby resulting in residual confounding. However, the numerator-denominator mismatch may be less of a problem in people of working age (our cases were 53–65 years at death).

We studied mortality in a middle-aged population and inaccuracies in death certification for both conditions are less likely to be a problem than in older age groups. Mortality is determined by both incidence and case-fatality so some caution is required in extrapolating to incidence. It should also be noted that we restricted the focus of the method to generations born in the 1930s and this cohort was exposed to quite unusual circumstances (economic depression, wartime) in their early life course.

The method described required an estimation of background risk, which is obtained from non-migrants. The smallest geographical scale at which this could be calculated satisfactorily was the county level. The method could be adapted to examine specific exposures or indicators of exposure (e.g. air pollution, stillbirth rate) and for these, finer spatial scales could be used providing data are available at these levels. However, the level of spatial analysis is restricted by confidentiality constraints on the LS.

Our results for cancer of the stomach suggest that there was an association with county of residence in later life but no evidence of an independent effect of county of residence in early life. The latter seems to be at odds with other studies which found clear associations between place of birth, or other indicators of the early life environment, and stomach cancer mortality. Evidence suggests that infection with Helicobacter pylori in childhood may be a possible explanation for the early life effect. There are a number of possible reasons for the discrepancy. We may not have had sufficient power to detect an association. The dilution effect of population movements may have been sufficient to mask an association. Other hidden biases and uncontrolled negative confounding may be other possible reasons. However, it is worth noting that we studied a relatively recent birth cohort (born in 1930–1939) and examined deaths among this cohort in middle age (53–65 years). There has been a substantial decline in mortality from stomach cancer over recent decades and it is possible that early life factors now play a less important part in explaining geographical variation in stomach cancer mortality relative to later life factors such as diet and salt intake in adulthood. It is also possible that the relative contributions of early and later life factors, and changes in the relative contributions over time, vary by place. It is worth noting that Coggon et al. did find a significant association with place of death as well as place of birth, albeit very small in magnitude. In addition, as they examined deaths from all ages in 1969–1972, they effectively amalgamated information from a wide span of birth cohorts, potentially masking changes in cohort effects.

With regard to stroke, a large number of studies have suggested that factors operating in early life may be important in predicting risk of stroke, and of risk factors for stroke, in adult life. Evidence implicates factors operating in the intra-uterine environment and in childhood, though it is not clear which period is more important in determining risk. In addition, an age-cohort analysis indicated that there is evidence of a cohort effect on stroke mortality at the population level. However, it is also widely recognised that a number of factors operate in adult life to influence stroke risk, including lifestyle factors such as smoking and alcohol consumption and a range of medical interventions. In addition, there is a strong link between concurrent socioeconomic deprivation at the small area level and stroke mortality, which diminishes with increasing age. We found strong evidence of an independent association of stroke mortality risk with place of residence in early life and to a lesser extent with place of residence in later life, consistent with the above literature.

In summary, we have developed a population-based case-control study design for examining early life influences on geographical variation in adult mortality in England and Wales using routinely collected data and demonstrated the method on mortality from stomach cancer and stroke. We plan to extend this work to investigate specific characteristics of areas of
enumeration in 1939 in relation to subsequent risk of mortality from a range of diseases.

Acknowledgements
We would like to thank ONS for allowing use of the ONS Longitudinal Study and members of the LS User Support Programme at the Centre for Longitudinal Studies (CLS), Institute of Education for assistance with accessing the data. This work uses data provided with the support of the ESRC and JISC, and census and boundary material, which are copyright of the Crown, the Post Office and the ED-LINE Consortium. A grant from the Joint Standing Research Committee, St Mary's Hospital, London was used to obtain mortality data. NHS Executive Trent provides core funding for the Public Health GIS Unit. The views expressed in this publication are those of the authors and not necessarily those of ONS, CLS or the funding organizations.

KEY MESSAGES
- We have developed a population-based case-control study design for examining early life influences on geographical variation in adult mortality in England and Wales using routinely collected data and demonstrated the method on mortality from stomach cancer and stroke.
- The method detected an association between place of residence in early life as well as current area of residence and stroke and an independent association between current area of residence, but not place of residence in early life, and stomach cancer.
- The method will provide a useful new way of assessing the relative contributions of factors operating in early life and in adulthood on geographical variation in adult mortality in England and Wales but is restricted in focus to generations born in the 1930s.

References
6 Ben-Shlomo Y, Davey Smith G. Deprivation in infancy or in adult life: which is more important for mortality risk? Lancet 1991;337:530–34.